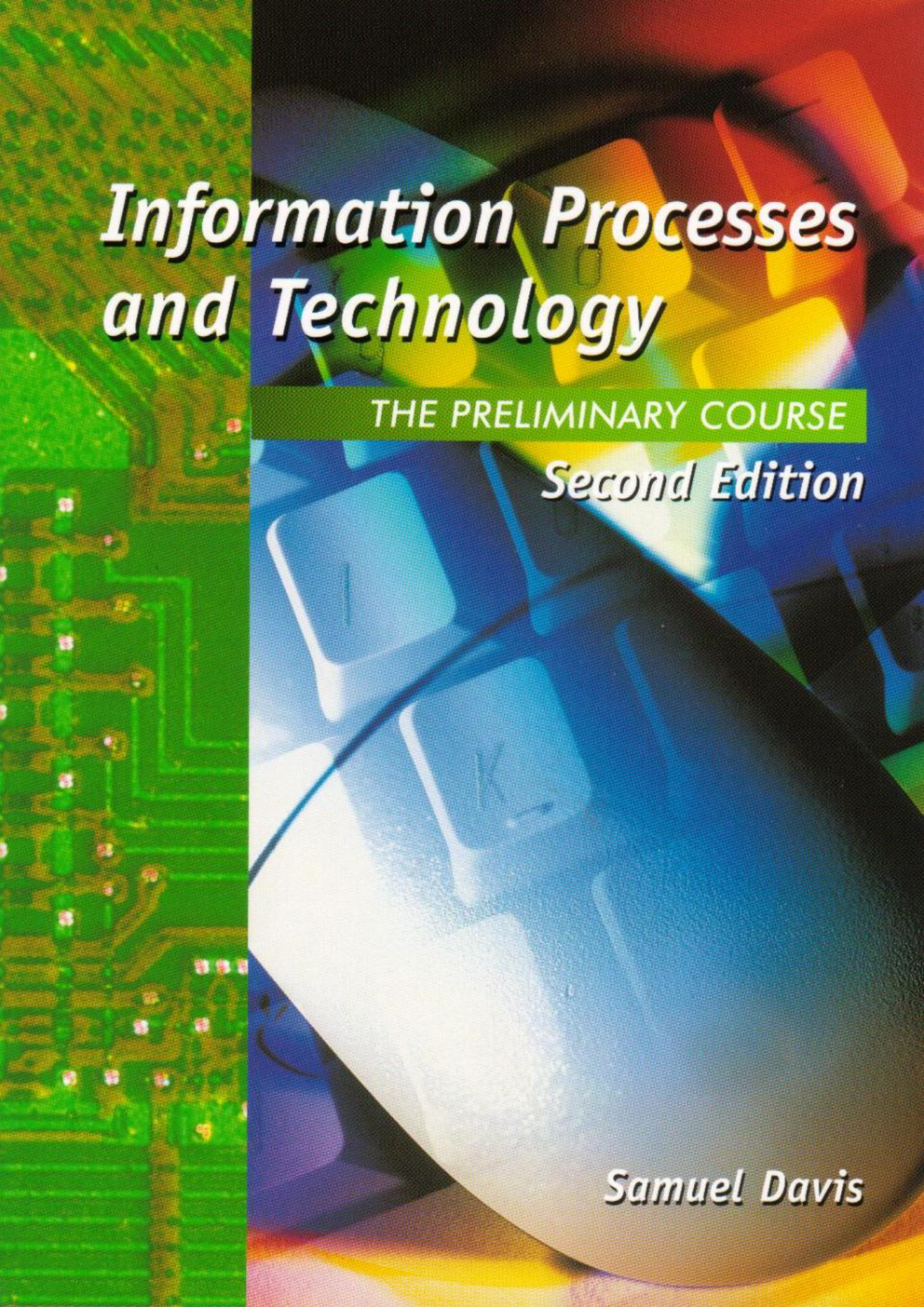




Information Processes and Technology

THE PRELIMINARY COURSE

Second Edition

Samuel Davis

First published 2009 by
Parramatta Education Centre
Tel: (02) 4632 7987 Fax: (02) 4632 8002

Visit our website at www.pedc.com.au

Copyright © Samuel Davis 2009

All rights reserved.

Except under the conditions described in the Australian Copyright Act 1968 (the Act) and subsequent amendments, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the copyright owner.

National Library of Australia
Cataloguing in publication data

Davis, Samuel, 1964-.
Information processes and technology: the preliminary course (second edition).

Includes index.
ISBN 9780957891081.

1. Information storage and retrieval systems. 2. Electronic data processing.
3. Information technology. I. Title

004

Reviewer: Stephanie Schwarz
Cover design: Great Minds
Printed in Australia by Ligare Pty. Ltd.

CONTENTS

ACKNOWLEDGEMENTS	vi
TO THE TEACHER	vii
TO THE STUDENT	vii

INTRODUCTION TO INFORMATION SKILLS AND SYSTEMS

1. INTRODUCTION TO INFORMATION SYSTEMS	3
What is a system?	3
What is information?	5
Information systems in context	6
Set 1A	16
Social and ethical issues	17
Set 1B	30
Chapter 1 review	32
2. INTRODUCTION TO INFORMATION PROCESSES AND DATA	35
Relationships between information processes	35
Collecting	38
Set 2A	42
Organising	44
Analysing	45
Storing and retrieving	46
Processing	47
Transmitting and receiving	48
Displaying	50
Set 2B	52
The nature of data and information	54
Set 2C	65
Digital representation of data	67
Chapter 2 review	78

TOOLS FOR INFORMATION PROCESSES

3. COLLECTING	81
Hardware used for collection	82
Set 3A	91
Set 3B	102
Software used for collection	103
Set 3C	111
Non-computer procedures in collecting	113
Social and ethical issues in collecting	118
Set 3D	125
Chapter 3 review	127

4. ORGANISING	129
The effect of organisation on software applications	131
Set 4A	139
Set 4B	153
Non-computer tools for organising	160
Social and ethical issues associated with organising	161
Set 4C	166
Chapter 4 review	167
5. ANALYSING	169
Hardware requirements for analysing	170
Software features for analysing	174
Set 5A	178
Non-computer tools for analysing	185
Set 5B	189
Social and ethical issues associated with analysing	190
Chapter 5 review	195
6. STORING AND RETRIEVING	197
The role of storing and retrieving	198
Hardware in storing and retrieving	202
Set 6A	204
Operation of secondary storage hardware	206
Set 6B	224
Software in storing and retrieving	226
Set 6C	237
Non-computer tools for storing and retrieving	239
Social and ethical issues associated with storing and retrieving	241
Chapter 6 review	244
7. PROCESSING	247
The integration of processing and other information processes	247
Hardware in processing	250
Set 7A	260
Software applications for processing	262
Non-computer tools for documenting processing	267
Set 7B	272
Social and ethical issues associated with processing	274
Chapter 7 review	277

8. TRANSMITTING AND RECEIVING	279
Communication concepts	280
Set 8A	290
Hardware in transmitting and receiving	291
Set 8B	304
Set 8C	316
Software for transmitting and receiving	317
Non-computer tools for transmitting and receiving	328
Set 8D	329
Social and ethical issues associated with transmitting and receiving	330
Chapter 8 review	337
 9. DISPLAYING	 339
Hardware for displaying	340
Set 9A	356
Software for displaying	358
Set 9B	369
Non-computer tools for displaying	371
Social and ethical issues associated with displays	373
Set 9C	380
Chapter 9 review	381
 DEVELOPING INFORMATION SYSTEMS	
 10. DEVELOPING INFORMATION SYSTEMS	 383
Introduction to system development	383
Traditional stages in developing a system	384
Understanding the problem	386
Planning	388
Set 10A	393
Designing	394
Implementing	400
Testing, evaluating and maintaining	405
Set 10B	410
Social and ethical issues	412
Chapter 10 review	416
 GLOSSARY	 417
INDEX	430

ACKNOWLEDGEMENTS

Writing a text such as this is an enormous task. Technology changes continuously, indeed some technologies that were in common use when the first edition was published in 2004 have now (2009) been completely replaced with new technologies. Fortunately, there are numerous professionals across the globe who were more than willing to assist with valuable information and feedback on these new technologies – without the Internet, email and newsgroups such information would be simply impossible to obtain and verify. There are too many of these people to mention by name, but thank you, your knowledge and comments has greatly enhanced the accuracy of the text.

The original text was reviewed by Stephanie Schwarz. Stephanie's review comments, as expected, were always accurate, insightful and right on target. She has an uncanny ability to express relatively complex ideas in a succinct, yet understandable manner. Stephanie and I worked together as senior HSC marker for numerous years. We've worked together at the NSW Board of Studies and on numerous Trial HSC Exams. Many of the HSC Style Questions included within the text were originally published within past trial exams. Stephanie's knowledge and enthusiasm for technology and education is legendary.

My wife Janine, together with my children Louise, Melissa, Kim and Luke have all made sacrifices so I can disappear to research and write. Janine provided much of the motivation to continue with such a detailed text, she also completed the final editing – thanks darling! My young son Luke has spent much of his life with a father sitting at a computer. Young man, you've got your Dad back!

Thanks also to the many companies and individuals who willingly assisted with the provision of screen shots, articles and other copyrighted material. Every effort has been made to contact and trace the original source of copyright material in this book. I would be pleased to hear from copyright holders to rectify any errors or omissions.

Samuel Davis

TO THE TEACHER

This book provides a thorough and detailed coverage of the revised NSW *Information Processes and Technology* Preliminary course. The revised syllabus was first examined at the 2009 Higher School Certificate. The text is written to closely reflect the syllabus, both in terms of content but also in terms of intent. In my view, the IPT syllabus is written in such a way that is relevant to students with a broad range of abilities. The best students will want to know the detail of how and why; this text includes such detail.

The text closely follows the syllabus and apart from Chapter 10, each builds on and refers to concepts introduced in previous chapters. The content of the text (and also the syllabus) is arranged around the seven, somewhat arbitrary syllabus information processes. To my mind this arrangement makes logical sense in terms of focussing on processes rather than hardware and software. However, it can mean that students place too much emphasis on arbitrarily splitting systems into these seven processes. Throughout the text I have endeavoured to downplay this tendency by considering the real and interrelated nature of information processes that occur in real systems. I applaud those who revised the syllabus for including specific content on the integration of processes. I specifically address this content at the start of Chapter 2 and again at the start of Chapter 7.

Numerous Group Tasks are included throughout the text. These tasks aim to build on both the theoretical and practical aspects of the course. A teacher resource kit is available that provides further details and discussion points for each of these tasks. The teacher resource kit also includes fully worked solutions for all sets of questions, blackline masters and a CD-ROM containing a variety of other relevant resources.

TO THE STUDENT

Information systems are all around us and we use them to meet our needs every day. In fact, meeting the needs of people is ultimately the purpose of all information systems. The *Information Processes and Technology* Preliminary Course focuses on the underlying processes occurring within information systems. These processes or actions are performed by computer hardware and software, together with people. IPT is not about learning how to use software applications – although you will develop some skills in this area. IPT is more about learning how and why things operate and how different components and processes can be combined to solve problems. It's a course about systems that process data in to information; information systems!

The Preliminary course will provide you with a thorough grounding in regard to the operation and design of information systems. In the HSC course you will apply this knowledge as you examine and develop particular types of information systems such as communication systems and database systems.

Best wishes with your senior studies, and in particular with your *Information Processes and Technology* studies. Hopefully this text will provide worthwhile assistance in this regard.

In this chapter you will learn to:

- diagrammatically represent a given scenario that involves an information system
- explain how an information system impacts on its environment and how it in turn impacts on the information system
- describe the environment and purpose of an information system for a given context
- explain how a given need can be supported by an information system
- describe an information system in terms of its purpose
- for a given scenario, identify the people who are:
 - in the environment
 - users of the information system
 - participants in the information system
- describe social and ethical issues that relate to:
 - information system users
 - participants
- ensure that relevant social and ethical issues are addressed
- identify and explain reasons for the expansion of information systems, including:
 - advances in technology
 - suitability of information technology for repetitive tasks

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies.

In this chapter you will learn about:

Information systems in context

- diagrammatic representation of an information system in context
- the environment – everything that influences and is influenced by the information system
- the purpose – a statement identifying who the information system is for and what it needs to achieve
- who the information system is for includes individuals and organisations
- the information system – a set of information processes requiring participants, data/information and information technology built to satisfy a purpose
- information processes – computer based and non-computer based activities
- information technology – hardware and software used in information processes
- data – the raw material used by information processes
- information – the output displayed by an information system
- user – a person who views or uses the information output from an information system
- participant – a special class of user who carries out the information processes within an information system

Social and ethical issues

- social and ethical issues arising from the processing of information, including:
 - privacy of the individual
 - security of data and information
 - accuracy of data and information
 - data quality
 - changing nature of work
 - appropriate information use
 - health and safety
 - copyright laws
- the people affected by social and ethical issues, including:
 - participants within the information system
 - users of the information system
 - those in the environment
- the ethical and social responsibility of developers
- current government legislation to protect the individual and organisations
- the use of information systems in fields such as manufacturing as well as the traditional fields of observation and recording
- global information systems:
 - where the purpose involves international organisations, or
 - where the data and processes are distributed across national boundaries

INTRODUCTION TO INFORMATION SYSTEMS

What is an information system? The answer to this question is the central aim of this chapter. To understand information systems let us first consider the broader questions of ‘What is a system?’ and ‘What is information?’

WHAT IS A SYSTEM?

A system is a collection of resources and processes that operate together to achieve some common purpose and hence fulfil some need. For example, the braking system in a car fulfils the need to slow down the car. Its purpose, or reason for existence, is to slow down the car. To achieve this purpose requires resources or components such as the brake pedal, brake pads, brake disks, together with tyres and many other components. Even the driver is an essential component of the braking system. These components or resources must work together to successfully slow down the car. The ways in which they interact are known as the processes of the system. Processes are actions that when systematically followed will cause the resources to achieve the specified purpose. In our braking system example the driver applies pressure to the brake pedal, which in turn causes fluid to move from the master cylinder to a calliper on each wheel. At each wheel calliper the fluid pressure causes the brake pads to push against the brake disk causing friction and hence slowing down the wheels rotation. Because the tyres are gripping the road surface the reduction in rotation speed also slows down the road speed.

Almost all systems are themselves made up of smaller sub-systems and similarly almost all systems are part of larger systems. Everything that influences or is influenced by the system is said to be in the environment. In our braking system the complete car is a larger system that has the braking system as one of its sub-systems. Most of these other sub-systems affect or are affected by the braking system and hence are in its environment. For example, the braking system interfaces with the electrical system via a switch that turns the brake lights on or off as the brakes are activated or deactivated. Each of the component parts of the braking system can themselves be seen as a system, for example the master cylinder. Even within the master cylinder there are a number of sub-systems that each achieves a specific purpose within the larger master cylinder system.



System

Any organised assembly of resources and processes united and regulated by interaction or interdependence to accomplish a common purpose.

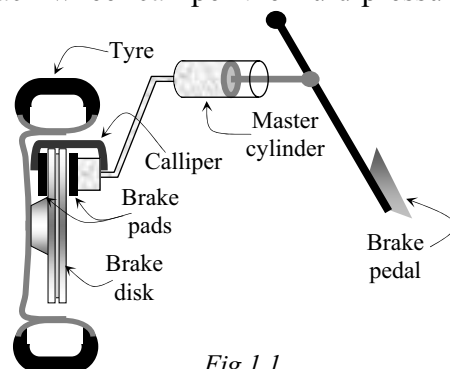


Fig 1.1

The braking system is a sub-system of the car and is also made up of sub-systems.

DIAGRAMMATIC REPRESENTATION OF A SYSTEM

System engineers from all fields use diagrams and models to describe systems. Different types of diagrams are used to describe different aspects of the system. The diagram at right *Fig 1.2*, describes an overview of the resources and processes of a system, together with its purpose and environment. The arrows on the diagram show that the resources are used by the processes and in turn these processes work to achieve the system's purpose.

There are many different methods for representing systems diagrammatically, including context diagrams, data flow diagrams, flowcharts and IPO charts. Context diagrams are used to model the data movements to and from the system and its environment. Data flow diagrams model the data movements within the system. Flowcharts describe the logic of the system's processes. IPO charts identify how specific inputs are transformed into outputs. Throughout the IPT course we shall learn to use a variety of these techniques.



Consider the following:

A backyard swimming pool contains a filtration system that includes a timer, a pump, a filter, various pipes and electrical connections and a skimmer box. These components work together to keep the pool water clean and healthy. *Fig 1.3* shows many of these components, together with the flow of water.

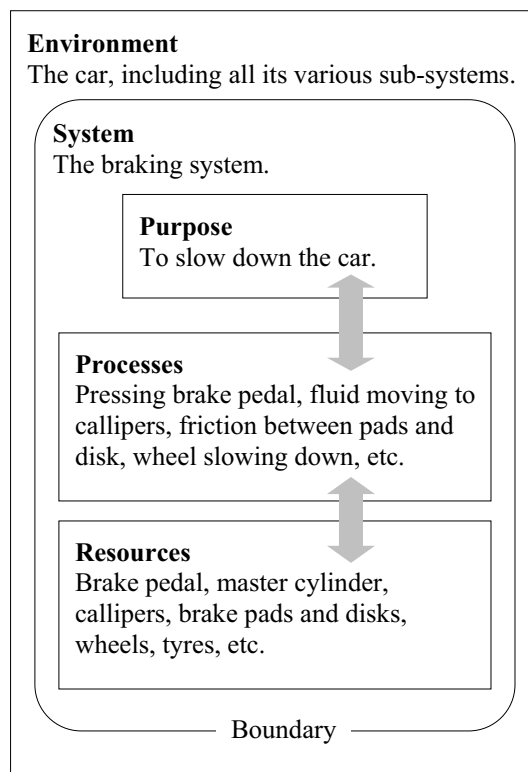


Fig 1.2
Diagrammatic representation of the
braking system on a car.

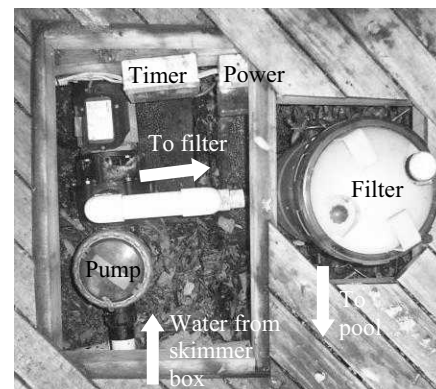


Fig 1.3
Pool filtration system.



GROUP TASK Discussion

What is the purpose of this filtration system? What are the resources and processes of the system? Describe this systems environment and how it achieves its purpose within this environment?



GROUP TASK Activity

Draw a diagram, like the one shown in *Fig1.2* above, to model the swimming pool filtration system.

WHAT IS INFORMATION?

The word ‘information’ appears to be the catchword of the century. Apparently we are living in the information age. Information is supposed to help us all and the more we have the more enlightened and fulfilled our lives are supposed to be. There are even charitable organisations devoted to making information more accessible to those in third world countries. Information is traded as a commodity, like oil or even gold. The Internet is often referred to as the information super highway. So what is this stuff called information?

Information leads to knowledge and knowledge is acquired by being aware of and understanding the facts. The facts or data must be processed into information before humans can use the data to obtain knowledge. We may have access to a large store of facts or data but it is not until these facts are understood and their meaning derived that we have information.

This is really the primary aim of this course, to examine the processes and technologies used to turn raw facts or data into meaningful information. We must be careful with our understanding of facts in this context, the information resulting from the data will only be correct if indeed the data is factual. The cliché ‘garbage in – garbage out’ holds true, if the data is rubbish then the resulting information will also be rubbish.

Information is therefore the output displayed by an information system that we, as human users, use to acquire knowledge. When we receive information concerning some fact or circumstance we interpret the information to acquire knowledge. For example, ‘123456.65’ is data; ‘your savings account balance is \$123,456.65’ is information; whereas ‘I’ve got enough money to buy that Ferrari’ is knowledge.



Fig 1.4

Data is transformed into information using information processes and technology.



Information

Information is the output displayed by an information system. Knowledge is acquired when information is received.



Consider the following list of data items:

- All the HSC results for a given year.
- The daily rainfall over the last ten years in your area.
- The number of cars passing your school each minute.
- Details on each take-off and landing at Mascot airport.



GROUP TASK Activity

List at least 2 types of information that may be derived from each of the above sets of data.



GROUP TASK Discussion

Discuss how humans may use the above information to acquire knowledge.

INFORMATION SYSTEMS IN CONTEXT

An information system is a system whose primary purpose is to process data into information. The data is collected, processed using various resources of the system and finally the resulting information is output.

In this section we examine the general nature of information systems including:

- the environment
- the boundary
- the purpose
- information processes and
- resources.

The resources used by all information systems includes the participants, the data and information, together with all the various forms of information technology. As computers are particularly suited to data processing tasks, it is common for the information technologies used to include computer hardware and software.

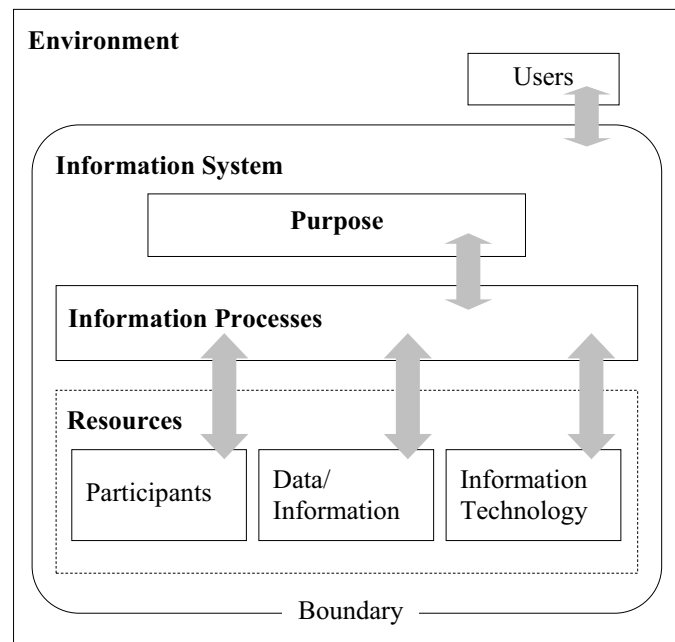


Fig 1.5
Diagrammatic representation of an information system.

ENVIRONMENT

The environment in which an information system operates is everything that influences, and is influenced by, the information system but is not part of the information system. It encompasses all the conditions, components and circumstances that surround the system. This includes those users who do not directly interact or perform processes within the system. That is, users who are not participants are part of the environment. The information system may collect data from and display information to these indirect users, however they do not participate in the information system's operation.



Environment

The circumstances and conditions that surround an information system. Everything that influences or is influenced by the system.

The word environment is often used in terms of the natural environment in which we or some plant or animal live. The natural environment contains many complex and interrelated systems that are so intricate that we can never hope to understand or control them in their entirety. The environment for most information systems is less complex yet in most cases it contains many aspects that cannot be controlled or even predicted. For example, many information systems require network access; however the network is commonly part of the information systems environment. Hence the system must know how to communicate using the network but correcting faults within the network is beyond the scope of the information system. Information systems must aim to minimise any environmental effects that could hinder the system as it operates to achieve its purpose.

BOUNDARY

The boundary defines what is part of the information system and what is part of the environment. It is the delineation between the system and its environment. For example, an online ordering system designed to process orders for a business may use the services of a payments system to process and approve credit card payments. The payments system is in the environment of the online ordering system, however the ordering system must be able to interface with the payments system but cannot affect how payments are processed by the payments system.



Boundary

The delineation between a system and its environment. The boundary defines what is part of the system and what is part of the environment.

When developing new information systems it is critical to define the boundaries of the system as clearly as possible so that all parties understand what a new system will do and often more importantly what it will not do. All the processes and resources that will form part of the new system are said to be within the scope of the system. If there is likely to be confusion about whether some process or resource is or is not included then a specific statement should be included to remove any doubts. Consider a new online ordering system. The system scope may include collecting order details from customers and storing them in a database. It is reasonable that a client may expect the system to include approval of credit card payments as payment approval is closely related to ordering. The developer would be wise to clearly state whether payment approval is or is not within the systems scope.



Consider the following:

Automatic Teller Machines (ATMs) are now common items in every bank, shopping centre and even in most service stations. An ATM is an information system, its primary purpose is to process data into information. Account details, PINs and transaction details are entered by the user and result in a combination of outputs in the form of cash, receipts and information displayed on the monitor. These processes occur within an environment that cannot be fully controlled by the ATM system.

Let us consider some aspects of the environment that could potentially cause disruptions to the ATM information system achieving its purpose:

- Power failure – consider the consequences of loss of power half way through a transaction.
- Problems with network connection – could be a physical loss of the complete connection or an issue with response times.
- Incorrect output of cash – could be the result of crumpled notes or notes sticking together.
- Insufficient receipt paper, receipt ink or cash – how can this be detected and what response is reasonable.
- Fraudulent use – consider techniques for dealing with incorrect PINs, physical tampering with the machine, unusual transaction patterns for individuals, etc.



**GROUP TASK Discussion**

How is it that each of the above points relates to the environment within which ATMs operate? Discuss.

**GROUP TASK Discussion**

Discuss suitable techniques that are, or could be used to overcome or at least lessen the impact should any of the above disruptions occur.

PURPOSE

The purpose of an information system is to fulfill some need or needs. To achieve this purpose is the aim or objective of the system. In fact the purpose of the system is the whole reason for the system's existence. To accurately realise the system's purpose requires an understanding of who the information system is for and what it is they need to achieve. Therefore the purpose of an information system is very closely linked to the needs of those for whom the system is created.

**Purpose**

A statement identifying who the information system is for and what it needs to achieve.

The purpose of an information system should be stated clearly and in achievable terms. The word purpose implies a conscious and determined act, which is achieved through guided and thoughtful processes. The purpose of the system should remain at the forefront during the creation and use of any information system.

Information systems can be designed for individuals or for organisations. Information systems for organisations must meet the common needs of the individuals that make up the organisation. Determining these needs and then translating them into a common purpose can often be quite a daunting yet crucial task.



Fig 1.6

Understanding needs leads to a clear and achievable purpose.

Determining the purpose of an information system involves the following steps:

1. Identify the people whose needs the information system should meet.
2. Formulate a list of needs that the information system should realise.
3. Translate these needs into objectives that form the purpose of the information system.

When developing new information systems the purpose is used as the basis for developing a series of definite and achievable requirements. If the requirements are achieved then the purpose has also been achieved.



Consider the following scenarios:

- The territory manager for an oil company has some 500 service station, factory and rural customers to service. Their job is to maintain contact with existing customers as well as to promote the oil company to potential customers. A

separate department processes all orders and deliveries so the territory manager's only input in this area occurs when a problem arises with one of their clients. Most of the territory manager's time is spent visiting each of their customers to ensure personal contact as well as to provide information on new products.

There are some twenty territory managers across the country and each is free to use any information system that suits their needs. Some territory managers use a traditional diary/planner whilst others utilise electronic versions and even laptop computers. The oil company provides either printouts or computer files containing all customer details and sales histories for their area.



GROUP TASK Discussion

Assume you have just gained employment as a territory manager. What are your information needs? How would you decide which form of information system you would use? Discuss.

- Each school needs a timetable to operate effectively. The requirements relate to each teacher and student knowing where they should be and what they should be doing at any given time.



GROUP TASK Discussion

What is the purpose of timetable systems in schools? What needs do they address? Discuss.



GROUP TASK Discussion

Discuss the processes and resources used at your school to create, maintain and publish the school's timetable. Do these processes and resources achieve their purpose successfully?

INFORMATION PROCESSES

Collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving and displaying are all examples of information processes. Together these seven basic activities are what needs to be done to transform the data into useful information. The bulk of the preliminary course deals with these information processes and their related tools. We therefore need to be crystal clear about the concept of information processes before we proceed further.



Information Processes

What needs to be done to collect and transform data into useful information. These activities coordinate and direct the system's resources to achieve the system's purpose.

In general, information processes are computer and non-computer based activities that are carried out using the resources or tools of the information system. These activities coordinate and direct the system's resources to complete the required task and achieve the system's purpose. Therefore information processes use participants, data and information and information technologies to achieve the system's purpose. Information processes are not necessarily performed by computer-based technologies; they can equally be performed using other means.



Consider the following:

Most of us own an address book; this is an example of an information system. Let us consider some of the information processes necessary for this information system to operate:

- We collect names, addresses and phone numbers of our friends, relatives and other acquaintances. This does not happen all at once, we revisit this information process each time we wish to add a new contact.
- We decide on the format we will use in our address book. Perhaps each page has three columns; one for names, another for addresses and a third for phone numbers. To enable us to later locate an individual we setup individual pages for each letter according to surname.
- We recognise the first letter of the surname to enable us to correctly store the data. We also isolate the name, address and phone number. This process, although it seems trivial in this example, is where we make sense of the data, that is, it is transformed into information.
- We locate the correct page in the address book and write in the new contacts details.
- We locate the correct page and then scan to the required contact's name and read their details.
- When a contact moves house or changes their phone number we find their name and edit the changed details.
- We skim through our address book and prepare a list of individuals to be invited to a party.
- We use the phone numbers or addresses to contact individuals.



GROUP TASK Discussion

Classify each of the above information processes as either collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving or displaying. Discuss your responses.

PARTICIPANTS

Participants are a special class of user who carry out or initiate the information processes. Users are all the people who view or make use of the information output from an information system. Participants also view or use information from the system; however they are also actively involved in the operation of the information system. The word participate involves sharing and having a part in something, therefore participants in an information system share and have a part in its operation. They perform or carry out the system's information processes.

For most information systems there are a variety of different personnel; some directly use the system, others indirectly use the system and some create or develop



Users

People who view or use the information output from an information system.



Participants

A special class of user who carries out (or initiates) the information processes within an information system.

the system. Participants are involved in the actual operation of the system and are integral to that operation, in essence they are resources used by the system. Thus indirect users and developers are not considered participants in the system. Indirect users are often a source of data for the system or they receive information from the system rather than being themselves involved in its operation. These indirect users are in the system's environment. They can influence or be influenced by the information system but they do not directly carry out its information processes. For example, a customer in a shop is an indirect user of the cash register system. They provide data and in turn are provided with receipts but they have no control over the information processes that occur to transform this data into information. The shop assistant, on the other hand is a direct user of the cash register, they carry out the information processes and are therefore a participant in the system.

Development personnel include system analysts, engineers, programmers and testers. The job of developers is to design, create and upgrade the system, rather than being involved in its operation. As a consequence most development personnel are not considered participants. Their job is over once the system is implemented and operational.

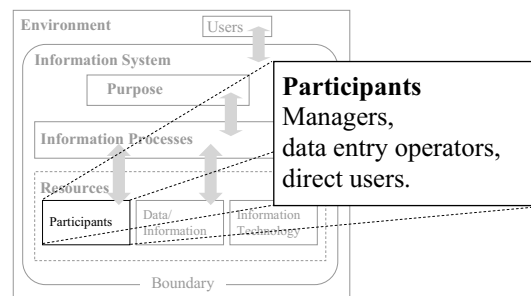


Fig 1.7

Participants are an integral resource of all information systems

We now have an understanding of the personnel who are not participants, let us now consider personnel who are participants in the information system. Participants are made up of all the personnel who are used by the information processes of the system. This includes managers, data entry operators and other users who initiate or perform information processes as part of the system's operation. Most of these personnel can be classified as direct users; they directly interact with the information system during its operation. Each of these groups of people are part of the information system's resources and without their contribution the system would not operate. Each is used as a resource during one or more of the system's information processes. For small systems an individual fulfills multiple tasks, whereas larger systems operating within organisations may have many personnel engaged in each task.



Consider the following:

Each school is required by law to maintain an accurate information system to monitor student attendance on a daily basis. There are many people involved in this information system including the principal, deputy principal, office staff, teachers, parents and government bodies including the NSW Board of Studies.



GROUP TASK Discussion

Consider each of the personnel mentioned above in relation to your school's attendance system. Are they participants in the information system? Justify your response by outlining the information processes in which each are involved.

DATA/INFORMATION

Data is the raw material of an information system in the same way as timber is the raw material for a carpenter building a deck. The whole aim of an information system is to process data into information. Thus data is a required resource for all information systems. This data is transformed using information processes into something useful that achieves the system's purpose. In the case of an information system the 'something useful' is information; in the case of the carpenter the 'something useful' is the finished deck.

Most data is itself the information derived from another system or process, and similarly the information output from a system is often used as data for another system or process. Consider *Fig 1.8*, which describes the issuing of a speeding ticket by a police officer. The speed is the information derived from the officer's speed radar. The speed is then used to determine the amount of the fine when issuing the speeding ticket. In turn this information is used as data by the driver when they pay their fine. *Fig 1.8* is a simplified dataflow diagram describing the flow of data and information through three processes. Note that each arrow indicates information out of a process and data flowing into a subsequent process.



Data

The raw material used by information processes.



Information

Information is the output displayed by an information system. Knowledge is acquired when information is received.

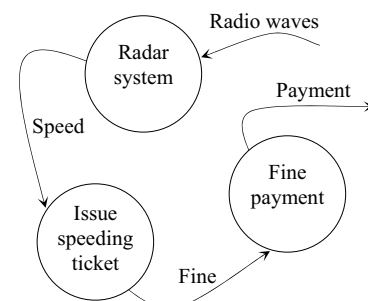


Fig 1.8
Simplified dataflow diagram for the speeding ticket scenario.

Earlier in this chapter we discussed information as being the meaning that a human assigns to data. This is the central purpose of information systems, to derive meaning from data. To do this requires the resources and information processes of the system. The system must be able to understand the nature of the data if it is to successfully transform it into information. In Chapter 2 we examine, in some detail, the nature of data and how it is represented in digital form.



GROUP TASK Activity

Data is represented in many different and varied forms. For example video can be stored in analogue form on videotape, sent using radio waves to a TV, transmitted using cable, or stored digitally on a DVD. Make a list of as many different forms of data as you can together with different ways in which the data is stored or transmitted.

INFORMATION TECHNOLOGY

We know what information is, but what is technology? Technology is the result of science being applied to practical problems. This is what engineers do; they apply scientific knowledge to practical problems to produce new technology. Therefore technology is things people create to assist them in solving some problem. A hammer is an example of technology; it is used to assist us to use nails and bind timber together. The scientific principles of force, leverage and momentum have been used



Technology

The result of scientific knowledge being applied to practical problems.

to engineer the hammer into a technology. Because the hammer is used to solve other problems it is also known as a tool. The items of technology we examine in this course are also used as tools to assist in the solution of problems.

Technology is not always physical objects, it can also be the way things are done or the steps undertaken to accomplish some aim. For example DNA technology is more about techniques than physical tools. In this case scientific principles have led to the development of these new techniques, in turn these new techniques are used in medicine, forensics, genetics, etc.

Information technology encompasses all the tools used to assist an information system to carry out its information processes. Most of this course is devoted to examining these tools. In general information technology can be split into hardware and software. The hardware being the physical equipment and the software being the instructions that coordinate and direct the operation of the hardware. Computer hardware is particularly suited to many information processing tasks, however there are many other forms of hardware that are also used as resources within information systems. Much of the preliminary course is devoted to examining various tools, both hardware and software, that are used to complete the various information processes.



Information Technology

The hardware and software used by an information system to carry out its information processes.



Consider the following:

Writing, publishing and printing a book, such as this one, requires various tools. These tools are forms of information technology and include computer hardware and software together with various non-computer based tools.



GROUP TASK Activity

Make a list of all the tools that would most likely be required during the writing, publishing and printing processes.



GROUP TASK Discussion

Your list created above probably contains many computer-based tools. How would the processes performed by these tools have been accomplished prior to the introduction of computers? Discuss.



Consider the following:

One of the major functions for most secretaries is to type various forms of letters for their bosses. The letters produced need to be stored in a logical manner so they can be later retrieved efficiently. In this information system the secretary is the sole participant. The boss and the recipients of the letters being indirect users.

Let us examine the information processes and technologies used by a particular secretary, who we'll call Sue, to complete these tasks.

Information Processes

- Letters being dictated by boss.
- Entering, editing and formatting letters.
- Printing and posting letters.
- Saving and retrieving letters.
- Backing up files.

Information Technologies

- Personal computer
- Laser printer
- Wooden in and out trays
- Word processing software
- Operating system, specifically the file management utility

Normally the sequence of events for preparing a letter is as follows:

1. Sue is called into the boss's office, when her boss wishes to dictate a letter.
2. She then returns to her office and places the work into her in tray according to its urgency compared to other work in her in tray.
3. Once the letter reaches the top of the in tray she types it into the word processor, edits, formats and prints it.
4. Sue is in the habit of saving new files as soon as they are created and then about every 15 minutes or so. Sue has her own system of filing where she has setup a hierarchy of folders on her hard disk, she also uses the date as part of each filename.
5. The completed letter is placed in her out tray.
6. Towards the end of each day Sue takes all the letters from her out tray into her boss to be signed.
7. She then prepares them for mailing and posts them.

**GROUP TASK Activity**

List the forms of information technology Sue uses during each of the seven steps listed above.



Consider the following:

A stationary supply company sells approximately 500 different products. Most of their sales are made either by phone or fax. The company has five telephone sales people who answer phones and respond to incoming faxes. The salesman then enters the order into their computer. Once an order has been entered by the sales staff a delivery docket is printed on the printer in the warehouse. The warehouse staff then pack the goods and send them off with the delivery docket. The day after the goods have been shipped all invoices from the previous day's orders are printed and posted.

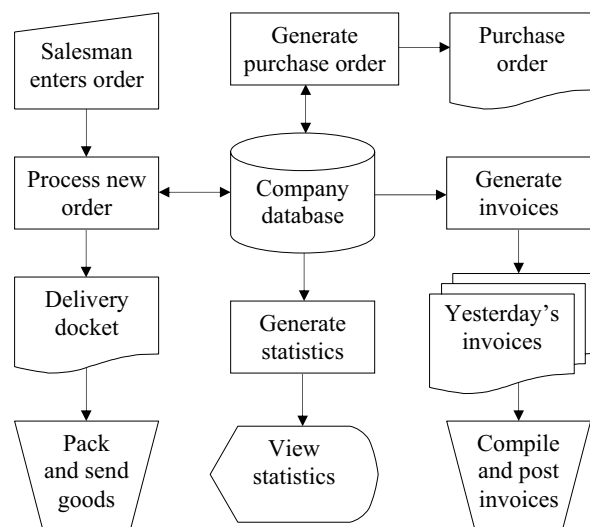


Fig 1.9

System flowchart for the stationary supply company.

Management are able to view up-to-date statistics on sales, stock levels and overdue accounts. The warehouse is able to view current stock levels, graphs on the sales history of each product and are able to produce orders to enable them to purchase

stock. In total the company employs 20 personnel, including a single technical person who oversees the computer network and its security.

The system flowchart, shown in *Fig 1.9* above, shows the logic and basic flow of data through this information system. For example: a salesman manually enters an order, the order is processed which would likely involve checking for available stock and its price and then storing the order details, the delivery docket is printed and finally the goods are packed and sent. System flowcharts are not a necessary part of the IPT course, therefore there is no need understand the meaning of each symbol. The flowchart is included to diagrammatically represent the system.



GROUP TASK Activity

List all the people involved at each symbol of the system flowchart above. Are there personnel involved who are not participants?



GROUP TASK Activity

List the forms of information technology used by this system. Include both computer and non-computer based technologies.



HSC style question:

Many newer mobile phones include GPS (Global Positioning System) navigation functions. The GPS receiver within the phone receives data from satellites to accurately determine the current location of the phone. The current location is used by the navigation software to plot the users current position on a map and also to direct them to other locations they specify. The initial map data and also regular updates are downloaded from a website and stored on the phone's flash memory card. Downloads occur either using the phone's 3G wireless connection or via a USB interface to an Internet connected computer.

Consider the GPS mobile phone navigation system as an information system.

- Identify the data collected by the system and the information displayed by the system.
- Identify the information technology within the mobile phone that is used to implement the GPS navigation functions.

Suggested Solution

- Data collected includes updates to map data in digital form, satellite data in digital form used to pinpoint position of phone and user entered locations of interest.

Information displayed includes plot of current position on displayed map, directions to selected locations (includes spoken instructions and path on displayed map).

- Information technology includes speaker (to generate speech output), LCD screen (to display maps), GPS receiver (to receive time/location data from satellites), keypad (for entry of locations or menu choices), wireless receiver/transmitter (connection to 3G network), USB interface (connection to computer), flash memory card (for storing/retrieving map data), CPU and RAM (for all processing, such as generating graphical maps), and navigation software to determine current location, process GPS data, produce maps and speech output.

SET 1A

1. The circumstances and conditions that surround a system but are not part of the system are known as the:
 - (A) system's purpose.
 - (B) environment.
 - (C) resources of the system.
 - (D) processes performed by the system.
2. Data entry operators are primarily involved in which information process?
 - (A) organising.
 - (B) storing.
 - (C) displaying.
 - (D) collecting.
3. "I'm going to be late for school" is an example of A, "The time is 9:25 AM" would be an example of B and "0925" would be C. A, B and C respectively could best be replaced with the words:
 - (A) data, information, knowledge.
 - (B) information, knowledge, data.
 - (C) knowledge, information, data.
 - (D) knowledge, data, information
4. Information technology is a term used to describe:
 - (A) the hardware and software resources of the system.
 - (B) the resources available to the system.
 - (C) each of the information processes occurring within the system.
 - (D) how scientific knowledge is applied to the solution of practical problems.
5. Participants in an information system commonly include all of the following:
 - (A) managers, end-users, programmers, engineers and data entry operators.
 - (B) direct users, indirect users, managers and data entry operators.
 - (C) managers, direct users, system administrators, engineers and data entry operators.
 - (D) managers, end-users, system administrators, network personnel and data entry operators.
6. A system can be best described as:
 - (A) a collection of connected sub-systems that work together to achieve a purpose.
 - (B) containing data, participants, information technology and information processes.
 - (C) an organised assembly of resources and processes that interact to achieve a common purpose.
 - (D) all the different types of organisms present in the environment that interact and are dependant on each others actions.
7. The purpose of an information system:
 - (A) is the reason for the system's existence.
 - (B) is to fulfill some need or needs.
 - (C) should be clear and achievable.
 - (D) All of the above.
8. Activities that coordinate resources to achieve the system's purpose are known as:
 - (A) Information Technologies.
 - (B) Information Activities.
 - (C) Information Resources.
 - (D) Information Processes.
9. The main difference between data and information is:
 - (A) Data is always digital whereas information is not.
 - (B) Data is the raw material that is processed by information systems to create information.
 - (C) There is no difference; they are interchangeable terms.
 - (D) Data is individual characters or numbers, whereas information is words, sentences and charts.
10. Which term defines the line between what is and what is not part of a system?
 - (A) users
 - (B) boundary
 - (C) information technology
 - (D) information processes
11. Define each of the following terms:

(a) Environment	(c) Information System	(e) Participants
(b) Purpose	(d) Information Processes	(f) Information Technology
12. Make a list of all the different types of information technology you use every day. Which of these utilise computer-based technologies to operate?
13. A telephone directory can be thought of as an information system. Are the names, addresses and phone numbers data or information? Discuss.
14. Think of the kitchen in your home as a system. What is its purpose? List all the resources used by this system? What processes are used in the kitchen?
15. Open a new document in a word processor with which you are familiar. Examine each of the menu items and classify each item as collecting, organising, analysing, storing and retrieving, processing, transmitting and receiving, or displaying.

SOCIAL AND ETHICAL ISSUES

We live together as a social group rather than in isolation. For this to occur harmoniously requires laws for correct conduct but it also involves many unwritten ways of going about the business of living. These unwritten morals are known as ethics. Some ethics have evolved over time into laws, however many others remain principles that are understood by society and that influence the conduct of its members. For example, most would agree that it is morally unacceptable to commence a new relationship whilst already in a relationship. It is not illegal, however most of us would look poorly on someone who does this. We would also accept that stealing is unethical, in this case society has, over time, created laws to ensure those who steal are punished.



Social

Friendly companionship.
Living together in harmony rather than in isolation.



Ethical

Dealing with morals or the principles of morality. The rules or standards for right conduct or practice.

In this section we examine social and ethical issues arising from the processing of information. These issues affect not only the participants within the system but also those outside the information system. It is the responsibility of system designers to ensure the information systems they create take account of social and ethical issues. Likewise participants must ensure they use systems in a socially and ethically acceptable manner. Some of these issues have been recognised by governments and as a consequence laws have been enacted to ensure compliance. Some of the major issues include:

- Privacy of the individual
- Security of data and information
- Accuracy of data and information
- Data quality
- Changing nature of work
- Appropriate information use
- Health and safety
- Copyright laws



Consider the following:

1. A website collects email addresses and subsequently sends out advertising emails.
2. A mail order company sells its customer details to another direct mail company.
3. An employee of an energy company views details of her friends' accounts.
4. A student downloads information from the web and uses it as part of an assignment.
5. An employee spends at least 8 hours per day at the keyboard.



GROUP TASK Discussion

Under certain conditions each of the above scenarios could be socially and ethically acceptable, and under others they would not. Discuss.

PRIVACY OF THE INDIVIDUAL

Privacy is about protecting an individual's personal information. Personal information is any information that allows others to identify you. Privacy is a fundamental principle of our society, we have the right to know who holds our personal information. Privacy is a feeling of seclusion, where we can be safe from observation and intrusion. For this to occur we need to feel confident that our personal information will not be collected, disclosed or otherwise used without our knowledge or permission.

Personal information is required, quite legitimately by many organisations when carrying out their various functions. This creates a problem, how do we ensure this information is used only for its intended task and how do we know what these intended tasks are? Laws are needed that require organisations to provide individuals with answers to these questions. In this way individuals can protect their privacy.



GROUP TASK Activity/Discussion

Make up a list of all the organisations that are likely to hold personal information about you. Do you know what information is held and how it is used?

In Australia, privacy is legally protected under the *Privacy Act 1988* and its subsequent amendments. This act contains ten National Privacy Principles, that set standards that organisations are required to meet when dealing with personal information; the text in *Fig 1.10* briefly explains each of these principles.

What are the ten National Privacy Principles?

The following briefly explains what the NPPs mean for you.

NPP1: Collection - describes what an organisation should do when collecting your personal information.

NPP2: Use and Disclosure - outlines how organisations can use and disclose your personal information.

NPP3: Data Quality & NPP4: Data Security - set the standards that organisations must meet for the accuracy, currency, completeness and security of your personal information.

NPP5: Openness - requires organisations to be open about how they handle your personal information.

NPP6: Access & Correction - gives you a general right of access to your own personal information, and the right to have that information corrected, if it is inaccurate, incomplete or out of date.

NPP7: Identifiers - says that generally, Commonwealth government identifiers (such as the Medicare number or the Veterans Affairs number) can only be used for the purposes for which they were issued.

NPP8: Anonymity - where possible, requires organisations to provide the opportunity for you to interact with them without identifying yourself.

NPP9: Transborder Data Flows - outlines privacy protections that apply to the transfer of your personal information out of Australia.

NPP10: Sensitive Information - requires your consent when an organisation collects sensitive information about you such as health information, or information about your racial or ethnic background, or criminal record. Sensitive information is a subset of personal information and special protection applies to this information.

Fig 1.10

The ten 'National Privacy Principles' briefly described from the Office of the Federal Privacy Commissioner's website at <http://www.privacy.gov.au>

Consequences of the *Privacy Act 1988* mean that information systems that contain personal information must legally be able to:

- explain why personal information is being collected and how it will be used
- provide individuals with access to their records
- correct inaccurate information
- divulge details of other organisations that may be provided with information from the system
- describe to individuals the purpose of holding the information
- describe the information held and how it is managed



GROUP TASK Research

Using the Internet, or otherwise, examine the privacy policies for a number of organisations that hold personal information. Do these policies address the above dot points appropriately?

SECURITY OF DATA AND INFORMATION

Security of most resources is about guarding against theft or destruction. For example, an alarm on your car aims to deter thieves and vandals. PIN and PUK codes on mobile phones are deterrents to theft. Similar techniques are used to protect data and information, however there is an additional problem; most data and information can easily be edited or copied without any noticeable change to the original. We therefore require additional techniques and strategies for dealing with the security of data and information.

Generally the larger the information system becomes the more crucial effective security of the data and information becomes. If your home computer crashes then the consequences are annoying but if a bank's computer system fails, even for an hour, the consequences are enormous.

Some possible security issues that all information system designers need to consider include:

- Virus attacks – Viruses are software programs that deliberately produce some undesired or unwanted result. Most viruses are spread via attachments to emails but also by infected media such as flash drives and CDs.
- Hackers – These are people, often with extensive technical knowledge and skill, who aim to overcome or get around any security mechanisms used by a computer system. This allows them to view, utilise and even edit data and information.
- Theft – Unauthorised copying of data and information onto another system. Also physical theft of hardware, and as a consequence, the data and information it contains.
- Unauthorised access by past and present employees – Past and present employees may maliciously tamper with data or they may view and use data of a private nature inappropriately.
- Hardware faults – Failure of hardware, and in particular storage devices, can result in loss of data. It is inevitable that hardware will eventually fail at some time.
- Software faults – Errors in programs can cause data to become corrupted. No software is completely free of errors.

Some strategies commonly used to address the above issues include:

- Passwords – Passwords are used to confirm that a user is who they say they are. Once verified the user name is then used by the system to assign particular access rights to the user.
- Backup copies – A copy of important files is made on a regular basis. Should the original file fail or be lost then the backup copy can be used. It is important to keep backup copies in a secure location.
- Physical barriers – Machines storing important data and information, or performing critical tasks are physically locked away.
- Anti-virus software – All files are scanned to look for possible viruses. The anti-virus software then either removes the virus or quarantines the file. The widespread use of networks, and in particular the Internet, has made anti-virus software a virtual necessity.
- Firewalls – A firewall provides protection from outside penetration by hackers. It monitors the transfer of information to and from the network. Most firewalls are used to provide a barrier between a local area network and the Internet.
- Data encryption – Data is encrypted in such a way that it is unreadable by those who do not possess the decryption code.
- Audit trails – The information system maintains records of the details of all transactions. The aim is to make it possible to work backwards and trace the origin of any problem that may occur.

To implement the above strategies requires that procedures be put in place to ensure their correct operation. For example: if an employee leaves, their user name and password needs to be removed, anti-virus software needs to be updated regularly to take account of any new viruses and backup copies need to be checked to ensure they are occurring correctly.



GROUP TASK Activity

Some strategies aim to prevent security issues occurring whilst others help correct the problem once it has occurred. Classify each of the strategies above as either prevention or correction.



Consider the following:

1. An employee works on a file on their home computer. They then email the file to work. Unfortunately the file contains a virus.
2. The network administrator for a company is reading employees' emails without their knowledge.
3. Scott likes trying to 'get around' the security on government computer systems. He never changes any of the data he finds, he just enjoys breaking in.
4. An employee, whose job is to chase overdue accounts, marks the account of a friend as paid.



GROUP TASK Discussion

What strategies could be used to stop, or at least discourage, each of the above scenarios from occurring? Discuss.

ACCURACY OF DATA AND INFORMATION

Inaccurate data results in incorrect information being output from the information system. The consequences of such incorrect information can be minor, for example a letter addressed incorrectly, or major, for example a country going to war. The term ‘data integrity’ is used to describe the correctness, accuracy and validity of data. All information systems should include mechanisms for maximising data integrity.

There are various techniques used including: data validation and data verification checks. Data validation involves checking the data is in the correct format and is reasonable as it is entered into the system. For example your HSC assessment mark in this course must be a number between 0 and 100, the software can perform such validation and ensure this is the case. However, knowing the mark entered is your actual result is a different matter. Data verification checks ensure the data entered is actually correct. For example, although 97 is a legitimate HSC mark, perhaps it was mistyped as 79, data verification aims to correct such errors. In this case, the data entry operator may be required to physically check each entered mark before pressing the submit button. Verifying data as correct is a much more difficult task than validating it as reasonable. Data can become inaccurate over time, for example addresses change, so verifying the accuracy of data is an ongoing process.

The accuracy of collected data is improved when the format of data collection forms ensure data is in the required format and required range. For example computer-based forms can use check boxes ☒, radio buttons ☐, or list boxes to ensure input is of the type required. These items are said to be ‘self-validating’ – they ensure the data entered is reasonable in terms of format and range. Both computer and paper-based forms can include masks that provide a template to indicate the format of the data required. For example a phone number mask could be (_ _) _ _ _ _ _ _ _ , a post code mask could be _ _ _ _ _ .

DATA QUALITY

Quality data meets the requirements of all information systems that will make use of the data. For example, a database that processes customer orders is not just used by the ordering system; it is also used for stock control, analysing sales patterns, marketing and numerous other tasks. Quality data meets the needs of all systems. Many organisations develop data quality policies and standards to ensure the data within its systems will meet the needs of all its current and future systems.

There are a number of perspectives that should be considered when assessing data quality. Accuracy, timeliness and accessibility are three common data quality perspectives (there are many others). The importance of each perspective is closely related to the particular information systems that will utilise the data. The different perspectives are not separate, rather they each have an effect on the others. For instance, inaccurate data occurs when data is not updated in a timely manner. In terms of accuracy, data quality encompasses the above section on “Accuracy of data and information”.

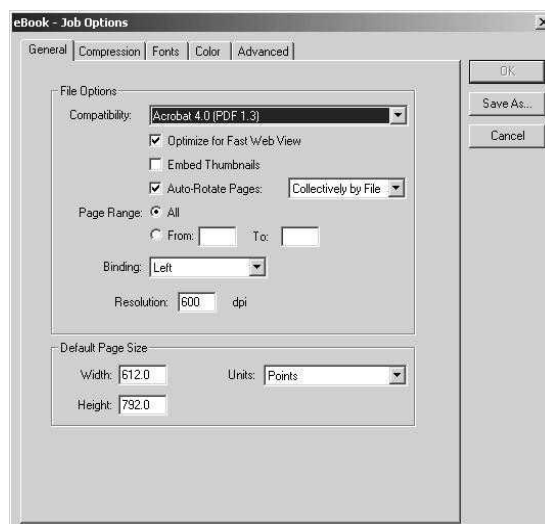


*Fig 1.11
Adobe Acrobat screen showing a number of
self-validating screen elements.*

The timeliness of data relates to how soon changes to data are actually made and also how soon such data changes are available to other processes or systems. For example, purchases made using a credit card can take some time to be reflected in both the purchasers account and the merchants account. If the purchase is processed using an online facility then both accounts are adjusted in close to real time, however if the purchase is processed manually then it can be some days for the account balances of the purchaser and merchant to reflect the change.

Accessibility of data refers to the availability and suitability of data for processing. For example, many organisations maintain separate databases at each branch. Management at head office requires access to all branch databases if it is to accurately produce sales totals. If the business only calculates monthly sales totals then online access to each branch database may not be a priority, however if sales totals are monitored on an hourly basis then online access is needed. In addition, if the organisation of the data within each branch database is different then it will be difficult for the head office system to calculate the sales totals efficiently. For instance, some branches may add GST to each product within each order whilst others may add GST to the total of each order.



Consider the following:

A time and motion study is being undertaken for a white goods manufacturer. Each worker on the assembly line is asked to keep records on the time taken to assemble each component. The results of the study are used to pinpoint bottle necks in the manufacturing process.

The results are compiled and as a consequence various recommendations are made to management. Management disagrees with many of the recommendations and doubts the accuracy of the data used. It is later found that the times submitted by many of the individual workers were inaccurate. When these times are totalled the result is far greater than the time they actually worked.



GROUP TASK Discussion

Discuss reasons why the workers recorded inaccurate times? What techniques could have been used to improve the quality of the data?

CHANGING NATURE OF WORK

The nature of work has seen significant change since the 1960s. These changes have been both in terms of the types of jobs available and also in the way work is undertaken. The widespread implementation of computer-based systems, including computer-based information systems, has been the driving force behind most of these changes. In the early 1970s many thought that the consequence of new technologies would be a reduction in the total amount of work needing to be done; this has not occurred. Rather new industries and new types of employment have been created. Many people are now working longer hours, in more highly skilled and stressful jobs than ever before.

The term 'Information Technology Revolution' has been widely used to describe changes occurring over the last few decades, however more recently the term 'Global Knowledge Economy' has emerged. Information and communication technologies can be regarded as truly global technologies; they provide the ability to code information and share it globally at high speed and at minimal cost. Consider the

growth in Internet usage (see *Fig 1.12*). In the ten years from 1997 to 2007 the percentage of Internet users globally has risen from 2% up to 22% and it continues to climb. Globalisation means markets have expanded and international competition has increased. Furthermore components, services and capital used by business can be sourced from a worldwide market place.

These changes in the nature of the economy are having profound affects on the nature of work for the majority of employees. They have altered the type of jobs available to employees as well as altering the way employees perform these jobs. Let us now examine what these changes are and how they affect workers.

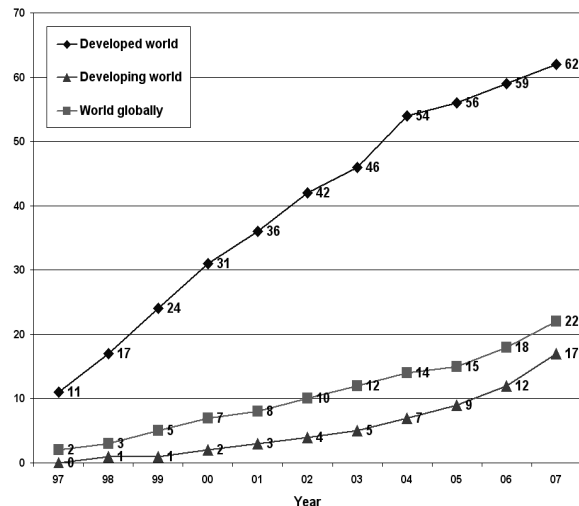


Fig 1.12

Internet users per 100 inhabitants 1997-2007
 Source: International Telecommunication Union



GROUP TASK Brainstorm

There are many jobs now that just did not exist in the 1960s, and there are also many jobs that have almost totally disappeared. Make up a list of all the different types of jobs that have been created since the 1960s and another list of those jobs that have virtually disappeared.

Changes in the type of employment

During the 1960s there was much concern in regard to the automation of many tasks traditionally undertaken using manual labour. These jobs were predominantly found within goods producing industries such as agriculture, mining, manufacturing, construction and utilities. The fear, at the time, was that unemployment rates would spiral out of control. Although there has been a significant decline in the number of jobs within goods producing industries there has also been a corresponding increase in knowledge and person based service industries. The data and graph shown *Fig 1.13* uses information from the Australian Bureau of Statistics to illustrate this trend. Knowledge and person based service industries include finance, property, education, health, entertainment and communication industries.

Jobs within knowledge and person based service industries require skills in regard to using technology rather than skills that substitute for technology. For example, a clerk no longer needs to manually search through filing cabinets, rather they need to be able to use software to query a database. In other words, the technology performs the search under the clerk's direction; the clerk requires more advanced skills to direct the search than were required to carry out the manual search. Similarly an increase in the importance of inter-personal skills and a decrease in the importance of manual skills is occurring. There is little need for physical strength and coordination in knowledge and person based service industries rather there is an increased need for people to communicate more effectively with each other.

Industry	1970	1975	1980	1985	1990	1995	2000	2005
Goods producing	44.4	40.3	37.4	33.4	30.7	27.8	26.9	24.7
Knowledge and person based services	26.1	29.5	32.4	36.4	39.2	42.5	48.4	50.6
Other – includes retail, government admin, transport and storage	29.5	30.2	30.2	30.2	30.1	29.7	24.7	24.7

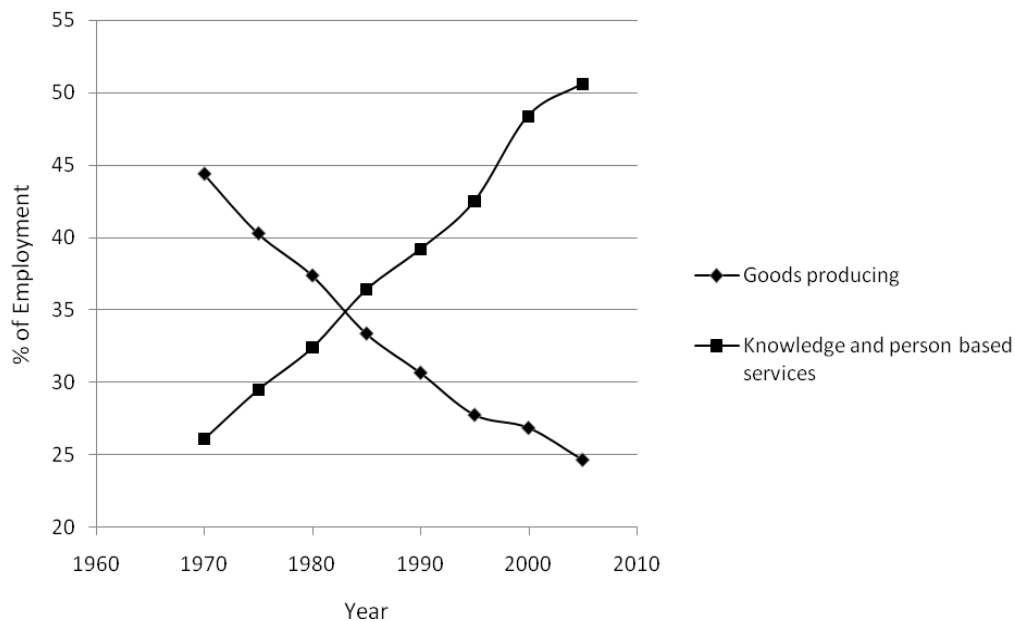


Fig 1.13

Employment in Australia by Industry Group. Data sourced from Australian Bureau of Statistics



Consider the following:

In 1967, when the HSC was first introduced, about 18,000 students sat for examinations in 28 different courses and only approximately 20% of Year 10 students completed the HSC. Now more than 65,000 students sit for examinations in around 70 courses and about 70% of Year 10 students complete the HSC.



GROUP TASK Discussion

Discuss possible links between the changes to the types of available employment and the HSC statistics in the above statement.

Changes in the way work is undertaken

Traditionally we think of employment providing a steady wage or salary and involving regular working hours, usually somewhere between 35-40 hours per week; the tasks performed at work being well defined, consistent and directed by management. Most people had the expectation that throughout their working life they would work for a single employer; changes in employment only occurring for grossly sub-standard performance or by choice. This is no longer the case.

The chart in *Fig 1.14* describes the number of hours worked per week by employees across Australia. The small percentage working 0 hours is made up of those on leave of some type during the survey week. The changes in working hours reflect the trend that many more people are now employed part time, as casuals or on fixed term contracts. From the chart we see that about 30% of people worked a 35-40 hour week in 2000, where as in 1970 this number was around 51%. It is becoming less common to work, a so called standard 35-40 hour week. It is common for many people to work a different number of hours in different weeks due to various flexible work arrangements such as rostered days off and nine day fortnights.

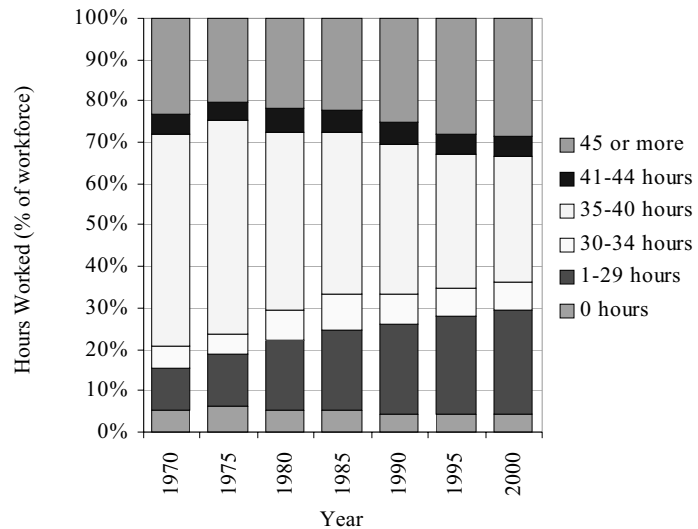


Fig 1.14

Hours worked by employees shown as percentage of total workforce. Data sourced from Australian Bureau of Statistics.

It is also interesting to note that the number of people working less than 30 hours or more than 45 hours per week has steadily increased. Although flexible work arrangements account for some of this change, it is also true that those working long hours are generally professionals in highly skilled and varied jobs. Research indicates that these people not only earn significantly larger incomes than average, but that they also experience higher levels of job satisfaction.



Consider the following:

The data used to produce the chart in *Fig 1.14* is based on a particular survey week in each of the included years; remember that today there are many workers whose hours are flexible; they may work 10 hours one week, 45 the next and 25 the following week. This may well account for some of the increase in employees working 1-29 hours however the significant increase from around 10% in 1970 to around 25% in 2000 requires examination.



GROUP TASK Discussion

Why do you think there has been such an increase in the number of employees working between 1 and 29 hours per week?



GROUP TASK Discussion

“The rich are getting richer and the poor are getting poorer”. Discuss this statement in regard to both income and job satisfaction.

APPROPRIATE INFORMATION USE

Information is created to fulfill some purpose, however often this same information is also useful to assist in achieving some other purpose. The possibility for inappropriate use of information arises. Inappropriate use of information can occur intentionally or it can be quite innocent and unintentional. It is vital to thoroughly understand the source, nature and accuracy of any information before it is used.

Some examples of inappropriate information use include:

1. Client's email addresses, collected by a business, are sold to a direct mail company. The direct mail company then sends out mass advertising or spam mail.
2. A student adds up their trial HSC marks, and converts the total to a percentage in an attempt to estimate their UAI.
3. Credit checks are made on all applicants for a job based solely on their name. The employer incorrectly culls some applicants when in fact it is someone else, who has the same name, that has the poor credit rating.
4. A graph showing a steady increase in sales over the past few years is used to predict future sales. Management insists each salesman increases their sales to match this future prediction.
5. A newspaper reporter uses the number of students who gained a band 6 in IPT to rank the effectiveness of schools.



GROUP TASK Discussion

Consider each of the above numbered points. In each case is the inappropriate use of information intentional or unintentional?

The appropriate use of information systems is often detailed as a policy statement for the organisation. The policy outlines inappropriate activities together with the consequences should a user violate any of the conditions. Typically such a policy statement would include the following activities as inappropriate usage:

- Unauthorised access, alteration or destruction of another user's data, programs, electronic mail or voice mail.
- Attempts to obtain unauthorised access to either local or remote computer systems or networks.
- Attempts to circumvent established security procedures or to obtain access privileges to which the user is not entitled.
- Attempts to modify computer systems or software in any unauthorised manner.
- Unauthorised use of computing resources for private purposes.
- Transmitting unsolicited material such as repetitive mass mailings, advertising or chain messages.
- Release of confidential information.
- Unauthorised release of information.



GROUP TASK Discussion

Consider the policy statements above in conjunction with the five examples at the top of the page. Would such a policy assist in ensuring the appropriate use of information for each of these five examples?

HEALTH AND SAFETY

All workers are exposed to potential health and safety problems whilst undertaking their work. Employers are responsible for ensuring these risks are minimised. In NSW the *Occupational Health and Safety Act 2000*, together with the *Occupational Health and Safety Regulation 2001* are the legal documents outlining the rights and responsibilities of employers and employees in regard to occupational health and safety. Workcover NSW administers this act in NSW to ensure and monitor compliance. Employers must setup a procedure for identifying and acting on occupational health and safety (OHS) issues. This requirement is often fulfilled by appointing either an OHS representative or by forming an OHS committee.

Ergonomics is the study of the relationship between human workers and their work environment, it is not just about the design and placement of furniture, rather it is about anything and everything that affects the work experience. This includes physical, emotional and psychological aspects of work.



Ergonomics

The study of the relationship between human workers and their work environment.

Most participants in information systems primarily work in offices at computer workstations. Some broad ergonomic issues relevant to this type of work environment include:

- Furniture and computer hardware design and placement should be appropriate to the task. This includes desks, chairs, keyboards, monitors, pointing devices, etc.
- Artificial lighting should appropriately light the work area. Outside and overhead lighting should not cause glare.
- Noise levels generated by equipment, but also from other workers, to be at reasonable levels. Research shows that conversations from fellow workers are a major distraction to most workers.
- Work routine should include a variety of tasks designed to minimise boredom and discomfort. Working continuously on the same task is the greatest cause of repetitive strain injury (RSI).
- Software design should be intuitive and provide shortcuts for experienced users. The user should drive the software, the software should not drive the user. Training should be thorough and ongoing.
- Procedures for reporting potential OHS problems should be in place and understood by all employees.

Further details in regard to ergonomic considerations particular to information systems will be examined in later chapters. Be aware that lack of job satisfaction has been shown to be closely linked to poor ergonomics. Health and safety is not just about minimising and dealing with injuries, rather it concerns the total work experience.



GROUP TASK Discussion

Do you work part-time? If so, consider each of the above dot points in relation to your job. Discuss potential problems in regard to health and safety at your work place.

COPYRIGHT LAWS

Copyright laws are used to protect the legal rights of authors of original works. The *Copyright Act 1968*, together with its various amendments, details the laws governing copyright in Australia. Copyright laws are designed to encourage the creation of original works by limiting their copying and distribution rights to the copyright owner. The copyright owner is normally the author of the work, except when the work was created as part of the author's employment; in this case the employing organisation owns the copyrights. Without copyright laws there would be little economic incentive for authors to create new works.

Copyright does not protect the ideas or the information within a work, rather it protects the way in which the idea or information is expressed. For example, there are many software products that perform similar processes, however these processes are performed in different and original ways, hence copyright laws apply. Generally copyright protection continues for the life of the author plus a further fifty years.

All works are automatically covered by copyright law unless the author specifically states that the copyrights for the work have been relinquished. The use of the familiar copyright symbol ©, together with the author's name and publication date is not necessary, however its use is recommended to assist others to establish the owner of a work's copyrights.

Computer software, data and information is easily copied, and the copy is identical to the original. This is not the case with most other products. As a consequence special amendments to the Copyright Act have been enacted.

In regard to software:

- One copy may be made for backup purposes.
- All copies must be destroyed if the software licence is sold or otherwise transferred.
- Decompilation and reverse engineering is not permitted. The only exception being to understand the operation of the software in order to interface other software products.

In regard to compilations of information (such as collected statistics and databases of information):

- The information itself is not covered.
- There must have been sufficient intellectual effort used to select and arrange the information; or
- The author must have performed sufficient work or incurred sufficient expense to gather the information even though there was no creativity involved.



Consider the following:

1. An employee takes a copy of a customer database with them when they leave.
2. A friend gives you a copy of a computer game they got for Christmas.
3. You create a digital phone book using name, address and phone numbers downloaded from Telstra's white pages web site.



GROUP TASK Discussion

Discuss the implications, in terms of Copyright Law, for each of the above scenarios.



HSC style question:

Car Guide is a business that collects details on the sale of motor vehicles and then sells this information to subscribers. Subscribers pay an annual fee and then receive regular printed reports through the mail. Many of the subscribers are motor vehicle dealers, although individuals are also welcome to subscribe.

Each report is personalised to suit the requirements of each individual subscriber. Subscribers can specify search criteria including date range, vehicle manufacturer, vehicle model, year of manufacture and postcode of sale. Also various other report details can be specified, such as summary information, charts or even the raw motor vehicle sales data.

Collecting the raw data is the most time consuming and costly part of Car Guide's operation. Car Guide pays dealerships to supply details of each vehicle sold, however they must telephone private sellers to obtain similar details.

- (a) Some subscribers have expressed concern with regard to the accuracy of the reports they receive from Car Guide. Identify areas where inaccuracies may be introduced into Car Guide's information system and suggest strategies to minimise these inaccuracies.
- (b) The collected data is entered into Car Guide's database and is then processed to create each subscriber's report. Discuss who owns the copyrights over the database and who owns the copyrights over the final reports.

Suggested Solution

- (a) There could be problems in the data entry of the received data from Dealerships and/or private sellers. This could be picked up by appropriate validation checks in the software to highlight obviously unreasonable values as data is entered. Appropriate verification processes should also be insisted on where the data entry people are trained to check every paper form carefully (or read back their entered data on the phone to private sellers) before pressing the submit button.

Private sellers may tend to exaggerate the sell price to reflect their advertised price rather than admit to selling for a lower price. Perhaps Car Guide should not even try to contact the people selling the car, but go to the RTA where all transfers of cars are registered and ask if the data can be transferred electronically (and of course ethically) to them.

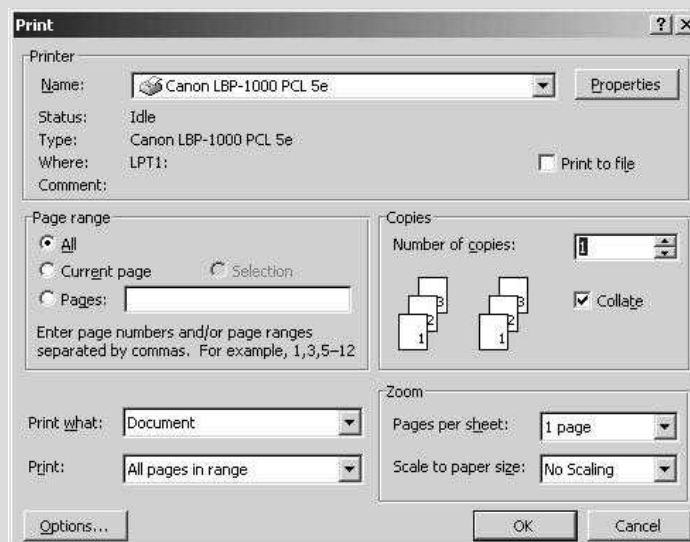
Payment to Dealerships could also cause errors in the data collected. They may be eager to submit extra data and include sales that are too high or even fictitious to impress Car Guide. This could be pre-empted by only accepting exports of real sales data transmitted directly from the various Dealerships computer systems.

- (b) There has been significant effort on the part of Car Guide to collect and compile the sales data, therefore they own the copyrights over the database. These copyrights are in terms of the organisation and other processing performed by their database system. The reports are also produced by this system hence Car Guide owns the copyrights over the reports (unless the subscriber contracts specify otherwise). The individual sales records within the database are raw facts so although this data originates from the dealerships and private sellers it is unlikely to be covered by copyright law as no significant intellectual activity was needed to create the data.

SET 1B

1. A business that is unable to explain why they are collecting personal information is in breach of the:
 - (A) Copyright Act 1968.
 - (B) Occupational Health and Safety act 2000.
 - (C) Privacy Act 1988.
 - (D) None of the above, they are just being unethical.
2. Passwords can be used to:
 - (A) increase security.
 - (B) protect the privacy of sensitive information.
 - (C) stop unauthorized copying of files.
 - (D) All of the above.
3. The difference between data validation and data integrity is:
 - (A) there is no difference, they are interchangeable terms.
 - (B) Validation ensures the data is reasonable and is in the correct format at entry time, integrity is about ensuring it is correct.
 - (C) Validation is about the screen items used to make up computer-based forms, whereas integrity is to do with the underlying data.
 - (D) Integrity checks ensure the data is reasonable and is in the correct format at entry time, validation is about ensuring it is correct.
4. Restoring files after the complete failure of a file server can only happen if which of the following has occurred:
 - (A) Anti-virus software was installed and regularly upgraded.
 - (B) All files have password protection.
 - (C) an audit trail is maintained by the system.
 - (D) Regular backups have been made.
5. Automation of many tasks traditionally undertaken by manual labour has resulted in:
 - (A) high unemployment within the total population.
 - (B) a decline in jobs available within goods producing industries.
 - (C) an increase in jobs within knowledge and person based service industries.
 - (D) manual labourers now working in knowledge and person based industries.
6. A government employee creates an information system for his department. In terms of copyright:
 - (A) he is able to sell licenses to other parties to use this system, provided he does not include the government department's data or information.
 - (B) he is the author and possesses the copyrights.
 - (C) the government owns the copyrights and he may not take the system if he leaves their employ.
 - (D) the law is not precise in this regard, he should seek the services of a copyright lawyer if he wishes to market the system.
7. Ergonomics is concerned with:
 - (A) furniture design and placement.
 - (B) reducing work place injuries.
 - (C) the total work environment.
 - (D) ensuring OHS principles are enforced.
8. For copyright law to apply, works must:
 - (A) display the copyright symbol ©.
 - (B) be copied and distributed for profit.
 - (C) contain original ideas or information.
 - (D) None of the above.
9. It is true to say that over the past 30 years or so:
 - (A) the number of people working longer hours has decreased and the number working shorter hours has increased.
 - (B) there has been very little change in the hours worked by employees.
 - (C) the number of hours worked by most people has decreased significantly.
 - (D) the number of people working longer hours has increased and so too has the number of people working shorter hours.
10. The term 'Global Knowledge Economy' has arrived as a consequence of:
 - (A) the increase in knowledge and person based service jobs compared to those in other industries.
 - (B) automation within goods producing industries resulting in lowered manufacturing costs.
 - (C) the widespread implementation of computer-based technologies across the globe.
 - (D) the ability to code and share information across the world at high-speed and low cost.

11. Consider each of the following scenarios. For each, describe suitable methods available for rectifying the situation:
 - (a) A number of data entry operators are experiencing muscle strain, particularly in their wrists.
 - (b) A software developer discovers that one of their products is being distributed illegally over the Internet.
 - (c) You continue to receive spam mail from a company despite informing them to remove you from their mailing list.
 - (d) A business continues to send you an invoice for products you never ordered or received. They are now threatening legal action.
12. A number of legal documents are discussed in the text. Make a list of these documents and briefly describe their purpose.
13. Examine the screen shot below:



There are a number of different types of controls on the above screen. Identify those that are 'self-validating' and those that are not. For the controls that are not self-validating describe appropriate checks that could be used to ensure the data input is reasonable.

14. Doctors hold much private information on each of their patients. It is therefore crucial that their patient files and records are kept secure and the information is used appropriately. List and describe a number of techniques suitable for ensuring this occurs.
15. 'Work is a necessary evil. You put in your 8 hours labour each day, get a pay packet at the end of the week and a bit of a holiday every so often. This just the way it is!'
Do you agree with the above quote? Discuss, in relation to the changing nature of work over the past 30 years or so.

CHAPTER 1 REVIEW

1. Information processes include:
 - (A) participants, data, information and information technology.
 - (B) environment, purpose, resources and participants.
 - (C) collecting, organising, analysing, storing, processing and displaying.
 - (D) data, information and knowledge.
2. A system that is itself an integral part of another system is called a(n):
 - (A) information system.
 - (B) system resource.
 - (C) sub-system.
 - (D) information technology.
3. Which of the following is true of information technology?
 - (A) It is the result of science being applied to a practical problem.
 - (B) It includes hardware and software.
 - (C) It is all the tools used to perform a system's information processes.
 - (D) All of the above.
4. When assessing data quality, which of the following should be considered?
 - (A) accuracy, timeliness, accessibility.
 - (B) privacy, security, copyright.
 - (C) text, numbers, images, audio, video.
 - (D) users, participants, developers.
5. Passwords are used to:
 - (A) assign particular access rights to users.
 - (B) confirm that a user is who they say they are.
 - (C) block unauthorised access to a network.
 - (D) restrict the activities of employees, both past and present.
6. Privacy of the individual is primarily concerned with:
 - (A) protecting an individual's personal information.
 - (B) ensuring all individuals have access to their personal information.
 - (C) making sure personal data held is accurate.
 - (D) enforcing the 10 National Privacy Principles specified in the *Privacy Act 1988*.
7. Systems that maintain an audit trail are doing so to ensure:
 - (A) fair and equitable access to information.
 - (B) all transactions can be traced to their source.
 - (C) copyright is respected at all times.
 - (D) employee's actions can be observed and monitored.
8. A hacker is someone who:
 - (A) knowingly and maliciously creates and/or distributes viruses.
 - (B) uses data and/or information inappropriately.
 - (C) uses their skills to circumvent the security of computer systems.
 - (D) has extensive knowledge of computer systems.
9. All information systems:
 - (A) contain participants, data/information and information technology.
 - (B) process data into information; this is their primary purpose.
 - (C) operate within an environment which influences, and is influenced, by the information system.
 - (D) All of the above.
10. Jack downloads some images from the web to include on a commercial website. Apparently, the images do not include any sort of copyright notice or license agreement.
Which of the following is true?
 - (A) As there is no copyright mark, notice or license agreement; Jack is free to use the images as he pleases.
 - (B) Images cannot be copyrighted, so it is legal for Jack to use the images.
 - (C) The images may or may not be covered by copyright, but as there is no copyright notice and the images were found on the web then it is reasonable to assume they are in the public domain.
 - (D) Jack should assume the images are covered by copyright. It would be wise to contact the website and find out the copyright status of the images before using them.

11. Define each of the following terms:

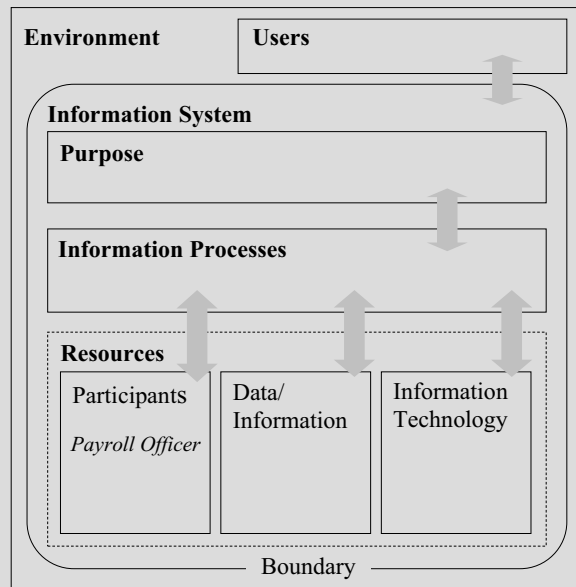
(a) system	(c) environment	(e) information technology
(b) information	(d) purpose	(f) information processes

12. Describe three factors that should be considered when assessing the quality of data.

13. Describe, at least five, strategies you would hope your bank uses to ensure the privacy and security of your personal information and money.

14. Copy and complete the system diagram at right for the scenario that follows. For example: The payroll officer is clearly a participant so “Payroll Officer” should be written in the participant section of the diagram.

A factory employs approximately 50 workers. There is also an attached office where the payroll officer has a computer attached to the company’s local area network. Her secretary, in the adjoining office, enters each employee’s hours worked into the payroll system each day. On Thursday mornings the payroll officer calculates the gross pay, tax and net pay for each employee. She then generates pay slips for each employee together with a summary page for the factory manager. The factory manager must sign the bottom of the summary before the transfer of any money is permitted. Given that the summary sheet is approved the payroll officer then checks for sufficient funds in the company’s accounts, electronically transfers each employee’s pay into their individual accounts and generates a cheque for the taxation office. The secretary then distributes the pay slips and posts the tax cheque to the taxation office.



15. Consider the following:

- The gas company reads each of their customer’s meters every 3 months.
 - These meter readings are used to calculate total consumption for the period and hence to generate accounts.
 - In about 10% of cases it is not possible to read the customer’s meter. When this occurs a notice is sent to the customer requesting them to do a “self read” of the meter. If no response is received within 14 days then an estimate of consumption is made based on the previous year’s consumption.
 - Sometimes a problem is revealed as a consequence of the meter reading. These problems are referred to the troubleshooting team for further investigation. Some examples of problems include:
 - ⇒ The reading is identical to the previous one, seemingly indicating there has been no consumption.
 - ⇒ The reading indicates consumption is radically different to previous years.
 - ⇒ The reading is excessively high.
- (a) Describe the data used by this system and the information generated by the system.
- (b) List the different groups of participants within this information system. Describe the tasks performed by each of these groups.
- (c) Each team of meter readers is allocated particular suburbs and towns. Their performance is monitored by head office in terms of total meters read and percentage of successfully read meters. Suggest possible problems that could emerge if performance is measured solely on these statistics.

In this chapter you will learn to:

- distinguish between, and categorise, the activities within an information system in terms of the seven information processes
- use an existing information system to meet a simple need
- manually step through a given information system identifying the information process
- for a given information system, describe how the following relate to the information processes:
 - participants
 - data/information
 - information technology
- schematically represent the flow of data and information through a given information system, identifying the information processes
- distinguish between data and information in a given context
- categorise data as image, audio, video, text and/or numbers
- identify the data and the information into which it is transformed, for a given scenario
- identify examples of information systems that use information from another information system as data
- explain why information technology uses digital data
- describe advantages and disadvantages for the digital representation of data

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies.

In this chapter you will learn about:

Information processes

- collecting – the process by which data is entered into or captured by a computer system, including:
 - deciding what data is required
 - how it is sourced
 - how it is encoded for entry into the system
- organising – the process by which data is structured into a form appropriate for the use of other information processes such as the format in which data will be represented
- analysing – the process by which data is interpreted, transforming it into information
- storing and retrieving – the process by which data and information is saved and accessed later
- processing – a procedure that manipulates data and information
- transmitting and receiving – the process that sends and receives data and information within and beyond information systems
- displaying – the process that controls the format of information presented to the participant or user

The nature of data and information

- data – the input to an information system
- data representation – the different types of media, namely:
 - image
 - audio
 - video
 - text
 - numbers
- information – the output which has been processed by an information system for human understanding
- the generation of information from data via the information processes
- how information from one information system can be data for another information system

Digital representation of data

- the need for quality data, including:
 - accuracy
 - timeliness
 - accessibility
- current data digitising trends, for example:
 - newspapers on the Internet
 - telephone system
 - video on DVD
 - facsimile
 - media retrieval management

INTRODUCTION TO INFORMATION PROCESSES AND DATA

In this course we, somewhat arbitrarily, split information processing into seven areas. Although this makes sense, in terms of understanding each of the information processes, it is rare for such a distinction to exist in reality. More often an individual process will involve actions from multiple ‘syllabus’ information processes. Think of these seven syllabus information processes as the basic building blocks of processing for any information system. Data is collected, organised, analysed, stored, retrieved, processed, transmitted, received, and displayed in virtually all information systems when viewed at almost any level of detail. We find examples of each of these information processes happening when we view an overall picture of a large system, such as the whole Internet, but we can also find instances of most of the information processes when examining the detailed operation of the central processing unit within a single computer.

In this chapter, we first consider the connections and relationships between each of the information processes specified in the syllabus. We then examine the actions performed during each of these information processes. Finally, we examine the data used by information systems. We examine different types of data and how it is represented and transformed by the system’s information processes.

RELATIONSHIPS BETWEEN INFORMATION PROCESSES

Remember information processes are actions that direct and coordinate the system’s resources to affect the data within the system in some way. As we are dealing with information systems there must be a flow of data and/or information into and out of the system. There must also be a flow of data into and out of each information process within the information system itself.

Fig 2.1 is a generalised context diagram, it shows data and information flows between the system and its environment. Each data flow arrow is labelled to describe the nature of the data. Data moves into the information system from an external entity in the environment. An external entity that provides data to an information system is known as a ‘source’. A source may be an indirect user, a communication link from another system or any other source of data that is external to the system. The information processes within the system perform their actions on this data and output the resulting information from the system. An external entity that is the recipient of output is known as a ‘sink’. An information system can have multiple sources and multiple sinks; it is also common for a single external entity to be both a source and a sink.

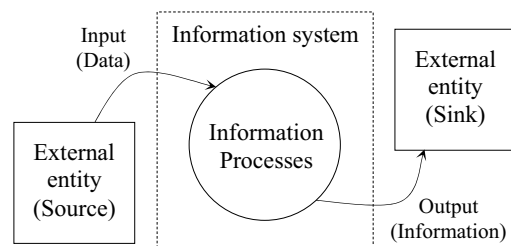


Fig 2.1

Context diagram for a typical information system.



Consider the following:

A keyboard can be considered to be an information system. It obtains input from the user as they type, it then processes these keystrokes into digital signals that are output to the computer.



GROUP TASK Activity

Draw a context diagram to illustrate the flow of data described above. Do you think a keyboard really is an information system?

Let us now discuss the flow of data between the information processes within a typical information system. *Fig 2.2* is a dataflow diagram that includes each of the seven information processes specified in the syllabus. The aim of this diagram is to illustrate the complex flow of data that occurs between information processes. For example, data may be collected, organised and stored. At a later time this data may be

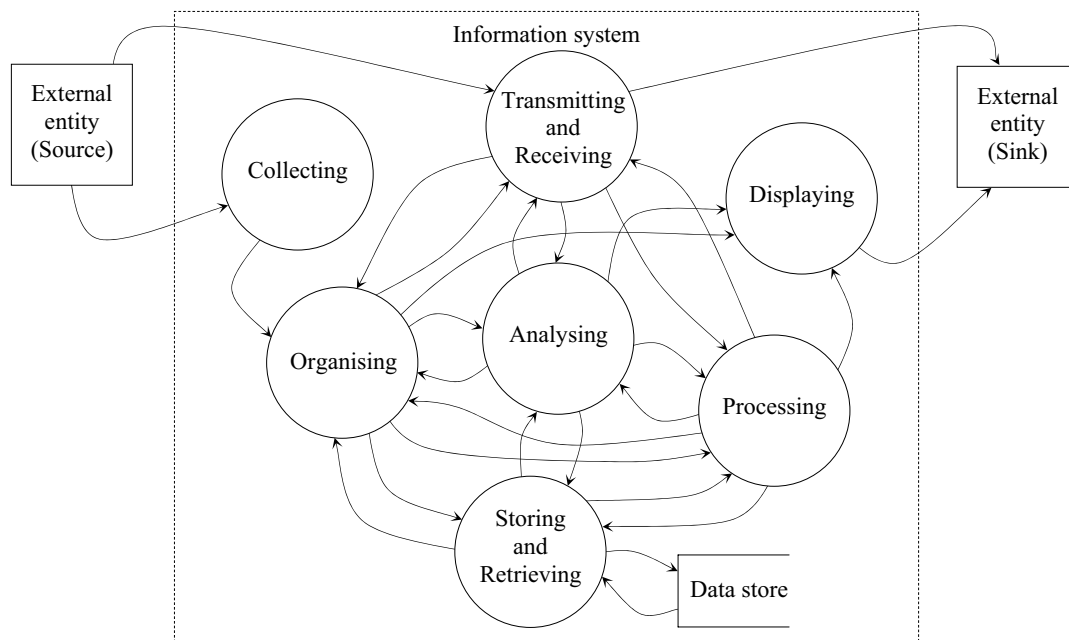


Fig 2.2

Dataflow diagram showing some of the possible data flows within an information system.

retrieved and processed, the results may then be analysed and displayed. This example is but one of an almost infinite number of ways of following the complex network of data flows shown on the diagram. It is important to realise that on data flow diagrams there is no attempt to describe the order in which the processes occur, as the name 'dataflow diagram' implies, they describe the movement of data between information processes. Despite this, it is often true that the nature of the processes involved tends to imply a particular order.

You may notice that not all the information processes on the diagram are connected to each other in both directions, and others are not directly connected at all. Why is this? In many cases the nature of the data output from a certain process requires further processing before it is suitable as input to other processes. Consider the collecting information process; before data collected can be stored it must be organised into an appropriate format for storage. Displaying is a process that outputs data from the

system, therefore it accepts input from other processes but only outputs to external entities. Data stores are locations where data is permanently stored; such as files, databases or even filing cabinets. It therefore makes sense that ‘Storing and Retrieving’ is the only information process that deals directly with data stores.



GROUP TASK Activity

Use a word processor to create, format, save and print a simple one-page document. Whilst performing this task make a note of each of the different information processes occurring together with the exchange of data between these processes.



GROUP TASK Discussion

No doubt your notes from the above activity indicate a certain sequence of events. Must this sequence be strictly followed? Discuss.

So far we have discussed the relationship between information processes in terms of data flowing between each of the information processes. Another important relationship between information processes concerns the order in which these processes occur. Each unique information system will have different processing requirements in regard to order, however there are some processing sequences that tend to exist in most information systems. The systems flowchart shown in *Fig 2.3* describes some of these sequences. For example, at least some data must be collected prior to commencing any of the other information processes. This data must also be organised appropriately prior to analysis or further processing. Although systems flowcharts are not part of the IPT syllabus they are useful tools to indicate the logic of the information processes, which is essentially the order in which each process occurs.

The systems flowchart shown in *Fig 2.3* seems to indicate that collecting must be complete before the organising process commences and similarly organising must be complete prior to analysing and processing beginning. This is not the intention, rather the intention of *Fig 2.3* is to show the path taken by an individual unit of data once it has entered the system.

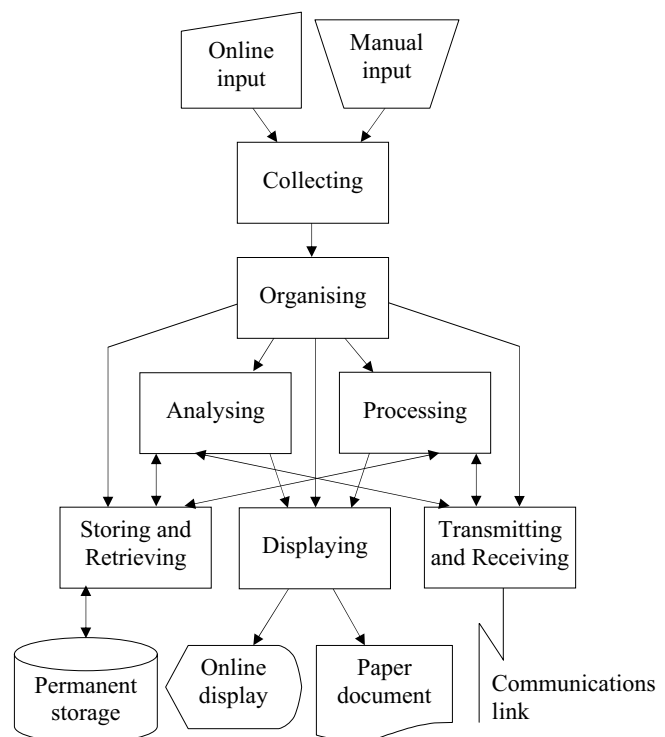


Fig 2.3

Systems flowchart showing some processing sequences that tend to exist in most information systems.



Consider the following:

A florist sells flowers over the phone, in person or via their website. Each time an order is received it is entered into the store's computer system. Once the sales assistant enters the customer's name they are presented with a list of possible matches from the customer database. The sales assistant can either select one of the possible matches or enter a new customer. Customer details are entered even for 'in person' orders; this provides a marketing tool for the florist whereby they can examine purchasing trends for individual customers and send them advertising brochures at appropriate times. If an order is to be delivered then it is printed and placed in an 'in tray' for later completion and dispatch. At the end of each day the owner generates a sales report detailing the number of each product sold, number of each product remaining in stock, together with the total value of all products sold. Examining this report each day assists the florist make suitable stock purchases when they visit the markets each morning.



GROUP TASK Discussion

Read through the above scenario and identify each of the syllabus information processes occurring. Reread the scenario and identify the data used by the system and the information produced.

COLLECTING

In previous sections we alluded to collecting as the information process that gathers data from some outside entity. This is true; collecting is essentially an input process, its purpose being to provide data from the environment to other information processes within the information system. For example, entering keywords into a search engine on the Internet is a collection process and so too is scanning an image using a flatbed scanner. In both cases data is entered or captures from the environment for use by the system's information processes.

To perform this collecting input process requires more than just the actual entry or capture of data; it requires an understanding of what data is required, from where it will come, together with how the data will be encoded. Consider

performing a search using a search engine – the search engine is the information system. In this case the search criteria entered must comply with a particular syntax defined by the collection process, that is; the search engine system defines the data it needs and users must comply with this definition. The source of the search criteria (data) is the user. The data is entered via the keyboard using a text box on a web page being viewed on the user's remote computer. During data entry the search criteria is encoded into a sequence of ASCII characters using the keyboard, web page, web



Collecting

The information process by which data is entered or captured by a computer system.



Fig 2.4

Collecting involves more than just gathering the data.

browser and various other information technologies within the user's computer system. The encoded data is then transmitted via the Internet to the search engine. In summary, to perform collection we need to know what data we require, where the data will come from and how we will get it into the system. Answers to all these questions must be determined when designing collection processes.

Let us now consider each aspect of the collection process, namely:

- deciding what data is the required,
- determine the source of the data, and
- determine how the data will be encoded.

DECIDING WHAT DATA IS REQUIRED

What data is needed by the system to achieve its purpose? To answer this question requires analysis of the system, and in particular its purpose, to determine its data requirements. For example, an invoice includes information about the customer, the supplier, the products and also the date and various calculated costs. The aim is to determine the necessary data that is required without collecting data multiple times or collecting data that can be derived or calculated from existing data. In our invoice example it makes sense to collect the address details of each customer just once and then reuse this data each time the customer places an order. Similarly the cost of each item need only be entered once, and can be used each time that product is ordered. The sub-totals, GST and totals do not need to be entered as each of these can be calculated using other data.

It is then necessary to consider the detail of each required data item. For example, if an address is needed then is it appropriate to collect it as a single data item or should each element of the address be collected separately? If an image is required then what resolution is needed and should it be colour? How can the validity and integrity of each data item be checked? Ask yourself questions, such as: what makes a data item legitimate? Does the value of one data item influence the value of another? For example, particular products may have accessories that can only be ordered with the product and only apply to that particular product.



Consider the following:

Addresses, phone numbers and dates are commonly required data for many information systems. There are various ways of defining each of these; some examples of the final output required could include:

1. 5/88 John Street, Mays Hill 2145	96355517	15/01/2003
2. Unit 5, 88 John St, Mays Hill, 2145	(02) 9635-5517	15-Jan-03
3. 5/88 John St. Mays Hill 2145 NSW	+61 2 96355517	15 January 2003
4. Unit 5/88 John St. Mays Hill 2145 Ph: 9635 5517 Date: 15/1/03	5. 5/88 John St. MAYS HILL NSW 2145 Tel: (02) 9635-5517 Wednesday, January 15, 2003	



GROUP TASK Discussion

What actual address, phone number and date data needs to be collected so that it is possible to later display the results in any of the formats shown above? Discuss.

DETERMINE THE SOURCE OF THE DATA

The source of the data is the place or origin from which the data is obtained, for example, temperature sensors, customers, web sites and government departments. The collection process must be able to identify this source prior to the actual data collection commencing.

Often there will be a choice of data source and it will be necessary to make a decision as to which source is the most suitable. In other cases, a variety of different data sources are used. Some issues to consider when determining the suitability of a data source include:

- **Availability** – Is the data source readily available? Perhaps information from an existing data source can be used rather than collecting data from scratch. For example, using rainfall data from the Bureau of Meteorology rather than installing rain gauges.
- **Data Quality** – Can you assess the accuracy of the data source and will it remain accurate? Secondary sources provide data that has previously been processed in some way; this can make assessing the quality of the data more difficult. On the other hand, primary sources of data often require substantially more effort to obtain; yet their accuracy can be determined more directly.
- **Cost** – What is the cost in terms of time, effort and money of using a particular data source? It may not be cost effective to collect data from every individual, however a significant sample of the population may be sufficient for the needs of the information system.



Consider the following:

1. A courier company uses an information system to assign particular jobs to each of its drivers. The system attempts to assign jobs in such a way that a minimum of time is spent travelling between drop off and pick up points; currently the straight line distance between points is used as the basis for these decisions.
2. Each time a customer places an order they submit their mailing details. In some cases these mailing details are different to the existing mailing details within the system. When this occurs the mailing details are updated and all subsequent communications are sent to the new address.
3. The state government is trying to assess the viability of building a new runway at Mascot airport. They require information in regard to projected air traffic as well as any increased noise and other environmental impacts.



GROUP TASK Discussion

Identify the possible sources for the data required, or used, by each of the above information systems.



GROUP TASK Discussion

Selection of data sources is a compromise between availability, quality and cost. Discuss using examples from the above scenarios.

DETERMINING HOW THE DATA WILL BE ENCODED

Once the required data has been defined and the source of this data has been identified it is necessary to determine how the data will be gathered or encoded for use by the system. To do this requires deciding on the most appropriate tools and procedures to use. In Chapter 3, we examine tools used for collecting in detail. Some of the hardware tools commonly used include: scanners and digital cameras for collecting images, microphones for collecting audio data, video capture devices, keyboards and optical character recognition devices for collecting text, together with a number of other specialised data collection devices. To collect data also requires software tools. In Chapter 3 we consider examples of different types of software used to interface with hardware devices resulting in data entry into various software applications.

Data is commonly encoded using a form of some sort. In this context a form can be either printed or it could be a computer based input form. The content of the form is used to prompt the user for each piece of required data. Often the user's response will be limited to a list of possible valid choices or the length of the data item will be indicated. For example, a list of possible products where the user ticks each one they require or a postcode indicated using a mask with 4 squares. These techniques are used where the data is collected directly from a user; this is not always the case. Consider image data that is to be captured using a scanner, it has its own special data requirements related to the purpose of the information system. Images to be used for publications require a far higher resolution than those to be used on a web page hence the gathering process must ensure the resolution is suitable. Voice mail data recorded over a telephone is of poor quality compared to that required for a music compact disk therefore quite different tools and procedures are required to encode each of these types of audio data.



Consider the following:

Fig 2.5 shows a credit card sales voucher that is used to collect sales transaction data. The card number, name and validity dates are collected directly by taking an imprint of the customer's card; details particular to the sale are entered by hand.

Fig 2.5

Credit card sales voucher used to collect required data from various sources.



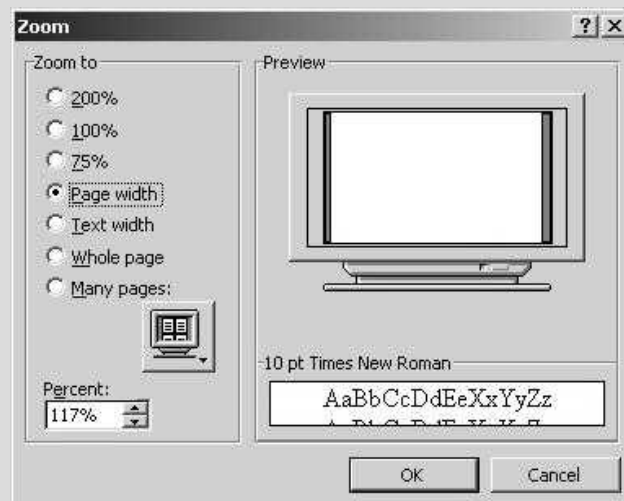
GROUP TASK Discussion

Examine the sales voucher above. List and describe each element on the voucher that assists to define the required data for the sale. Identify the source of each of these data items. What techniques are used to ensure the data on the voucher is valid?

SET 2A

1. Context diagrams:
 - (A) show the relationships between information processes.
 - (B) show data movements between the system and its environment.
 - (C) describe the logical flow of data through an information system.
 - (D) are used to model processes that transform data into information.
2. Dataflow diagrams:
 - (A) are used to describe data connections between the system and its environment.
 - (B) aim to describe the logical sequence of processes.
 - (C) describe the movement of data in and out of each information process.
 - (D) must include all seven of the information processes.
3. Deciding on the information technology that will be used during a collection process is part of:
 - (A) deciding what data is required.
 - (B) determining the source of the data.
 - (C) determining how the data will be encoded.
 - (D) All of the above.
4. The seven syllabus information processes:
 - (A) must all be present in any information system.
 - (B) are seldom all present in a single information system.
 - (C) are present in most information systems.
 - (D) usually occur
5. A typical real world information process:
 - (A) is likely to perform processing from multiple syllabus information processes.
 - (B) would only include processing from a single syllabus information process.
 - (C) will always collect and display information, and may also utilise other information processes.
 - (D) is composed of hardware, software, data, information and people..
6. Before actually collecting data, one needs to determine the detailed nature of each data item. This would be considered part of:
 - (A) deciding what data is required.
 - (B) determining the source of the data.
 - (C) determining how the data will be encoded.
 - (D) the general collecting information process.
7. The information processes that communicate with the environment are:
 - (A) collecting and displaying.
 - (B) transmitting and receiving.
 - (C) Both (A) and (B).
 - (D) All seven information processes.
8. Collecting the same data multiple times:
 - (A) is a good idea, as it can be checked for accuracy.
 - (B) is necessary if the data is later to be displayed using different formats.
 - (C) indicates a poor design and is always unnecessary.
 - (D) should be restricted to crucial data.
9. Hardware tools for collecting data include:
 - (A) keyboards, scanners, microphones and monitors.
 - (B) printers, monitors, plotters and speakers.
 - (C) questionnaires, interviews, meetings and observation.
 - (D) keyboards, scanners, microphones and barcode readers.
10. In terms of collecting, the data required by the information system should always be:
 - (A) derived or calculated from existing data so that it can be reproduced should the system crash.
 - (B) gathered using the same format in which it will be displayed so further processing is minimised.
 - (C) validated after it has all been entered to save time during the data entry process.
 - (D) None of the above.

11. Consider each of the following scenarios. For each, draw a suitable context diagram:
- A speed camera detects speeding motorists and takes their photo. Each week an RTA officer collects these photos and delivers them to the RTA's fines department.
 - Enrolment forms are distributed to possible new students. Those interested complete the forms and return them to the school. One of the school secretaries sorts the forms based on the school's enrolment policy and passes them onto the Principal.
 - The Australian Taxation Office (ATO) processes tax returns based on various government legislation together with rulings from the high court.
 - Fred is an author of technical books. During the writing process he consults other books, the Internet and various experts in the field. Fred's final manuscript is emailed to his publisher.
12. The process of answering this question involves many information processes. Identify and describe these processes.
13. Examine the following screen shot from Microsoft Word.



Describe how each element on this screen assists the collection process.

14. Before the actual collection of data commences it is necessary to:
- define the required data,
 - determine the source of the data, and
 - determine how the data will be encoded.

Describe what needs to be done to accomplish each of these tasks.

15. At the start of this chapter we discussed the seven information processes as building blocks for processing in any information system. These building blocks operate together to achieve the system's purpose.

Discuss the types of relationships and connections that exist between these information processes.

ORGANISING

Organising is the information process that determines the form in which the data will be structured and represented; it applies this structure to the data within the information system. Organising converts the data by structuring it and representing it in a new form. The organising process



Organising

The information process by which data is structured into a form appropriate for the use of other information processes.

does not alter the data itself rather it modifies the way the data is structured and represented. For example, data entered into a text file is structured as a sequence of characters where each character is represented using its ASCII code; the data remains the same, it is the organisation of the data that has changed.

Organising is required after collection, however it is also common for data to be reorganised at other times to make it suitable for use by other information processes. The aim of organising is to provide data to other information processes in the most efficient format relative to the data needs of that process. For example, if a graph is required to display the total sales per month then the date of each sale needs to be represented in such a way that the month can easily be extracted and the sales data needs to be structured so that all dates and total sales can efficiently be analysed.

To assist in understanding the organising process let us consider structuring and representing as separate processes. In reality, both these processes usually occur virtually simultaneously.

STRUCTURING

In this context, structuring is the process that arranges the data in some specific and logical way. The structure is designed to best suit the requirements of the information processes that utilise the data. Programmers design data structures so that their programmers can efficiently access and process individual data elements.

Examples of common data structures include:

- In a spreadsheet the data is structured into rows and columns. This arrangement makes it easy to reference data items in terms of their columns and rows. For example, C4 refers to the individual data item stored in the cell at the intersection of column C and row 4.
- In a database table the data is structured into records and fields. Each record contains all the data about a particular entity, and each field holds a particular attribute of that entity. For example, in a customer table each record holds all the data about a particular customer, and the Surname field holds that customer's surname.

REPRESENTING

Each individual data item must be coded so that it can be understood and used efficiently by other information processes. The coded data represents or symbolises the actual data. Different types of data are represented in different ways depending on their intended purpose; later in this chapter we consider the digital representation of data and in Chapter 4 we examine specific tools used to accomplish this process.

Examples of representing include:

- When writing, we use combinations of letters to represent or symbolise words.
- When doing maths we use the digits 0 through to 9 in various combinations as symbols to represent numbers.
- Each picture element (pixel) within a digital photograph is stored as a binary number that represents the colour of that pixel.



Consider the following:

During your senior school studies you will complete various assessment tasks for each of your courses. You'll need to know when these tasks are due, together with the details of each task. Organising this data effectively should help you plan your time so that each task is completed on time and to the best of your ability.



GROUP TASK Activity

Organise your assessment task due dates and details into a form that assists you to efficiently schedule your time. How have you structured the data and how has each data item been represented?

ANALYSING

Analysing is the information process that transforms data into information. It interprets the data, so it makes sense to people and they can understand it. Analysing is the process of methodically examining the data to study its contents and interrelationships. It includes such processes as: searching, selecting, sorting and comparing data; some possible aims being to identify trends, model or simulate a scenario or to study the effects of change. The resulting information is then displayed in such a way that it can be understood and used to increase knowledge. For example, graphs and charts are often used to describe trends. These tools, compared to tables of information, visually describe trends and hence better facilitate the acquisition of knowledge.



Analysing

The information process by which data is interpreted, transforming it into information.

Some examples of analysing include:

- Searching for all clients who have not made a purchase in the past 3 months.
- Sorting student's results in an exam to determine their ranks.
- Comparing the contents of two files to determine differences.
- A hotel information system automatically allocates vacant rooms at check in.
- Predicting future sales based on past sales data to assist in estimating future company profits.
- Graphing minimum and maximum daily temperatures for the past 12 months.

Notice that in each of the above examples, the information is generated from the data, however the data itself is not altered. This is true for all types of analysis; the data remains unchanged, that is, the analysis process transforms the data to produce information, but does not modify the data.

The information returned after analysis will only be accurate if the data used is known to be accurate and complete. For example, if there are 3 classes studying IPT at your school and only 2 classes' results have been entered then sorting these results to determine ranks will yield incorrect information. Incorrect or dubious information results from analysis that does not take account of all factors influencing the outcome. It is often not possible to consider all relevant data, and hence the information resulting from analysis will not be precise. In these cases the information is used as a guide for decision making. For example, predicting a company's future profits when new major competitors have entered the marketplace and their effect is not fully known. The information is based on the best available evidence and hence is useful as a guide to management.



Consider the following:

When a new housing estate is opened various extra government services must be established; such as schools, hospitals and transport. Various data sets from a variety of different sources are used to predict the timing, location and size of each service, however despite the accuracy of the data sets the predictions are often inaccurate. The reality of the situation is such that forward planning is required and governments must base their planning decisions and priorities on information of some sort.



GROUP TASK Discussion

Discuss different sources and types of data that could be used. How could this data be analysed to provide the information necessary to plan the provision of new government services? What factors, do you think, influence the accuracy of this information?

STORING AND RETRIEVING

The ability to store and retrieve data is central to the activities of all information systems. Without this ability it would not be possible to reuse data without it continually needing to be re-entered. Before data can be stored it must be organised into a suitable format, similarly any processes that will later retrieve the data must understand this format. For example, a graphic saved as a jpeg file can only be retrieved successfully by an application that understands the format of jpeg files. Storing and retrieving does not modify the data, rather it represents the data in a form that is suitable to the storage device; for example, CD-ROMs represent data as microscopic pits, whereas RAM chips use different levels of voltage to represent data.



Storing and retrieving

The information process by which data and information is saved and accessed later.

Storing is the process of copying or saving data onto a storage device and retrieving is the process of reloading previously stored data. Storage devices can store data permanently (non-volatile storage) or temporarily (volatile storage). Examples of permanent storage devices include hard disk drives, floppy disks, tape, optical disks, flash memory, and even filing cabinets and other paper-based media. Permanent storage means the device does not require any type of energy to maintain the data, the storage is stable or non-volatile. For example, a hard disk stores data magnetically; the

data remains when the power is turned off. This is in contrast to volatile or temporary memory, such as random access memory (RAM), where electrical energy is required to maintain the data.

To successfully store or retrieve data requires information in regard to:

- The location of the storage device.
- The format of the data.
- How to communicate with the storage device.
- Methods used to secure and protect the data.

In Chapter 6, we answer these questions in regard to various different storage devices and their related software.



GROUP TASK Activity

There are many different file types that are used when storing and retrieving data, for example, jpeg, gif, txt. Create a list of as many different file types as you can. Classify each of these file types as holding image, video, audio, text and/or numeric data.

PROCESSING

There are seven information processes discussed in this course, as all of them are processes then surely they all perform processing? This is of course true, a process is a series of actions that bring about some result, all seven information processes clearly do this. For our purposes we shall confine the ‘processing process’ to encompass those actions that change data. More precisely: processing is the information process that manipulates data in various ways to produce a new value or result.

Processing is the only information process that alters the actual data present in the system. For example, at the conclusion of each school year the front office updates the current year level for all students. Those in year 7 are updated to year 8, year 8 is updated to year 9, year 9 to year 10, year 10 to year 11, year 11 to year 12 and finally the year 12 records are archived and removed. We define this to be a processing task as the data itself is altered. Other information processes may alter the way the data is represented but they do not change the data itself, that is, no data is lost and no data is changed.



Processing

The information process by which data can be manipulated in different ways to produce a new value or result.

As with each of the information processes, processing often occurs as an integral part of another information process. As data is collected, it is common for alterations to be made to existing data to reflect the new data. For example, making a withdrawal from an ATM results in changes to the account balance; calculating the new account balance is a processing task. When image data is organised into a format suitable for saving it is common to compress the data using techniques that alter the original image; as this action alters the data it is considered to be a processing task.

The processing process, in a computer-based information system, is performed by the central processing unit (CPU) in conjunction with primary memory. The speed of the CPU and its related resources is crucial to the efficiency of processing. In Chapter 7, we examine the hardware used for processing, such as the CPU and RAM, and specific features that improve processing performance. Of course, processing can also be performed using non-computer tools, it is just that computers are particularly well

suited to processing tasks because of their incredible speed and ability to follow procedures precisely with virtually total accuracy.



Consider the following scenarios:

- Sending an email message.
- Editing an essay using a word processor.
- Paying for goods using EFTPOS.
- A payroll system calculating tax on each employee's weekly salary.



GROUP TASK Activity

For each of the above, list any processes occurring that produce a new value or result by altering or updating data.



GROUP TASK Discussion

Each of the above dot points describes a scenario that makes use of a computer. Discuss appropriate alternative methods of processing should the computer fail.

TRANSMITTING AND RECEIVING

Transmitting and receiving is the information process that transfers data and information within and between information systems. Transmitting is the process of sending data or information and receiving is the process of acquiring data or information. Both these processes allow for communication between different devices, these devices may be components within a single computer or the devices themselves maybe different computers. For example; transmitting and receiving occurs between the CPU and random access memory (RAM) and it also occurs between a home computer and other remote computers using the Internet. The communication could also be between non-computer devices, such as telephones, mail, radio, television or even speech.



Transmitting and receiving

The information process that transfers data and information within and between information systems.

All successful communication requires 3 basic components, namely, a sender, a medium and a receiver. The sender encodes the message and transmits it over the medium. The receiver subsequently receives the message via the medium and decodes it. For example, when having a conversation messages are encoded into sound waves, which are sent using the air as the medium, the receiver uses their ear to detect and then decode these sound waves. The encoding process organises the data into a form suitable for transmission over the medium. In our speaking example, language is transformed by the sender's voice box into sound waves. Similarly, the

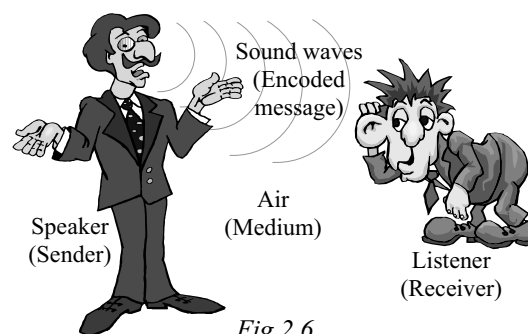


Fig 2.6

Speech is an example of transmitting and receiving.

receiver must retrieve the encoded message from the medium and make sense of it; in our example, the receiver's ear detects the sound waves, and their brain decodes these sound waves back into language.

Successful communication only occurs when messages are received accurately and on time. Both sender and receiver must understand the precise nature of the transmission together with when each transmission will commence and end. For this to occur requires both parties to agree on the way the data is represented as well as how the data is to be transferred across the medium. Important considerations include:

- The direction of the transfer.
- Format of the data.
- Speed of the transmission.
- Rules governing the transmission.
- Methods for directing messages to their destination.
- Techniques for dealing with communication errors.

In Chapter 8, we examine each of the above points in some detail.



Consider the following:

Sending an email is, on the surface, a transmitting process, however it actually includes all of the seven basic information processes. Typing the message, entering the email address of the recipient and hitting the send button are all part of the collecting process. The message is then organised into a suitable format for transmission to the sender's mail server. If an ADSL connection is used then the message is modulated for transmission via the phone lines. Once the message arrives at the mail server it is decoded back into its original digital form. The mail server then analyses the recipients email address and determines the address of the associated mail server. If the mail server's address is successfully recovered then the message is reorganised into packets in preparation for transmission onto the Internet. Eventually the message reaches its destination mail server where the whole process is repeated in reverse. Finally the message is stored on the recipient's computer and appears in their inbox. Amazingly this all seems to work, most of the time!

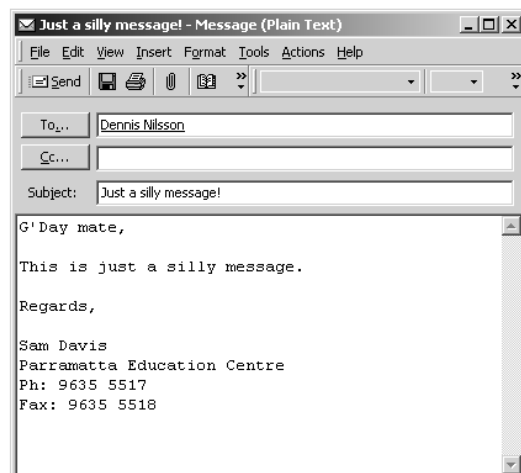


Fig 2.7

Sending an email is, on the surface, a transmitting process, however it really includes all seven information processes.



GROUP TASK Discussion

Identify each action in the above scenario as one of the seven information processes. As all seven information processes are occurring then is it reasonable to classify sending an email as a transmitting process? Discuss.

DISPLAYING

The word display, in terms of computers, usually implies a screen or monitor. So the process of displaying would mean the act of presenting information on a screen. In this course displaying has a far broader meaning, in fact it's meaning is closer to the general meaning of the word displaying. Displaying means to show, to put into view or to exhibit. This is primarily what the displaying process does; it puts information on show so people can view it. For our purposes the displaying process outputs information from an information system for presentation to the user (or participant). The information could be any combination of text, graphics, video, sound or any other type of output. The displaying process is vital to the achievement of the system's purpose, it controls what the end users see. The displayed information provides a window into the system for users, it is their only view of the system and hence its impact is significant.



Displaying

The information process by which information is output from the system to meet a purpose. It controls the format of the information presented to the participant or user.

To display information requires decisions in regard to the form in which the information will be displayed. Questions such as how text will be formatted, what resolution is needed for an image, or the most suitable volume for playing audio must be considered prior to actually displaying this information. Other questions will relate to the hardware that is to be used for displaying. For example, will a video be played on digital hardware or analog hardware, is the information designed to be displayed on a monitor or printer, if printed then what resolution is needed? The information needs to be displayed in a manner that will best achieve the purpose of the information system.

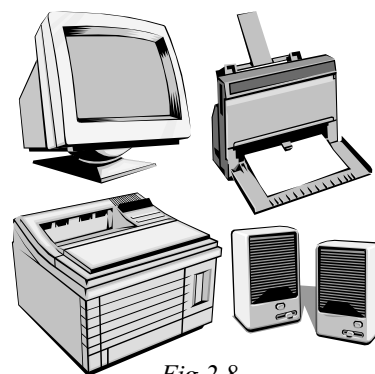


*Fig 2.8
Some common display devices:
a monitor, inkjet printer, laser
printer and speakers.*



Consider the following:

Each of the following scenario's fundamental purpose is to display information, however various other information processes are used to achieve this purpose.

1. Designing a personal web page.
2. Formatting a school assignment.
3. Recording a voice mail greeting.
4. Creating a graph to convey the results of a survey.
5. Taking a video of a friend's wedding.



GROUP TASK Activity

For each of the above, identify the information processes occurring that lead to the final display of information.



GROUP TASK Activity

For each scenario, identify likely hardware and/or software tools that would be used to display the final information.



HSC style question:

A local preschool intends to install digital video cameras within each of its rooms. The video is to be broadcast live on the Internet so parents can monitor their children throughout the day.

- (a) Identify the information technology required for such a system.
- (b) Identify and outline the significant information processes occurring within this system.
- (c) Many of the preschool teachers and some of the parents are strongly against the installation of the video cameras. Describe the likely nature of their concerns.

Suggested Solution

- (a) Video cameras in each room, computer in each room, webserver (or streaming server) software, high performance computer to run web or streaming server software, LAN hardware and software to link each classroom to the high performance machine, high speed Internet connection.
- (b) Video data is collected using the video cameras. This data is processed and compressed on each classroom computer and then transmitted over the LAN to the main computer. The main computer receives requests from machines on the Internet and responds by streaming the desired video to that machine where it is displayed.
- (c) Preschool teachers would be concerned about their own and their student's rights to privacy. Knowing that parents, and perhaps other unknown persons, are viewing the classroom would be intimidating for teachers. They will feel their every word and action is being watched and would be concerned about possible issues parents may have with their classroom techniques. Parents would not like other parents and other persons to be able to monitor the activities of their children, particularly if their child misbehaves. There may also be legitimate concerns in regard to paedophiles watching and targeting children.

Comments

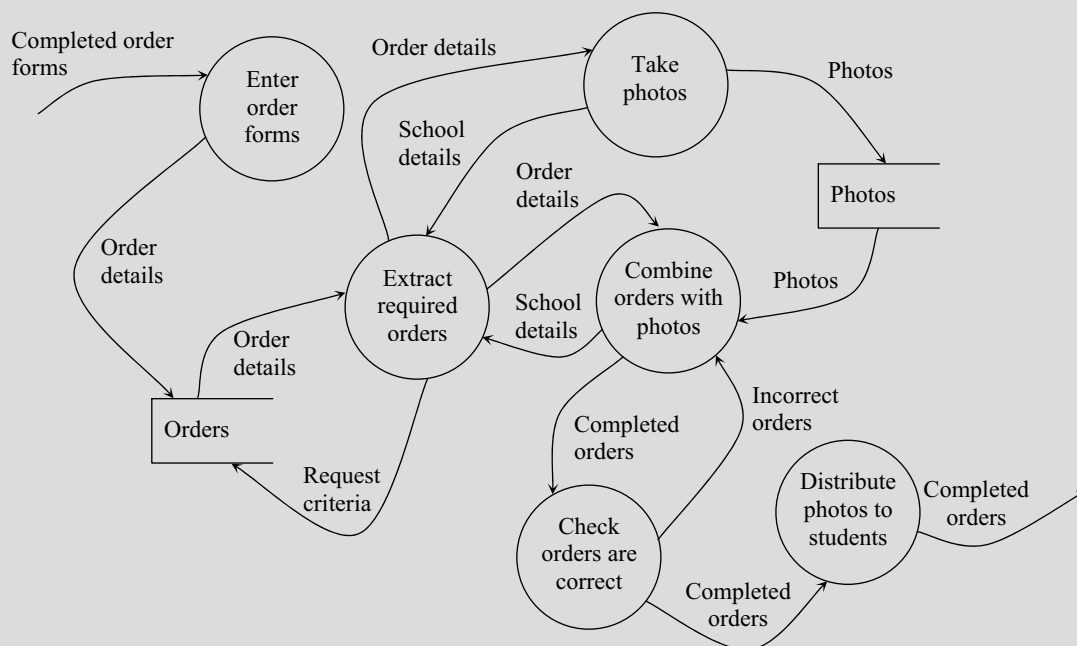
- (a) Assumptions about the pre-school computers and web or streaming server are necessary for this system even though they are not specifically detailed in the question. Simpler information technology could also have been described. In an examination this question would likely attract 2 or 3 marks.
- (b) The suggested solution includes examples of collecting, processing, transmitting, receiving and displaying. The processes described relate to a system that uses the information technology identified in part (a). In an examination this question would likely attract 3 or 4 marks.
- (c) It is important to address the concerns of both parents and teachers. In an examination this question would likely attract 3 or 4 marks.

SET 2B

1. The information process that arranges and represents data is:
 - (A) Organising.
 - (B) Analysing.
 - (C) Storing and retrieving.
 - (D) Processing.
2. The only information process that alters the actual data is:
 - (A) Organising.
 - (B) Analysing.
 - (C) Storing and retrieving.
 - (D) Processing.
3. The information process that transforms data into information is:
 - (A) Organising.
 - (B) Analysing.
 - (C) Storing and retrieving.
 - (D) Processing.
4. Determining the maximum value within a set of values could best be described as an:
 - (A) Organising process.
 - (B) Analysing process.
 - (C) Processing process.
 - (D) Displaying process.
5. An existing customer's name is entered, the customer's record is then located and deleted from a database located on a file server, finally a confirmation message is generated. This description includes:
 - (A) collecting, analysing, processing and displaying
 - (B) all seven information processes.
 - (C) all information processes except organising.
 - (D) all information processes except analysing.
6. A query returns a set of records that meet certain criteria. The main information process occurring is:
 - (A) Organising.
 - (B) Analysing.
 - (C) Transmitting and receiving.
 - (D) Displaying
7. During a normal telephone conversation the main information processes are:
 - (A) collecting and displaying.
 - (B) collecting, transmitting and receiving, and displaying..
 - (C) collecting, processing and displaying.
 - (D) collecting, and transmitting and receiving.
8. Non-volatile storage:
 - (A) requires energy to maintain the data.
 - (B) is used to hold instructions and data during processing.
 - (C) is permanent and does not require energy to maintain the data.
 - (D) is often called main memory.
9. A raw image file is compressed and saved onto a local hard disk. The main information processes being used are:
 - (A) collecting, organising and displaying.
 - (B) analysing, processing and storing.
 - (C) processing, transmitting and storing.
 - (D) organising, processing and storing.
10. The displaying process results in:
 - (A) output to a monitor.
 - (B) output directed to any of the other information processes.
 - (C) output of any type directed to an entity outside the system.
 - (D) the output of information from the system in a form suitable for humans.

11. For each of the following information processes, describe the general nature of the actions taking place:
- Collecting
 - Organising
 - Analysing
 - Storing and retrieving
 - Processing
 - Transmitting and receiving
 - Displaying
12. Classify each of the following scenarios according to the information process that best describes the actions taking place. Justify your answer in each case.
- A driving instructor completes a student's Learner Driver Log Book at the completion of each driving lesson.
 - Email software downloads new messages from a mail server.
 - The sound card in a computer converts digital signals to analog before sending them to a speaker.
 - Universities calculate ATARs based on HSC results from the Board of Studies.
 - A retailer increases the price of all their products by 5%.
13. A pocket calculator can be thought of as an information system. Identify and describe the information processes occurring during a simple calculation such as $2 + 9$.

Refer to the dataflow diagram below when answering Question 14 and 15.



14. Describe the processes that occur once a student completes their order form.
15. What syllabus information processes are occurring during each process on the dataflow diagram? Explain your answers.

THE NATURE OF DATA AND INFORMATION

In Chapter 1, we discussed data as the raw material for an information system; the data being the input to an information system. In the last section, we examined the seven information processes that operate on the input data to transform it into information; information being the output from an information system that is both meaningful and understandable. This information coming out of one information system can then be used as the data going into a further information system.

The data, and information, used by an information system is of various types, each being suited to different tasks. In this section we consider the different types of media commonly used as data, and information, within information systems. For each media type, we examine the nature of the data and how the data is represented digitally. Digital data is data that is coded using numbers, in the case of digital computers binary numbers are used. Binary is the base two number system; this means it uses just two digits, namely 0 and 1. These binary digits are known as bits. Ultimately all data and information is represented within computers as a series of bits.

DIFFERENT TYPES OF MEDIA

Media is the plural of medium; in the context of information, media refers to something in the middle that is used to transmit a message of some sort. This is what the press does; it transmits news, a form of information, using television, radio or print media. The term ‘multimedia’ is used to refer to information that combines text, sound, graphics and/or video. For example, the worldwide web makes extensive use of multimedia; the types of media used are chosen to best communicate the intended information.

In this section we consider different types of media commonly used by information systems, namely:

- text,
- numbers,
- image,
- audio and
- video.

These media provide a method for representing data and communicating information. Each media type conveys different information and is used to represent different types of data, yet computers represent all types of media in binary. Binary is a number system, just like the familiar decimal system, except rather than ten digits it uses only two, namely 0 and 1. Computers ultimately represent all the different types of media as a sequence of 0s and 1s. It is the way this data is organised that makes it meaningful and therefore able to be transformed into information.



Consider the following:

This book is primarily composed of text, hence the name textbook. In reality it uses other media, together with text, to communicate information. During your IPT studies your teacher uses this textbook together with other media to teach IPT.



GROUP TASK Discussion

Identify different types of media used during your IPT classes. For each type discuss advantages and disadvantages compared to straight text.

Text

The text media type is used to represent characters. These characters can be printable, such as letters of the alphabet, or non-printable, such as carriage returns or tabs. A sequence of characters is used to represent words, paragraphs or complete books, however text can also be used for many other purposes, for example, phone numbers are usually represented as text, as the sequence in which the digits appear is vital, yet each identical digit's meaning is the same.

What makes data a candidate for the text media type? Any data that is composed of a string of distinct characters where the order of the characters is important but each character, when considered in isolation, has a constant meaning regardless of this order. For example, the string of characters 'The cat sat on the mat.' is composed of 23 distinct characters, the meaning is derived as a consequence of the order in which these characters appear, yet each occurrence of say, the letter 'a', has the same meaning. In contrast consider the number 2320, the first occurrence of the character '2' means 2 thousand and the second means 2 tens. Numbers are therefore not good candidates for the text media type.

There are numerous methods for representing text digitally; all these methods code each unique character into a number. The two most commonly used methods are ASCII (American Standard Code for Information Interchange) pronounced *as-kee* and EBCDIC (Extended Binary Coded Decimal Interchange Code) pronounced *ebb-see-dik*. IBM mainframe and mid-range computers, together with devices that communicate with these machines, use EBCDIC. The ASCII system of coding text is used more widely and has become the standard for representing text digitally.

Standard ASCII represents each character using a decimal number in the range 0 to 127. This range is used as each character can then be represented in binary using just seven bits (binary digits). The table in Fig 2.9 shows the standard ASCII character set together with the decimal code for each of these characters. We can see in this table that the decimal number 65 represents 'A', 65 in decimal is equivalent to the seven bit binary number 1000001. The text 'The cat sat on the mat.' would likewise be represented in ASCII as 84 104 101 32 99 97 116 32 111 110 32 116 104 101 32 109 97 116 46 and in binary as a sequence of 23 seven bit binary numbers. Notice that in ASCII each of the characters in the alphabet are arranged in order, as are the digits, this greatly simplifies the sorting of text into alphabetical order. Also, the non-printable characters occupy the decimal values from 0 to 31.

Char	Dec	Char	Dec
NUL	0	@	64
SOH	1	A	65
STX	2	B	66
ETX	3	C	67
EOT	4	D	68
ENQ	5	E	69
ACK	6	F	70
BEL	7	G	71
BS	8	H	72
HT	9	I	73
LF	10	J	74
VT	11	K	75
FF	12	L	76
CR	13	M	77
SO	14	N	78
SI	15	O	79
DLE	16	P	80
DC1	17	Q	81
DC2	18	R	82
DC3	19	S	83
DC4	20	T	84
NAK	21	U	85
SYN	22	V	86
ETB	23	W	87
CAN	24	X	88
EM	25	Y	89
SUB	26	Z	90
ESC	27	[	91
FS	28	\	92
GS	29	]	93
RS	30	^	94
US	31	_	95
Space	32	`	96
!	33	a	97
"	34	b	98
#	35	c	99
\$	36	d	100
%	37	e	101
&	38	f	102
'	39	g	103
(	40	h	104
)	41	i	105
*	42	j	106
+	43	k	107
,	44	l	108
-	45	m	109
.	46	n	110
/	47	o	111
0	48	p	112
1	49	q	113
2	50	r	114
3	51	s	115
4	52	t	116
5	53	u	117
6	54	v	118
7	55	w	119
8	56	x	120
9	57	y	121
:	58	z	122
;	59	{	123
<	60		124
=	61	}	125
>	62	~	126
?	63	DEL	127

Fig 2.9
The ASCII character set.



Consider the following:

The ASCII table in *Fig 2.9* shows the decimal code for each character, but in reality computers represent these numbers using binary. Binary is the base 2 number system whereas decimal uses a base of 10. The decimal number 465, means 4 hundreds, 6 tens and 5 ones. Hundred, ten and one are all powers of ten, namely 10^2 , 10^1 and 10^0 , so $465 = (4 \times 10^2) + (6 \times 10^1) + (5 \times 10^0)$. In binary rather than powers of ten we use powers of two, hence the binary number 1101 in decimal really means $(1 \times 2^3) + (1 \times 2^2) + (0 \times 2^1) + (1 \times 2^0) = 8 + 4 + 0 + 1 = 13$. As computers generally work on groups of 8 bits, called a byte, it would be common to see the binary number 1101 written as 00001101, this is similar to writing 465 as 00000465, any leading zeros can be ignored.



GROUP TASK Activity

The following 8 bit binary numbers are used to represent a portion of text using standard ASCII. What does it say? Once you work it out you have my permission to call out your answer!

```
01001001 01000000 01101100 01101111 01110110 01100101 01000000
01001001 01010000 01010100 01000001
```



GROUP TASK Discussion

Is the sequence of binary numbers in the above activity data or information? Discuss.

Numbers

The number media type is used to represent integers (whole numbers), real numbers (decimals), currency and even dates and times. In fact any quantity that can be expressed on a numerical scale can be represented using numbers; ask yourself, is it possible to place a single example of this data on an ordered continuous line and is it possible and desirable to perform mathematical operations with this data? If the answer to these questions is yes then this data is a prime candidate to be represented as a number. Numbers have magnitude, that is, the concept of size is built into all numbers, for example, '15 is bigger than 10 but smaller than 20' describes the magnitude of 15. The digits that make up numbers have different meanings dependant on their position relative to other digits in the number. These attributes are not present in the other types of media. For example, images do not have magnitude and nor does text, to say that a photograph of a bird is greater than one of a building or to say this sentence is greater than the last is meaningless.

Ultimately all data stored and processed by digital computers is represented as numbers. Computers, at their most basic level, process binary numbers by adding and comparing them, consequently all media types must be represented and processed as binary numbers.

```
456
-345
16.0004440550066
-0.002
$65.45
$5,000,000
11/07/2003 4:44:47 PM
11-July-03
```

Fig 2.10
Data suitable for use by the
number media type.

Computers are finite devices, they cannot represent or calculate every possible number, there is a limit to the accuracy with which they represent and calculate numbers. As a consequence the manner in which they represent numbers is a compromise between space, speed and accuracy. As the needs of different information systems and their processes require different types of numbers and different levels of accuracy various different methods of representing numbers are in common usage. For example, if we are counting the number of cars that pass by a given point then our data is positive whole numbers; we have no need to store decimal fractions. If we are calculating the average of a set of numbers then the fractional part of the answer is significant and a real number representation method is required.

Let us briefly consider the storage requirements, range, strengths and limitations of commonly used methods for representing integers, real numbers, currency and dates/times:

- **Integers**

Commonly integers are represented using the two's complement system, this system codes the sign of each number in such a way that binary calculations need not consider the sign of the numbers. Each integer is represented using either 16 bits or 32 bits; the range for 16 bit integers is from -32768 to 32767 and for 32 bit integers from -2147483648 to 2147483647 . Whole number calculations within these ranges are perfectly accurate, however calculations outside the range are not possible. Any calculations resulting in fractional answers cannot be stored as integers without loss of the fractional part. For example, the result of simple divisions, such as 2 divided by 4, cannot be stored as they are not whole numbers.

- **Real numbers**

Real numbers are commonly represented using a system known as 'floating-point'. Floating-point numbers are represented using a technique similar to scientific notation. For example, 1234.5678 is written in scientific notation as 1.2345678×10^3 , 1.2345678 is called the mantissa and the 3 is known as the exponent; the position of the decimal point changes (or floats) depending on the value of the exponent. There are two common standards; single precision floating-point which represents each number using 32 bits and double precision floating-point which uses 64 bits. Single precision has an approximate range of -3.4×10^{38} to 3.4×10^{38} and double precision has an approximate range of -10^{308} to 10^{308} . Be aware that not all numbers within these ranges can be represented precisely, even simple fractions, such as $\frac{1}{3}$, have no exact floating-point equivalent. Single precision representations are accurate to around 7 significant figures and double precision to 15 significant figures, therefore in single precision $\frac{1}{3}$ is represented as 0.3333333 and in double precision as 0.3333333333333333 , be aware that repetitive calculations can multiply inaccuracies significantly. Floating-point calculations are more processor intensive than integer calculations; consequently most CPU designs include a dedicated floating-point unit (FPU).

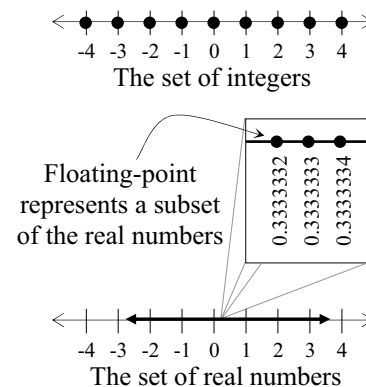


Fig 2.11
Integers, real numbers and floating-point.



GROUP TASK Investigation

Investigate the accuracy of calculations performed by a spreadsheet with which you are familiar. What type of representation do you think is being used for numbers?

- **Currency**

Financial calculations require very precise calculations but within a relatively restricted range. For most currency calculations accuracy must be perfect up to two decimal places. To achieve these requirements a system similar to integer representation is used but with the decimal point moved four places to the left; essentially integers are scaled by a factor of 10000. This results in a representation that is accurate to the required two decimal places. Commonly each data item is represented using 64 bits (8 bytes), resulting in an effective range of – 922,337,203,685,477.5808 to 922,337,203,685,477.5807. Every decimal number with up to four decimal places can be represented precisely within this range.

- **Dates/Times**

Many older systems coded dates and times using separate numbers for the day, month, year and time, it is now common for a single date and time to be represented as a double-precision floating-point number. For example, 37816.25 converts to 6am on the 14/7/2003, the whole number part is the number of days that have elapsed since the 30/12/1899 and the fractional part is the fraction of the day that has elapsed. The method of representation is identical to the double precision floating-point system; this is the way dates/times are organised. The analysing process transforms these numbers into dates and times that we humans understand.


GROUP TASK Activity

Using a spreadsheet, enter various numbers and then format them as dates and times. Verify if the system used is the same as the one outlined above.

Images

The image media type is used to represent data that will be displayed as visual information. Using this definition all information displayed on monitors and printed as hardcopy is represented as images. This is true, all monitors and printers are used to display image media, however text and numbers are organised into image data only in preparation for display. Photographs and other types of graphical data are designed specifically for display; this is their main purpose. In these cases the method of representing the image is chosen to best suit the types of processing required. For example, the representation used when editing a photograph to be included in a commercial publication is different to that used when drawing a border around some text in a word processor. There are essentially two different techniques for representing images; bitmap or vector; let us consider each of these in turn.

- **Bitmap**

Bitmap images represent each element or dot in the picture separately. These dots are called pixels (short for picture element) and each pixel can be a different colour and is represented as a binary number. The number of colours present in an image has a large impact on the overall size of the binary representation. For examples, a black and white image requires only a single bit for each pixel, 1 meaning black and 0 meaning white. For 256 colours, 8 bits are required for each pixel so the image would require 8 times the storage of a similarly sized black and white bitmap image. Most colour images can have up to 16 million different colours, where each pixel is represented using 24 bits. The number of bits per pixel is often referred to as the image's colour depth; the higher the colour depth, the more colours it includes and the larger the storage requirements for the image will be.

The other important parameter in regard to bitmap images is resolution. The resolution is the number of pixels the image contains and is usually expressed in terms of width by height. The image of the Alfa Romeo in *Fig 2.12* has a resolution of 505 pixels by 391 pixels, when the image is enlarged each pixel is merely made larger, e.g. the jaggy looking grille inset at the top right of the photo. When using bitmap images it is vital to consider the likely display device to be used to determine the resolution required.



Fig 2.12

The resolution of bitmap images should be appropriate to the display device.

Bitmap images are often compressed to reduce their size prior to storage or transmission. Many different bitmap image file formats are available; some reduce the size of the image file without altering the image (lossless compression) whilst others alter the image data as part of the compression process (lossy compression). For example the Alfa Romeo image in *Fig 2.12* takes up 578 kilobytes when stored as a standard uncompressed Windows BMP file and only 28.4 kilobytes when stored using lossy compression as a JPEG file.



GROUP TASK Investigation

Load a photograph into a photo editor such as MS-Paint. Save this image using different formats and colour depths. Observe and document the differences in terms of storage size and clarity of the resulting images.

• Vector

Vector images represent each portion of the image mathematically. That is, the data used to generate the image is a mathematical description of each shape that makes up the final image. Each shape within a vector image is a separate object that can be altered without affecting other objects. For example, a single line within a vector image can be selected and its size, colour, position or any other property altered independent of the rest of the image. For example, the body of the cat in *Fig 2.13* has been drawn using a single filled line whose attributes can be altered independently from the rest of the image.



Fig 2.13

Vector images are represented as separate editable shapes.

The total size of the data required to represent a vector image is, in most cases, less than that for an equivalent bitmap image however the processing needed to transform this data into a visual image is far greater. In fact all vector images must be transformed into bitmaps before they can be displayed on a monitor, printer or any other output device. Vector images can be resized to any required resolution without loss of clarity and without increasing the size of the data used to represent the image. Vector graphics are generally unsuitable for representing photographic images, as the detail required is difficult to reproduce mathematically.

Audio

The audio media type is used to represent sounds; this includes music, speech, sound effects or even a simple ‘beep’. All sounds are transmitted through the air as compression waves, vibrations cause the molecules in the air to compress and then decompress, this compression is passed onto further molecules and so the wave travels through the air. Our ear is able to detect these waves and our brain transforms them into what we recognise as sound. The sound waves are the data and what we recognise as sound is the information.

All waves have two essential components, frequency and amplitude. Frequency is measured in hertz (Hz) and is the number of times per second that a complete wavelength occurs. Sound waves are made up of sine waves where a wavelength is the length of a single complete waveform, that is, a half cycle of high pressure followed by a half cycle of low pressure. In terms of sound, frequency is what determines the pitch that we hear, higher frequencies result in higher pitched sounds and conversely lower frequencies result in lower pitched sounds. The human ear is able to discern frequencies in the range 20 to 20,000 Hz, for example, middle C has a frequency of around 270 Hz.

Amplitude determines the volume or level of the sound, very low amplitude waves cannot be heard whereas very high amplitude waves can damage hearing. Amplitude is commonly measured in decibels (db). Decibels have no absolute value; rather they must be referenced to some starting point. For example, when used to express the pressure levels of sound waves on the human ear, 0 decibels is usually defined to be the threshold of hearing, that is, only sounds above 0 decibels can be heard, sounds above 120 decibels are likely to cause pain.

Let us now consider how audio or sound data can be represented in binary. There are two methods commonly used, the first is by sampling the actual sound at precise intervals of time and the second is to describe the sound in terms of the properties of each individual note. Sampling is used when a real sound wave is converted into digital, whereas descriptions of individual notes is generally used for computer generated sound, particularly musical compositions.

- **Sampling**

The level, or instantaneous amplitude, of the signal is recorded at precise time intervals – each sample is stored as a binary number. This results in a large number of points that can be joined to approximate the shape of the original sound wave. There are two parameters that affect the accuracy and quality of audio samples; the number of samples per second and the number of bits used to represent each of these samples. For example, stereo music stored on compact disks contains 44100 samples for each second of audio for both left and right channels and each of these samples is 16 bits long. This means that an audio track that is 5 minutes long requires storage of $44100 \text{ samples} \times 300 \text{ secs} \times 16 \text{ bits per sample} \times 2 \text{ channels}$; this equates to approximately 50MB of storage. A normal audio

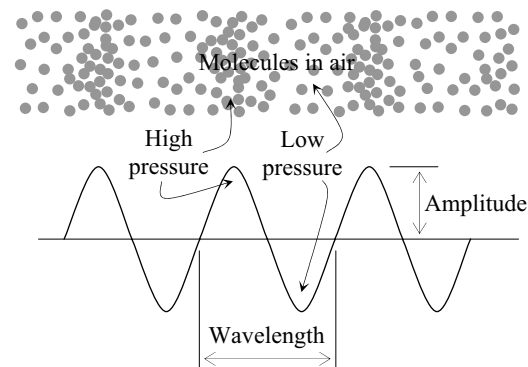


Fig 2.14
Sound is transmitted by compression and decompression of molecules.

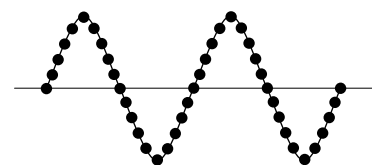


Fig 2.15
Samples are joined to approximate the original sound wave.

CD can hold about 650MB of data, therefore it is possible to store up to around 65 minutes of music on an individual CD. 44100 samples are taken each second because this ensures at least two samples for each wave within the limits of human hearing; remember humans can hear sounds up to frequencies of about 20000Hz, so 40000 samples would ensure at least two samples for all sound waves less than this frequency.

It is now common for music and other sound data to be recorded using 6 channels (surround sound), without compression these recordings require three times the storage of a similar stereo sound. Consequently various compression techniques have been devised to reduce the size of sampled sound data; however greater processing power is then required to decompress the sound prior to playback.

• Individual Notes

This type of music representation is similar to a traditional music score (see Fig 2.16). The position of each note on a music score determines its pitch and the symbol used determines its duration. Different parts of the score are written on their own staff (set of five horizontal lines), for example, in Fig 2.16 the top staff contains the notes played by the right hand and the bottom staff those played by the left hand.



Fig 2.16
Traditional music scores are represented
digitally as a series of individual notes.

In binary each note or tone in the music is represented in terms of its pitch (frequency) and its duration (time). Further information for each note can also be specified such as details in regard to how the note starts and ends, and the force with which the note is played. These extra details are used to add expression to each note. Particular instruments can be specified to play each series of notes. The most common storage format for such files is the MIDI (Musical Instruments Digital Interface) format; most digital instruments, including computers, understand this format. Extra files are available that either specify the distinct tonal qualities of a particular instrument or that contain real recordings (digital sound samples) of the instrument playing each note. These files are used in conjunction with the notes to electronically reproduce the music.

Generally, binary representations that use individual notes are significantly smaller than similar sound samples, however greater processing power is required to convert the data into information in the form of sound waves.



GROUP TASK Activity

Listen to a variety of different sounds digitised using samples (e.g. .WAV files) and digitised as individual notes (e.g. .MID files). What differences, in terms of sound quality, can you hear?



GROUP TASK Discussion

There are similarities between image bitmaps and sound samples and there are similarities between vector images and sound represented as individual notes. Discuss the similarities.

Video

The video media type combines image and sound data together to create information for humans in the form of movies or animation. To create the illusion of movement images are displayed one after the other in a particular sequence. Images entering the human eye persist for approximately one twentieth of a second, therefore for humans to perceive smooth movement requires displaying at least 20 images per second, most movies are recorded at 24 frames per second. Video data is composed of multiple images together with an optional sound track. The images and sound must be synchronised for the overall effect to work convincingly. All this information must be represented in such a way that it can efficiently be displayed as video information for humans.

Motion pictures, as viewed in cinemas, largely use 35mm wide photographic film to represent the images. Each image or frame measures approximately 35mm wide by 19mm high, hence each second of the movie requires a piece of film $24 \times 19\text{mm} = 456\text{mm}$ long. Let us consider the length of film required for a two hour movie; there are $2 \times 60\text{min} \times 60\text{sec} = 7200\text{sec}$ in two hours and each second requires 0.456m of film, so the total length for the film is $0.456 \times 7200 = 3283.2\text{m}$ or approximately 3.2832km of film. The sound track for the movie is stored, in digital, alongside the images, commonly three different formats are included; Dolby stereo, Dolby surround and Sony surround sound (see Fig 2.17). Notice that the video information is stored as completely separate images, the sound as a sequence of sound samples and it is all synchronised by its sequence and location on the film.

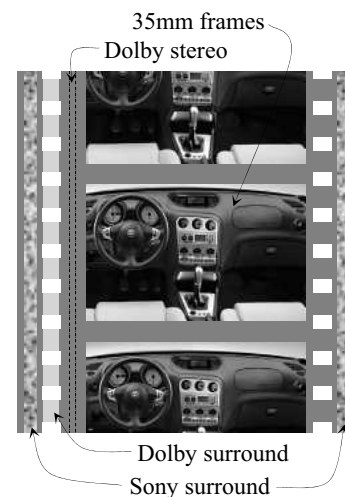


Fig 2.17
Typical 35mm film used
for motion picture.

Let us now consider techniques used to represent video in binary. Essentially video data is a combination of multiple images combined with a sound track. The images, in their raw form, are represented as bitmaps; this results in enormous amounts of data. Consider 1 minute of raw video; if there are 24 frames per second then 1440 frames ($24 \text{ frames/sec} \times 60 \text{ sec}$) or bitmaps are needed. If each bitmap has a resolution of 640 by 480 pixels and each pixel is represented using 3 bytes (24 bits) then a single minute of video requires a staggering 1,327,104,000 bytes, or more than 1.2GB of storage (see Fig 2.18). Plus we have neglected to include the sound track; the sound track uses sound samples, so if the sound track were recorded at CD quality we'd need to add a further 5MB or so. A two-hour movie, even at this rather meagre resolution, would therefore require approximately 150 gigabytes of storage. Clearly this data, particularly the images, must be represented more efficiently.

$$\begin{aligned}
 \text{Total Frames} &= 24 \text{ frames/sec} \times 60 \text{ sec} \\
 &= 1440 \text{ frames} \\
 \text{Data/frame} &= 640 \times 480 \text{ pixels} \times 3 \text{ bytes/pixel} \\
 &= 921600 \text{ bytes} \\
 \text{Total storage} &= 1440 \text{ frames} \times 921600 \text{ bytes} \\
 &= 1327104000 \text{ bytes} \\
 &= 1327104000 \div 1024 \text{ kilobytes} \\
 &= 1296000 \div 1024 \text{ megabytes} \\
 &= 1265.625 \div 1024 \text{ gigabytes} \\
 &\approx 1.2 \text{ gigabytes}
 \end{aligned}$$

Fig 2.18
Calculating the total storage for one
minute of raw video image data.

We require an efficient method of compressing and more importantly decompressing the data. Various standards exist for carrying out this process, perhaps the most common being the set of compression standards developed by the Moving Picture Experts Group (MPEG). Most of the commonly used standards utilise similar techniques to the MPEG standards, it is the detail of how these techniques are implemented that is different. Compressing video involves removing repetitive data and also removing data from parts of images that the human eye does not perceive. Some of these standards are able to compress data at a ratio of 5 to 1 whilst others can compress by as much as 100 to 1. Compression is somewhat of a balancing act; too much compression and the quality of the video deteriorates noticeably, not enough and the size of the file will be too large.



GROUP TASK Investigation

Examine the size of various video files together with their duration, resolution and colour depth. Calculate the compression ratio by calculating the size of the raw data and comparing it to the actual file size.

The most common technique used to compress video data is known as ‘block based coding’; this technique relies on the fact that most consecutive frames in a sequence of video will be similar in most ways. For example, a sequence of frames where a dog runs across in front of the camera will have a relatively stationary background, that is, the data representing the portions of the background not obscured by the dog is virtually the same for all frames, so why store this data multiple times? Block based coding is the process that implements this idea.

Let us consider a simple block based coding process:

- The current frame is split up into a series of blocks; each block contains a set number of pixels, commonly 16 pixels by 16 pixels.
- The content of each block is then compared with the same block in a past frame.
- If the block in the past frame is determined to be a close match then presumably no motion has taken place in that area of the frame, and a zero vector is stored as an indicator. Vectors indicate direction as well as size of movement, so a zero vector indicates no motion at all.
- Should the blocks not match then other like sized blocks, in the past frame, within the general vicinity of the original block are examined for possible matches. If a match is found then a vector is stored indicating the change in position of the block.
- If no match is found within the search area then the block in the current frame must be stored as a bitmap.

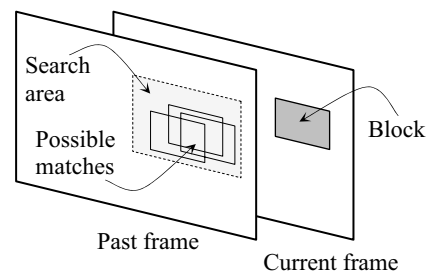


Fig 2.19

Block based coding compares blocks in each frame with those in a similar position on past frames.

Once a complete frame has been coded it is further compressed using various compression techniques commonly used for any binary data. Each frame of data is therefore represented separately but requires that past frames be known before the frame can be reconstructed and displayed. With video data this is always the case as each frame is viewed in a specific linear sequence.



HSC style question:

Currently sound recordings are almost always processed and distributed using a digital format. However the original sound is usually collected from an analog source and the final sound is ultimately output in analog form.

- (a) Distinguish between analog and digital representations of sound.
- (b) Discuss reasons why digital formats are preferred for the processing and distribution of sound recordings.
- (c) The widespread distribution of digital sound recordings has resulted in an increasing number of illegal copies of these recordings being made.

Discuss likely reasons why this increase in illegal copying has occurred.

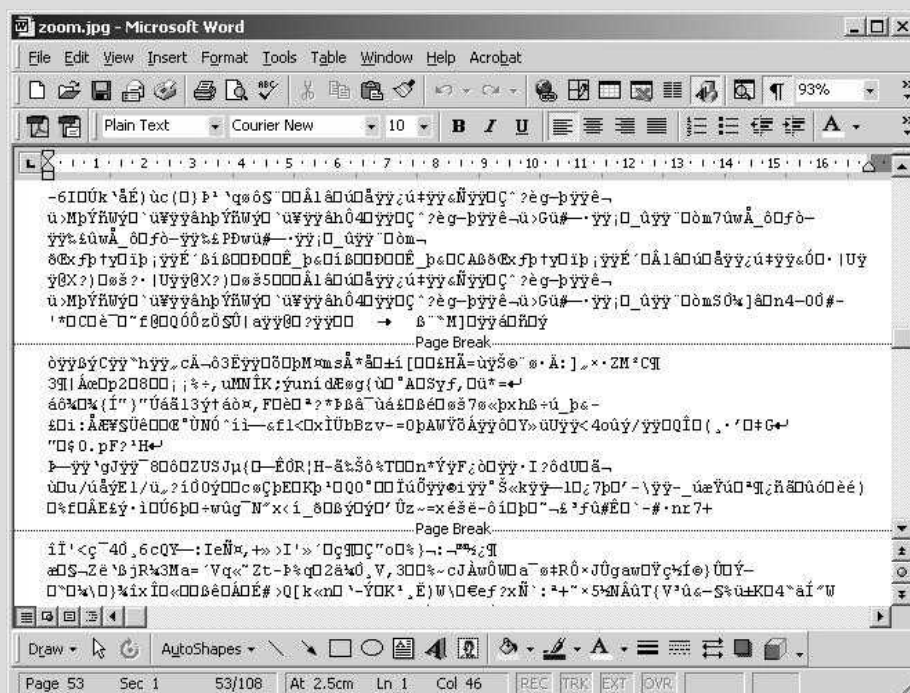
Suggested Solution

- (a) Analog sound is a continuously variable wave. In air sound is represented as a sequence of compression waves and this is converted to an analog electromagnetic wave by microphones – the frequency determines the pitch and the amplitude determines the level (or volume). Digital representations of sound are samples of the corresponding waves at regular time intervals – each sample is a binary number which represents the amplitude of the original sound wave at a particular point in time. The digital representation of the sound is a sequence of these binary numbers.
- (b) Reasons why digital formats are preferred for sound recordings include:
 - The same digital sound files can be used and played by a large variety of different hardware technologies, e.g. PCs, CD players, DVDs, MP3.
 - When processing using digital formats no new noise is introduced by the equipment. This is not true of analog sound processing technologies.
 - Distribution of digital data is precise – quality does not alter as more copies are made or as the recording ages.
- (c) Likely reasons for the increase in illegal copying include:
 - Copies of digital files are identical to the original in terms of sound quality. With analog formats, such as tape, each time a copy is made the quality of the recording deteriorates.
 - The widespread use of computers and in particular CD burners and MP3 players means copying is a simple process that can be performed anonymously by most people at home.
 - As most of the population now own the technology required to make illegal copies it is very difficult to identify and subsequently prosecute those performing illegal copying.

SET 2C

1. Digital data:
 - (A) is usually represented using the binary number system.
 - (B) includes any text, numeric, image, audio or video data.
 - (C) is easily understood by humans.
 - (D) is able to represent continuous quantities precisely.
2. EBCDIC is a system used to code:
 - (A) text.
 - (B) numbers.
 - (C) audio.
 - (D) images.
3. Postcodes in Australia always contain four digits. Postcodes would be represented as:
 - (A) numbers, as this restricts data entry to just digits.
 - (B) text, as some postcodes commence with one or more zeros and their numerical order is not significant.
 - (C) numbers, as this allows them to be used as part of mathematical calculations.
 - (D) text, as they are often combined with other textual items to form complete addresses.
4. Floating-point numbers are:
 - (A) used to represent any real number precisely.
 - (B) able to perform calculations on integers precisely.
 - (C) only suitable for representing dates and times.
 - (D) used to represent a subset of the real numbers.
5. A bitmap contains:
 - (A) a mathematical description of each shape within an image.
 - (B) a number to represent the colour of every pixel in the image.
 - (C) data that must be converted to digital prior to display.
 - (D) a series of pixels, where each pixel describes a different colour within the image.
6. Sound waves are a form of:
 - (A) digital data.
 - (B) radio waves.
 - (C) compression waves.
 - (D) light.
7. In relation to the audio media type:
 - (A) frequency determines the pitch and amplitude the volume.
 - (B) frequency determines the volume and amplitude the pitch.
 - (C) frequency determines the duration and amplitude the note.
 - (D) frequency determines the note and amplitude the duration.
8. Block based coding can best be described as:
 - (A) a method for compressing image files where some of the original information is lost.
 - (B) a sampling technique used to represent audio data in a compressed format.
 - (C) a compression technique that uses past frames in a video sequence to generate current frames.
 - (D) a system used to compress video data so that none of the original information is lost.
9. Amounts of money are commonly represented using:
 - (A) the two's complement system.
 - (B) a scaled version of the two's complement system.
 - (C) floating-point representations.
 - (D) the ASCII code of each digit.
10. When using a particular graphics program, it is possible to alter the thickness of a line without changing any other attributes of the image. The image is most likely represented as a:
 - (A) bitmap image.
 - (B) JPEG file.
 - (C) vector image.
 - (D) sampled image.

11. Convert the following binary numbers into their decimal equivalent:
 - (a) 11101101
 - (b) 10101
 - (c) 11001100
 - (d) 00011100
12. How would the word “Blonk” be represented in binary if ASCII were used as the coding system?
13. Describe how audio is represented on audio CDs.
14. The screen shot below shows the result after a JPEG image file was opened using a word processor.



Explain reasons why the JPEG file appears a bit of a mess when viewed in this way.

15. For each of the media types suggested in the following scenarios
 - Identify the media type,
 - describe a suitable form of digital representation, and
 - if necessary describe a suitable method of compression.
 - (a) Creation of a company logo for use on letterheads, folders and even the company's website.
 - (b) Composing a new piece of music.
 - (c) Removing an out of favour relative from a photograph and then emailing the photo to them.
 - (d) Preparing a small video for viewing over the Internet.
 - (e) Calculating the total a business is owed by each of its customers.

DIGITAL REPRESENTATION OF DATA

Why is data so often represented in digital form? The simple answer is that most computer based information technology only understands binary digital data, hence the data must be represented in this manner if it is to be used by these tools. This is true, however this presents us with a second question; what are the reasons why all these tools use binary digital data? In this section, we consider some answers to this question. Then in later chapters we examine, in some detail, how particular information technologies represent and process digital data.

The answers to why information technology uses digital data gives us a clear insight to the advantages of digital data, but of course there are also disadvantages. We consider some of these disadvantages. Finally, we examine the trend towards the use of digital data by considering examples from industries that have been revolutionised by their change from analog to digital data.

WHY INFORMATION TECHNOLOGY USES DIGITAL DATA (ADVANTAGES OF DIGITAL REPRESENTATION OF DATA)

- **Similar hardware design**

Binary data is made up of just two numbers, namely one and zero, this means that devices that process binary data need only be able to represent and process these two basic data items. The design of digital devices can therefore be based on similar technologies. For instance, most digital processing devices represent a zero as low voltage and a one as high voltage. Also, digital processing, at its lowest level, involves knowing how to add binary digits, knowing that a one is larger than a zero and being able to alter the state of an individual bit. These low level processes are built into the hardware of digital devices; by using them in various complex ways all the different information processes are accomplished. This means that the basic design of these information technologies can be reused in all types of digital devices.

- **Data quality**

Prior to the widespread use of digital data strategies to monitor and maintain the quality of data were time consuming and often uneconomical – physical copies of paper-based documents were needed if data was used for more than one purpose and in most cases this was simply impractical. When using digital data copies can be made effortlessly or databases can be shared and accessed by a variety of different software applications. With digital data it is now common for data changes to be reflected almost immediately across all connected systems, which improves the timeliness and accessibility of data. Also each data item, such as an individual's address, is only stored once therefore only one edit is required and all attached systems will now see the corrected data. This greatly improves the accuracy of data.

- **Ability to use different types of media**

Earlier in this chapter we learnt how all the various media types are represented using just the two binary digits one and zero, the consequence is that devices that can manipulate binary data have the potential to manipulate data from a variety of different media types. The low level processes used to manipulate binary data are the same regardless of the media type. This is not the case with most analog or non-computer based technologies; for example, a 35mm film projector cannot be used to process text or numbers, it is dedicated to displaying 35mm photographic film. This allows many digital information technologies to be multipurpose; today they can process numbers, tomorrow video and yesterday text. What's more they can be used to combine multiple media types together, for example, presentation software means

just a laptop and digital projector can replace an overhead projector, slide projector, film projector and even a blackboard.

- **Efficient data transfer**

As the data is represented using binary, and all digital devices understand binary, then it follows that transmitting and receiving data is greatly simplified. Different transmission media use different techniques for representing binary data during transmission; for example, light is used for optical fibre, voltage changes are used for communication between local devices and microwaves for many wireless transmissions. All these media are representing the same binary data, the conversion process therefore just needs to deal with transforming the data from one media onto another rather than considering the detail of the data itself. For example, a mobile phone conversation during its transmission is converted from microwave to a landline in exactly the same way as an email message or even a digital video. The ability for different digital devices to communicate effectively and without the data being degraded is a major advantage of digital data over analog representations. For example, each time a copy is made of a printed photograph some detail is lost, when a copy is made of a digital image file, the copy is identical to the original.

- **Storage of data**

Prior to the widespread use of digital data different media types were stored using quite different techniques and tools. For example, customer information would have been stored in individual files held in a bank of filing cabinets and movies were stored on photographic film. Digital representation of these media allow them all to be stored using the same technologies. That is, a database of customers can be stored on a hard disk alongside a movie file. The physical size of the storage device is relatively insignificant compared to that required previously, for example, a whole bank of filing cabinets is replaced by a single hard disk. Digital storage allows fast access to the data and it also allows the data to be reorganised and analysed in ways that were not practical using prior technologies. For example, resorting a large number of customer files by each customers address is impractical when stored in a filing cabinet, however when stored digitally this is a simple process. Watching particular scenes in a movie involved rewinding or fast-forwarding the film through the projector, when stored in a digital format we can jump directly to the required scene.

- **Speed and accuracy of processing**

Perhaps the most significant advantage of digital data is the speed and accuracy with which it can be processed. This speed and accuracy is due to the design of the integrated circuits within each processing unit. For example, a CPU operating at 2GHz is able to perform some two thousand million instructions per second, and each of these instructions is performed with virtually total accuracy. No other calculating device known to man can compete with this sort of performance.

DISADVANTAGES OF DIGITAL REPRESENTATION OF DATA

- **Not human friendly**

As humans we are not used to processing digital data, rather our brains are built to deal with various types of analog data. Our brains do not process data according to strict predefined rules and sequences; rather we make inferences and have educated guesses based on past experiences. A young child quickly learns to discriminate between dogs and cats, yet this is a very difficult task for a digital computer to accomplish. We can understand incredibly complex relationships between real world data yet we have much difficulty expressing these relationships in a logical manner

suited to digital processing. Walking across a busy street safely is something we can all do; yet representing all the data required for this task digitally is very difficult if not impossible. The conversion of real world data and processes into an equivalent digital form is suited to specific types of data and processes. Data must be ordered and the processes must follow strict rules; the human brain does not work this way.

- **Accuracy**

Digital representations of continuous data can never be as accurate as the original. Consider an audio CD; it is composed of a series of sound samples, it is not a continuous reproduction of the original sound. No matter how many times the sound is sampled the recording will never include all the detail of the original sound waves heard at a live performance. This is true for all continuous data, when it is represented digitally only a portion of the original information can ever be collected.

When using digital data we are relying on some third party to provide the information technology tools to process this data. A personal computer uses an operating system from one company, applications from other companies and also different companies have been involved in the design and production of each hardware component. It is inevitable that at times some of this technology will not operate as expected, and hence data will be processed incorrectly or could be lost completely.

- **Cost**

The information technology required to use digital data requires a large amount of expenditure up front. Manual systems grow as more data is added, for example, more folders are purchased as a company gains new customers, whereas a company with one hundred customers requires similar computer hardware as one with a few thousand customers. When the limits of the current hardware and software are reached it is necessary to again spend large amounts on new or upgraded technology.

The divide between the 'haves' and the 'have nots' continues to increase, those who have the economic means to purchase digital technologies, in particular those required to access the Internet, have access to a world of information. Those who do not have the funds to finance such technology cannot gain knowledge from this vast store of digital data. This applies equally well to entire countries as it does to individuals; companies within countries that have a significant digital data infrastructure can market to the whole world whilst those in countries without the necessary infrastructure fall further behind.

- **Security**

As binary data is so easily transferred and updated, the security of the data becomes a major concern. New technologies have emerged to deal with security concerns; for instance, virus detection and removal, backup systems and Internet firewalls are just three such technologies that are currently viewed as almost mandatory for any computer system. These digital security measures are in addition to all the existing physical measures that were used to protect the security of manual systems.

The problems with security are further exacerbated, as it is not usually obvious that the data has been copied or altered. The nature of digital data means any copies are not only identical to the original, but the original remains in the same location during the copying process. This is in contrast to most other representations where the original must be removed for copying and the copies are inferior to the original. As a consequence issues in regard to privacy and copyright become far more difficult to enforce and even detect. Who knows how often breaches occur when the breach cannot always be detected.

- **Training**

Working with computers and digital data is still a relatively new task for many people. Often people are afraid to use computers; they don't understand how they operate and they often fear damaging the equipment or the data. Such fears result from a lack of experience and knowledge; training can solve most of these problems.

Although training is necessary for all types of technology it is more crucial for most computer-based technologies. Computers are multipurpose machines, and as such they are able to perform many varied tasks. Computer users really direct the computer to perform a task rather than actually performing the task themselves, the task being performed by the computer in a split second. This means that many more different tasks can be performed and hence many more tasks must be understood. Consequently more training is required if users are to master a larger set of skills.

DATA DIGITISING TRENDS

Currently many traditionally analog systems are being digitised. In this section we consider a number of different systems that have made the change to digital data.



Consider newspapers on the Internet

Newspapers have traditionally been, as the name suggests, a printed media; as such their publication requires vast quantities of paper each day, large printing facilities, together with an extensive distribution network. Wouldn't it be nice for newspapers to rid themselves of all these costs and just publish over the Internet? Many newspapers are now publishing in digital form on the Internet however the printed version still remains, why is this?

When papers first started appearing in digital form on the Internet some concern was expressed by various newspaper publishers that the Internet would significantly reduce the need for printed versions, this concern has not been realised. Research has actually indicated that the opposite has occurred; in general the circulation of printed newspapers has continued to rise. This research indicates that the digital Internet version acts as a marketing tool, whereby customers then choose to subscribe to the related print version.

Many readers access both the printed and digital versions of the same newspaper on a regular basis. Most newspaper readers prefer the printed version when reading news items but utilise the Internet version for classifieds, in particular job advertisements. Both formats provide advantages and currently appear to be complimenting each other.



GROUP TASK Activity

Visit the Internet sites for some of Australia's large daily newspapers as well as the Internet site for a local paper in your area. Compare and contrast the information online with the printed equivalent.



GROUP TASK Discussion

Identify reasons why people may prefer printed newspapers for their news and online versions for searching classified advertisements. Do you think these reasons are sufficient to ensure printed newspapers continue? Discuss.



Consider telephone systems

The telephone system truly is amazing, just by dialling a particular sequence of numbers you are able to contact almost anyone on the face of the earth. The telephones in our homes today use essentially the same technology as those used over 100 years ago. It has a microphone, a speaker, some sort of bell and a simple switch to connect the phone to the telephone network. In fact in most homes, the wire connecting the phone to the network is essentially the same as those used for the past one hundred years, it is what happens once this wire reaches the local telephone exchange that has changed.

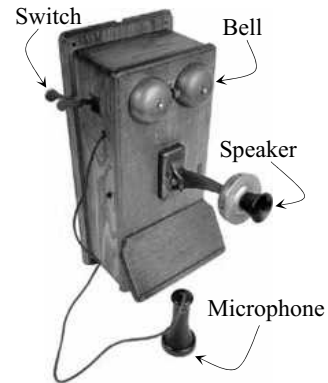


Fig 2.20

An antique telephone has similar components to today's phones.

In the past circuit switching was used to connect your home phone directly with other phones. Circuit switching creates a direct connection or circuit between the two phones. In the days of manual switchboards, an operator would manually connect the wire running from your home with the wire running to the person's phone you wished to call. Although this manual switching system has been automated for some time, a circuit switched network operates using the same principle, that is, a direct connection is setup and maintained whilst the conversation takes place. For much of the time this circuit is not really being used; during a typical conversation we spend less than half the time listening, less than half the time speaking and the remaining time is silence. Digital systems make much more efficient use of the lines.

Digital systems use packet switching to more efficiently utilise the line resources. Packet switching involves converting the analog signal into digital and then splitting this digital data into small chunks or packets, each packet being addressed and sent as an independent unit. As the packets are really sequences of binary digits they can be compressed to further reduce line usage. This means many many conversations can share the same line simultaneously. Each packet is routed through the system based on its address, once a packet reaches the telephone exchange closest to the recipient's home it is converted back to analog and placed on the wire leading to their phone. Packet switching is the same process as that used to transmit and receive data over the Internet; in fact most of the transmission resources used for Internet communication are owned and operated by telephone companies.

The telephone and data communication systems are becoming more and more integrated; today your phone conversation is possibly sharing a line with email messages, web pages or any other type of digital data packets. As all the data is digital the method of communication used to send these data packets can also be identical.



GROUP TASK Investigation

ISDN and DSL are both systems for digital transfer between home or office and telephone exchanges. Investigate how these systems operate using the existing telephone wires.



GROUP TASK Investigation

Phone cards are now available with incredibly low call rates; these companies use the Internet to provide the link. Investigate how phone cards work and why their call rates are so inexpensive.



Consider video on DVD and Blue-ray

The ability to store video at all has only been around since the mid 1950s. Prior to this time all television was broadcast live. During the 1970s video cassette recorders (VCRs) became commercially available, this led to the creation of a whole new industry. We now have video stores all over the place; at the time of writing these stores have largely converted from the old analog VHS format to DVD. Most are now stocking Blue-ray versions of movies. So what was the problem with VCRs and why all the fuss about DVD and Blue-ray?



*Fig 2.21
VCRs contain many complex
mechanical parts.*

VCRs are incredibly complex in a mechanical sense; the drive mechanism must be able to extract the tape, wrap it around various rollers and read/write heads, and then manage to move the tape at a constant speed; furthermore the motors and gears required to just eject tapes are themselves an incredible piece of engineering. A VHS tape has quite a complex set of moving parts; two spools to hold the tape, a spring-loaded door and various locks and rollers to ensure the tape remains at the correct tension. Contrast all these components with those required for DVD or Blue-ray playback. The optical disk has no moving parts and the player has just two, one to spin the disk and another to move the laser in and out. Mechanical complexity is still a relatively minor reason why DVD and Blue-ray has largely replaced VHS.

The main reason is the picture and sound quality available on DVD and Blue-ray. On DVD pictures are stored at approximately twice the resolution of VHS and Blue-ray much higher again. The audio track on optical disks typically contains six channels of surround sound compared to two channels on VHS tapes. DVDs store binary data as bumps on a continuous spiral track just 340 nanometers wide. Each bump is permanently stamped onto the disc and is around 400 nanometres long. A single sided/single layer DVD has a track that if unravelled would be approximately 12km long. Each track is able to store some 4.38GB of data, so a double sided/double layer DVD can store around 16GB of data. On Blue-ray disks each track stores around 25GB of data and it is possible to have multiple layers; up to 20 layers on some Blue-ray implementations. It is difficult to corrupt this data as the bumps are physical marks and the whole disk is coated with plastic. On the other hand analog tape deteriorates with time, the tape stretches, dust and magnetic interference corrupts the data; none of this is an issue with optical disks. Unlike VHS, an optical disk can store more than just video data; it is able to store any type of digital data. With VHS it is necessary to rewind or fast-forward to locate a particular scene, with optical disks the laser can simply move directly to any location on the disc.



GROUP TASK Discussion

“The mechanical complexity of VCRs has been replaced by the digital complexity necessary to process vast quantities of digital video data.”
Discuss the validity of this statement.



GROUP TASK Activity

Using the information above calculate the approximate number of bits per millimetre that can be stored on a single DVD track.



Consider facsimile

Alexander Bain first patented the basic principle of the facsimile, or fax machine, in 1843; this is some 33 years before the telephone was invented. It was some twenty years later that the first operational fax machines and transmissions commenced. These early machines required the image to be printed on tinfoil using non-conductive ink. This image was mounted on a drum where an electrode would scan across the image; the circuit being completed for blank portions of the image and not completed for inked portions. Once a horizontal line had been scanned the drum would rotate slightly and the process would be repeated. At the receiving end was a drum moving at precisely the same speed as the sender's drum, an electromagnet being used to control a pen; when current flowed the pen was off the paper and when no current was present the pen would contact the paper, in this way the original image was slowly recreated. These principles are still the basis of modern facsimile.

It wasn't until the late 1960s that fax machines became commercially viable; these machines adhered to the CCITT Group 1 standard, which used analog signals and took some 6 minutes to send each page. The message was sent as a series of tones, one for white and another for black, these tones were then converted to an image using heat sensitive paper. By the late 1970s the fax machine had become a standard inclusion in most offices. A new Group 2 standard was introduced; these Group 2 machines generated digital signals and used light sensors to read images on plain paper originals. Soon after machines were developed that could print the images directly onto plain paper. The Group 3 standard was introduced in 1983; it contained various different resolutions together with methods of compressing the digital data.

Today's computers are commonly used to produce, send and receive faxes; in fact most modems have built in fax capabilities. There are even Internet sites that allow a single fax to be broadcast to many thousands of fax machines simultaneously. It is not uncommon for a business to not own a fax machine at all; rather they use a computer for all their facsimile tasks. Devices are also available that integrate scanning, faxing and printing into a single peripheral device.



Fig 2.22

The Canon D620 combines the functions of a scanner, fax and laser printer.



GROUP TASK Discussion

Most business faxes are substantially text documents, yet faxes are not transmitted as text data at all, rather they are transmitted more like images. Discuss reasons why this is the case.



GROUP TASK Discussion

When fax transmissions first became commonplace there were concerns expressed in regard to the security of the data whilst in transit. These concerns have never been properly addressed, yet facsimile is routinely used for sensitive legal and medical documents. Discuss reasons why fax is used despite its lack of security.



Consider media retrieval management

Schools, universities, hospitals, libraries, businesses; they all utilise different media types from a variety of different sources. Examples include; video from VCRs, DVDs or even direct broadcasts from video cameras, audio from CDs, images and other types of media from computer files. Media retrieval management systems integrate all of an institution's different media into a centralised system. The purpose of such systems is to provide users with efficient access to all the institution's media resources via a single integrated interface.

Media retrieval management systems come in a variety of different flavours to suit the existing media resources and infrastructure existing within the organisation. Some systems are totally digital, whilst others allow a mix of analog and digital data. Let us consider examples of possible configurations.

1. Analog/digital mix

Many schools have an existing coaxial cable network linking all the televisions on the campus to a bank of VCRs and other video sources such as laser disks and DVDs. Manual switches being used to direct a particular video source to a particular room. The operation of this network together with its various VCRs and other video sources can be automated.

The network of coaxial cables can be utilised not only to transmit analog video signals, but also to send digital control signals from each room back to the centralised system. A set top box is installed on top of each monitor in every room; these boxes receive control messages from infrared remote controls and send them down the cable to the central computer. The central computer has an interface to each data source (e.g. VCRs); some sources have serial ports and so can be controlled directly, whilst others are controlled using infrared signals from the central computer. Essentially the task of the central computer is to connect a room to a particular data source and then to pass on control commands from the users remote to this data source.

Another common configuration is used when the school has an existing computer network throughout the school. In this case communication with the central computer can be made via this network. Some systems use a computer in each room to communicate with the central computer and for others a set top unit is connected to the computer network. In either case, the actual data is received as an analog signal to each room's television.



*Fig 2.23
Rauland's Telecenter IP integrates
digital and analog technologies.*



GROUP TASK Discussion

Brainstorm a list of all the various information resources within your school. Classify each as either analog or digital resources. Develop a series of recommendations on how these resources could better be managed and accessed using a media retrieval management system.

2. Digital

Totally digital systems utilise the institution's computer network for transmission of control signals to the central computer and also for transmission of the actual data. Obviously large quantities of digital data will be transferred; hence the network normally requires extensive upgrade to cater for this need. All the data is stored on a central server using large and fast hard disk storage. The server shown in *Fig 2.24* has some 3000 gigabytes of storage and is able to transmit data at a rate of up to 2500 megabits per second. Systems, using servers such as the one pictured, are commonly used for large organisations such as universities, schools delivering distance education or for movie systems such as those found in large hotels. Digital media provides the most flexible method of delivery and also allows for comprehensive monitoring and security of data access.

Digital systems are not compatible with older analog technologies such as VCRs, any analog data must be converted to digital prior to transmission. A totally digital system means that all media content is available at all times to the entire site, there is no need to insert or eject tapes or disks during normal operation. This type of system allows various rooms, or even different campuses, to view and control the same data source simultaneously but independent of each other; for example, one class can be watching the first scene in a movie whilst another is viewing a later scene, in fact if students are at their own computers then each student can be viewing different scenes from the same movie.

Many digital systems still provide the facility to utilise existing analog display devices. The central computer converts the digital signal to analog and broadcasts this signal on a particular analog channel. Classes wishing to view the program select the required channel on their TV set. Control of the signal, such as pausing, rewinding, or fast forwarding, is performed via a nearby computer.

Various other functions can be integrated into digital media retrieval systems; for example,, announcement and intercom systems, monitoring student progress and delivery of various types of digital files. Because the system is based on a single central machine upgrading to incorporate new and evolving needs is simplified. As all types of information is now being digitised a system based on such data is more likely to meet the long term media needs of organisations.



Fig 2.24
This server has some 3000 gigabytes of storage and is able to transmit up to 2500 megabits per second.



GROUP TASK Discussion

Digital media retrieval management systems require servers with large storage, together with fast rates of access. Why is this not true for most mixed analog/digital systems? Discuss.



GROUP TASK Discussion

'Totally digital media retrieval systems are likely to totally dominate the market in the near future.' Do you agree with this statement? Discuss.



HSC style question:

Dinky Di Sheds manufactures a range of customised steel framed sheds. The sheds are sold through a network of showrooms (and salespersons).

The following sequence of steps occurs for each customer-shed purchase:

1. Customer arrives at one of the showrooms and discusses with a salesperson to decide on the details of the shed they require.
 2. Salesperson enters dimensions, colours, window and door positions and various other details into 'Dinky Di Sheds' custom software application.
 3. Software outputs basic plans of the shed together with a quotation.
 4. Customer signs quotation and pays 10% deposit using EFTPOS system.
 5. Software transmits shed details to 'Dinky Di Sheds' head office software system.
 6. Head office software creates structural engineered plans, which are approved by a structural engineer and transmitted back to the salesperson as a PDF file.
 7. Salesperson prints copies of structural engineered plans for customer to manually submit to their local council for approval.
 8. Council officers check plans comply with their various legal requirements, inform neighbours and finally approve the plans. The customer then informs the salesperson and pays the balance owing.
 9. Salesperson transmits order for the shed to 'Dinky Di Sheds' head office.
 10. Head office manufactures the shed and subsequently delivers the shed to the customer's home. Customer takes delivery and signs delivery docket.
- (a) Consider one of the showrooms as a complete system.
 - (i) Identify all the inputs into one of the showroom systems.
 - (ii) Identify all the outputs from one of the showroom systems.
 - (b) Identify and describe processing information processes occurring from when a customer arrives at a 'Dinky Di Sheds' showroom until their shed is delivered.
 - (c) Within the above 10 steps, identify where information output from one system is used as input data to another system.
 - (d) Steps 1 through to 7 are commonly completed within 1 hour. Prior to the use of digital data the equivalent steps required a minimum of 2 weeks to complete.
 Apart from time savings, discuss other advantages of the use of digital data within this system.

Suggested Solution

- (a) (i) Inputs to the showroom system include:
- Details of shed from customer (includes dimensions, colours, window and door positions, etc).
 - Customer Signature on quotation.
 - Card and PIN for EFTPOS deposit transaction.
 - Engineered plans PDF from head office.
 - Customer notification of council approval.
 - Balance owing from customer.
- (ii) Outputs from the showroom system include:
- Basic plans for customer.
 - Quotation for customer.
 - Deposit total, card details and customer PIN to EFTPOS system.
 - Shed details to Head Office system.
 - Printed Engineered Plans for customer.
 - Shed order transmitted to head office.
- (b) Processing information processes include:
- Creating basic plans and quotation based on entered shed details.
 - EFTPOS system approving deposit transaction.
 - Head office software created structural engineered plans.
 - Council approving plans using legal requirements.
 - Head office processing shed order - such as determining required parts/materials and then ordering from different suppliers.
- (c) Examples where output from one system is input to another include:
- 10% of quotation total is used as input to the EFTPOS system.
 - Shed details from showroom are input to head office system.
 - Engineered PDF plans from head office are input to showroom and then once printed are input to Council approval system.
 - Council approval is input to showroom to commence order.
 - Shed order from showroom is input to head office to commence shed production.
- (d) Advantages of digital data in this system include:
- Initial shed details can be easily altered to view the effect on total cost.
 - No qualified engineers are needed at each showroom as the data is sent to head office from multiple showrooms for approval by a single engineer.
 - Prior to production, shed details can be altered at no (or minimal cost) to customer as the quotation and plans are generated by the software rather than having to be redrawn manually by an engineer.
 - Customisation is possible with minimal direct input from human engineers.
 - Extra copies of plans are easy to obtain because they are stored in digital form as PDF files.

CHAPTER 2 REVIEW

1. Digital data has all the following advantages over other types of data EXCEPT:
 - (A) Ease of transmission.
 - (B) Used to represent all media types.
 - (C) Easily understood by humans.
 - (D) Superior processing speeds.
2. Information processes must be connected to each other if the information system is to achieve its purpose. These connections are based on:
 - (A) data being passed between different information processes.
 - (B) the actions they perform often being the same or similar.
 - (C) the logical order in which the processes are performed.
 - (D) Both (A) and (C).
3. Which of the following is true for all information processes?
 - (A) All information processes transform data into information using various actions.
 - (B) All the information processes alter the data within the system, that is, once completed the original data has been changed in some way.
 - (C) Each information process can be uniquely classified using one of the seven syllabus information processes.
 - (D) Each information process requires input, on which it performs its actions; finally a corresponding output is produced.
4. In relation to the organising process, structuring means:
 - (A) arranging the data logically to meet the needs of other information processes.
 - (B) sorting the data into alphabetical or numerical order.
 - (C) coding each individual data item into its binary digital equivalent.
 - (D) All of the above.
5. In relation to the transmitting and receiving information process, a medium is:
 - (A) the type of data being transferred.
 - (B) the technique used to encode and then decode the message.
 - (C) the physical components used to accomplish the process.
 - (D) the resource that carries the message during its transmission.
6. The information process that has the most obvious effect on the user's view of the system would be the:
 - (A) collecting process.
 - (B) processing process.
 - (C) analysing process.
 - (D) displaying process.
7. The best digital media type for storing bird sounds would be:
 - (A) video compressed using block based coding.
 - (B) audio stored using sound samples.
 - (C) audio stored using notes.
 - (D) a description of the sound using text.
8. The purpose of compressing a file is to:
 - (A) remove parts of the file that humans are unlikely to perceive.
 - (B) improve the appearance of the final information.
 - (C) increase the amount of processing needed to view the file.
 - (D) reduce the size of the file for both storage and transmission.
9. In relation to telephones and circuit and packet switching, which of the following statements is most valid?
 - (A) Circuit switching makes far better use of each line as both parties are directly connected.
 - (B) Packet switching causes breaks in the conversation whilst each person waits for packets to arrive.
 - (C) Packet switched sections of the network utilise line resources far more efficiently than circuit switched sections.
 - (D) Packet switching is used for Internet connections but is not suitable for telephone systems therefore telephones must use circuit switching.
10. The most significant reason for the conversion of most video media from VHS tapes to DVDs is:
 - (A) DVDs have far fewer moving parts compared to VHS tapes and VCRs.
 - (B) DVDs store data digitally, which is a far better and more up-to-date system.
 - (C) The image and sound tracks on DVDs are of a far higher quality.
 - (D) DVDs are smaller and more durable than VHS tapes.

11. Define each of the following terms:

(a) digital	(c) integer	(e) vector image
(b) binary	(d) sound sample	(f) media

12. List three example scenarios where each of the following media types would be used.

(a) text	(c) image	(e) video
(b) numbers	(d) audio	

13. List two examples of actions that could occur during each of the following information processes.

(a) collecting	(d) processing	(e) displaying
(b) organising	(e) storing and retrieving	
(c) analysing	(f) transmitting and receiving	

14. Making a withdrawal from an ATM involves each of the seven information processes.
 - (a) List, in sequence, each action that the ATM must be performing in order to complete the transaction.
 - (b) Classify each of the actions in part (a) as belonging to one or more of the seven information processes.
 - (c) Identify the media types used during the processing of the withdrawal.

15. Read the following passage and answer the questions that follow.

What is Digital: Flexible and Unpredictable

By Peter Dunn
FACSNET High Tech Adviser

Perhaps the most remarkable and important aspect of the digital revolution is its flexibility and unpredictability, and the way in which it allows mass-produced semiconductors and computers to be used for surprising purposes. This goes back to the early 1960s, when students at MIT were given access to one of the first Digital Equipment Corp. computers -- a system designed for scientific and engineering work, with a programming language (GLOSSARY LINK) that was, for its time, very flexible and powerful. One student wrote a program that allowed the user to enter text, and create document files that could be stored and edited. Today, we would call it a "word processor;" the student named the program "Expensive Typewriter." Another student found a way to display two small space ships on the machine's screen, and allow players to manoeuvre them and shoot at each other -- the first video game. DEC had never envisioned that its machine would be used for these sorts of things, but the potential was there and people used their imaginations (and programming skills).

Bottom line: when any area of human endeavour "goes digital," it gains access to the ever-increasing power of silicon, and to the imagination of hardware and software engineers. This is a potent combination, and its effects will continue to be felt across society for the foreseeable future.

- (a) Peter Dunn refers to the digital revolution as being "flexible and unpredictable". Discuss how he justifies this premise throughout the article.
- (b) Do you agree with Peter Dunn's "bottom line"? Use examples of systems that have "gone digital" to support your answer.

In this chapter you will learn to:

- for a given scenario, identify alternatives for data collection and choose the most appropriate one
- use a range of hardware collection devices to collect different data types
- describe the operation of a range of hardware collection devices
- make predictions about new and emerging trends in data collection based on past practices
- choose the most appropriate combination of hardware, software and/or non-computer tools to collect data from a given source
- use the Internet to locate data for a given scenario
- design forms that allow data to be accurately recorded and easily input into software applications
- select and use appropriate communication skills to conduct interviews and surveys so that data can be accurately collected
- identify existing data that can be collected for an information system for a given scenario
- recognise personal bias and explain its impact on data collection
- identify the privacy implications of particular situations and propose strategies to ensure they are respected
- predict errors that might flow from data inaccurately collected
- predict issues when collecting data that might arise when it is subsequently analysed and processed

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Collecting – the process by which data is captured or entered into a computer system, including:

- deciding what data is required
- how it is sourced
- how it is encoded for entry into the system

Hardware used for collection

- scanners and/or digital cameras to collect images
- microphones and/or recording from peripheral devices to collect audio
- video cameras and/or peripheral devices with appropriate interfaces to capture video
- keyboards and/or optical character readers to collect numbers and text
- data capture devices such as counters for counting cars on a road
- historical and emerging trends in hardware collection devices

Software used for collection

- device drivers that allow hardware to interface with the operating system
- software that allows participants to enter or import data
- software that allows participants to move data between applications

Non-computer procedures in collecting

- literature searches
- surveys and interviews
- form design for data collection
- manual recording of events
- existing non-computer data

Social and ethical issues in collecting

- bias in the choice of what and where to collect data
- accuracy of the collected data
- copyright and acknowledgment of source data when collecting
- the rights to privacy of individuals on whom data is collected
- ergonomic issues for participants entering large volumes of data into an information system

TOOLS FOR INFORMATION PROCESSES: COLLECTING

Collecting is essentially an information input process; it gathers data from the environment for use by the system. In Chapter 2 we discussed aspects of the collecting process that need to be understood prior to the actual collection of data commencing; this includes deciding on what to collect, from where it will be collected as well as how it will be encoded during collection. In this chapter we focus on the tools available for use during the collection process. Each of these tools is suited to the collection of particular media types from particular sources using particular collection techniques. For example a microphone is used to collect audio data from sound waves. It does this by sensing the compression waves and converting them into changes in voltage; this data is then suitable for conversion into digital sound samples. Conversion from analog into digital, is strictly speaking, an organising process, however as it is integral to the operation of many collection tools it makes sense to describe this operation as we discuss each tool.

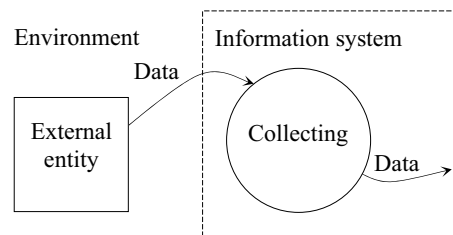


Fig 3.1

Collecting gathers data from the environment for use by the system.

Many of the tools considered in this, and subsequent chapters perform actions from more than one of the seven information processes; these tools, as is the case with the information processes themselves, do not operate in isolation. For example the CPU is clearly involved in all seven of the information processes yet we examine its operation in the processing chapter. Each tool has thus been categorised according to the particular information process that most closely aligns with the actions it performs.

In this chapter we first consider the operation of various hardware devices used to collect each of the different media types. In terms of collecting, hardware includes such common input devices as keyboards, scanners, microphones and video cameras together with various specialised data capture devices. We then examine the software used when collecting data; the software used to interface or communicate with input devices, software that forms the user interface for data entry and software that imports data so it can be moved between applications. We examine a number of non-computer procedures for collecting and finally discuss some social and ethical issues relevant to the collection process.



GROUP TASK Activity

Brainstorm a list of input devices. Categorise this list according to the different types of media each device is designed to collect

HARDWARE USED FOR COLLECTION

It is not possible to examine all types of input device, so in this section we restrict our discussion to include at least one type of device for each of the five media types.

- Keyboard for collecting of text.
- Mouse for collecting various media types.
- Scanner for collecting images.
- Digital camera for collecting images.
- Microphone and sound card for collecting audio.
- Video camera for collecting video.
- Vehicle counting and monitoring for collecting various media types.

Be aware that often the data collected by these devices is organised into a different type after collection; from the users' perspective this is often not obvious. For example when using a spreadsheet, we think we are entering numeric data, in actuality the keyboard collects the data as text; it is the spreadsheet application that converts this text data into numbers and displays it on the screen. Similarly a barcode scanner is really collecting image data, organising it into text (usually digits) and sending it to the computer for further processing.

KEYBOARD

In essence a keyboard is a collection or matrix of switches; each switch completes a circuit to indicate a particular key, or combination of keys, has been pressed. A digital code representing the key is then sent as an electrical signal to the computer. Sounds relatively simple, however in reality the keyboard is an amazing mix of ergonomic and technological design.

To structure our discussion let us work through the operations occurring as a single character is collected. That is, from the time the user presses a key until the computer receives the information.

First the user decides which key to press and locates that key. This may seem obvious but there are many aspects of keyboard design that facilitate this process. Consider the standard design of the keys; in most cases a QWERTY layout is used, the layout of the keys needs to be familiar if the user is to efficiently locate the correct key. Consider the physical size and shape of each key and the way each row is staggered relative to other rows (see Fig 3.2); these attributes are common to almost all keyboards, they allow users to transfer their keyboard skills from one keyboard to another. At first glance most keys appear to be cubes; actually they are tapered, with the top surface of each pad slightly concave; these design elements assist the fingers to positively locate the required keypad without touching adjoining keys.

So the user now presses the key; during this process the keyboard provides feedback to the user. Finger pressure moves the key down then upon release the key springs back to its original position, at the same time an audible 'click' is often produced. This feedback is a major factor in determining the general feel of a keyboard and is perhaps the most significant reason conventional keyboards are considered superior to most notebook keyboards; notebooks minimise the downward throw of each key to reduce their thickness.



Fig 3.2

*Section of a QWERTY keyboard.
Note the staggered rows and
standard size and shape of each key.*

Contained under each keypad is a switch which completes a circuit indicating precisely which key has been pressed. There are various designs of key switch used for this process; older designs use a matrix of mechanical switches, each switch being similar to those used for other applications such as door bells. At the time of writing, the most common designs utilise flexible rubber or silicone domes. Some use separate domes containing a carbon button for each key (see *Fig 3.3*). When a key is pressed the dome flexes, causing the carbon button to complete a circuit on the underlying circuit board, when released the dome springs back to its original shape. Other designs utilise two printed circuit boards separated by a thin film containing a hole for each key (see *Fig 3.4*); the domes are contained within a single silicone membrane. When a dome is depressed the contacts touch through the hole in the film to complete the circuit. All these designs are far simpler, and cheaper to produce, than traditional switches; furthermore the domes protect the actual contacts from dust and liquid spills.

The circuit board is really a matrix of wires; the intersection of a row and a column identifying a specific key. Each row and column is connected to the keyboard's internal controller which is a microchip contained within the keyboard case. The controller's job is to make sense of these signals and convert them into binary data for transmission to the computer. In actuality the controller detects changes in voltage; as a key is pressed the voltage in that circuit goes from low to high, and similarly when a key is released the voltage returns from high back to low. Every key is associated with a pair of scan codes; the 'make code' is generated as the key is pressed and the second, known as the 'break code', is generated when the key is released. The controller produces these scan codes, stores them in its own internal memory and sends them to the computer via an interface cable.

The interface cable contains four active wires; two are used to power the keyboard, one is a clock signal and the last is used for transmission of the scan codes and other data. Commonly the cable connects to the motherboard using a USB port; USB ports use synchronous serial communication; synchronous means the data arrives and departs at a steady rate controlled by the clock signal and serial means a single wire is used to transmit data one bit after the other.

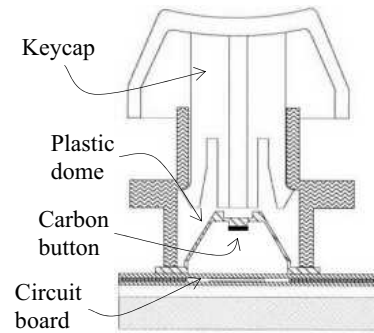


Fig 3.3

Detail of a typical flexible dome key switch similar to those used on many keyboards.

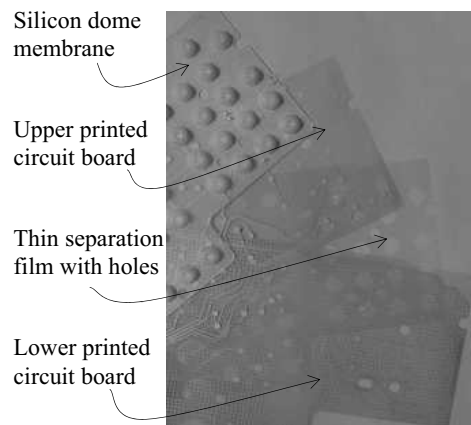


Fig 3.4

Detail of a keyboard design utilising a silicon membrane of domes and two circuit boards separated by a thin film.

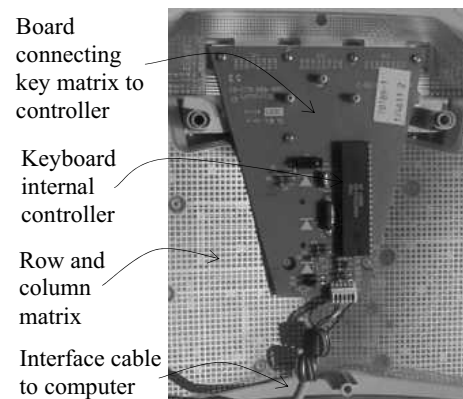


Fig 3.5

Detail of the keyboard controller within Microsoft's 'Natural' Keyboard.

When a series of scan codes arrive at the motherboard they are stored in memory and the operating system is notified using an interrupt request. The operating system, with assistance from the keyboard driver software, then examines the scan codes and responds accordingly. In most cases the scan codes are converted into a representation that includes the key's ASCII code together with information in regard to any modifier keys that may have been used. This data is passed to the currently active application. In other words the operating system transforms the scan code data into information that is meaningful to the application. This means different keyboard layouts are specified at the operating system level rather than at the keyboard itself; *Fig 3.6* shows a screen used to implement this facility within Microsoft's Windows XP; obviously the labels on each key would require alteration to reflect the changes made to such settings. The operating system also intercepts keystrokes that are intended for system level tasks, such as switching between applications, starting new applications or even rebooting the system.

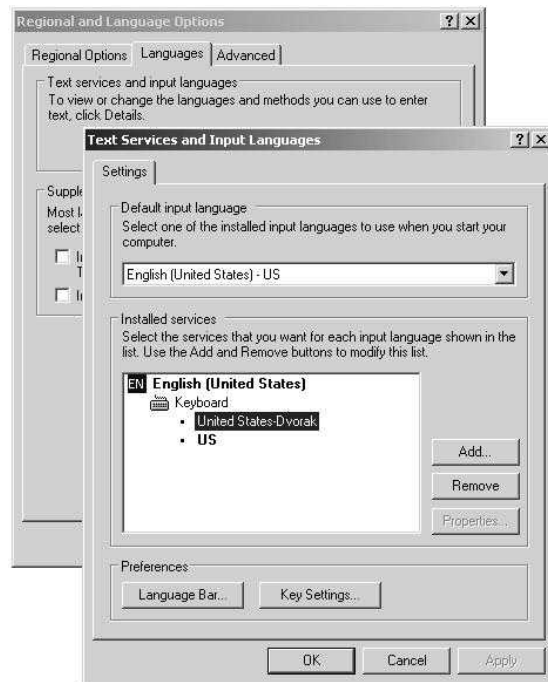


Fig 3.6
Changing the keyboard layout to Dvorak using the control panel in Microsoft's Windows XP.

So far we have only considered the transfer of data from the keyboard to the computer, however data also travels in the other direction. For example when the caps lock is pressed the operating system responds to these scan codes by sending the keyboard a message to turn on or off the caps lock light. There is also data returned to the keyboard each time an error occurs in the transmission of a scan code; each error message signals the keyboard's internal controller to resend the last scan code.



Consider the following:

All keyboards contain groups of keys that perform related actions. Consider the following groupings:

- Alphanumeric and punctuation keys (e.g. A-Z)
- Modifier keys (e.g. Shift)
- Numeric keypad (e.g. 0-9)
- Function keys (e.g. F1)
- Cursor control and navigation keys (e.g. Page Up)
- Other specialised keys (e.g. keys for Internet access)



GROUP TASK Activity

Most standard keyboards contain at least 104 keys. Examine the keyboard you use and classify each of the keys using the above bulleted list.



Consider the following:

The QWERTY layout was allegedly designed to slow typists; old typewriters used hammers that would get caught on each other so all the common letters were moved away from the centre or home row. Other sources indicate the real reason for the design of the QWERTY layout was somewhat less technical; all the letters in the word typewriter were deliberately located in the top row to allow typewriter salesmen to type the word typewriter at incredible speeds! Regardless of the reasons, the QWERTY layout is an inefficient design, but through consistent usage it has and is likely to remain the most widely used. *Fig 3.7* shows a standard Dvorak layout; notice that this layout has the most commonly used letters located in the home row.



Fig 3.7
Dvorak keyboard layout.

Some claims made in favour of the Dvorak layout compared to QWERTY include:

- Finger travel distance is up to 20 times less.
- 70% of letters occur in the home row in Dvorak, compared to 31% in QWERTY.
- The error rate for QWERTY typists is about twice that of Dvorak typists.
- The costs of retraining typists to use Dvorak layouts could be recovered in 10 days, due to increased productivity.
- Dvorak typists experience lower instances of repetitive strain injury (RSI).



GROUP TASK Discussion

Each of the above claims can only be substantiated using solid reliable evidence. Suggest techniques that could be used to collect such evidence.

MOUSE

The basic design of the mouse were first conceived by Douglas Englebart in 1964; it was some 20 years later, when Apple released the Macintosh, that the mouse became the input device of choice. Today it is hard to imagine using a computer without a mouse.

The mouse is primarily used to collect movement data in two dimensions; usually this data is used by the computer to control the position of the cursor on the monitor. In addition mice include a number of buttons and many also include a scroll wheel that doubles as an extra button.



GROUP TASK Activity

There are many other input devices that collect movement data similar to that collected by a mouse. Create a list of such devices.

So what happens when we move a mouse; that is, how does the mouse detect this movement and transform it into digital data for use by the computer? Currently there are two common designs; one based on a rolling ball and another using a purely optical design.

Let us first consider the rolling ball design (see *Fig 3.8*). A ball, inside the mouse case rolls along the desktop as the mouse is moved. The case contains two rolling shafts, one for the X direction and another for the Y direction; these shafts are in contact with the ball hence they revolve as the ball moves. A disk with many small slits around its circumference is attached to the end of each shaft; as each shaft spins then so too does the attached disk. A light emitting diode (LED) is mounted on one side of each disk and an LED sensor on the other side; as the disk revolves the slits allow pulses of light from the LED to reach the LED sensor. The LED sensor, in simple terms, opens a

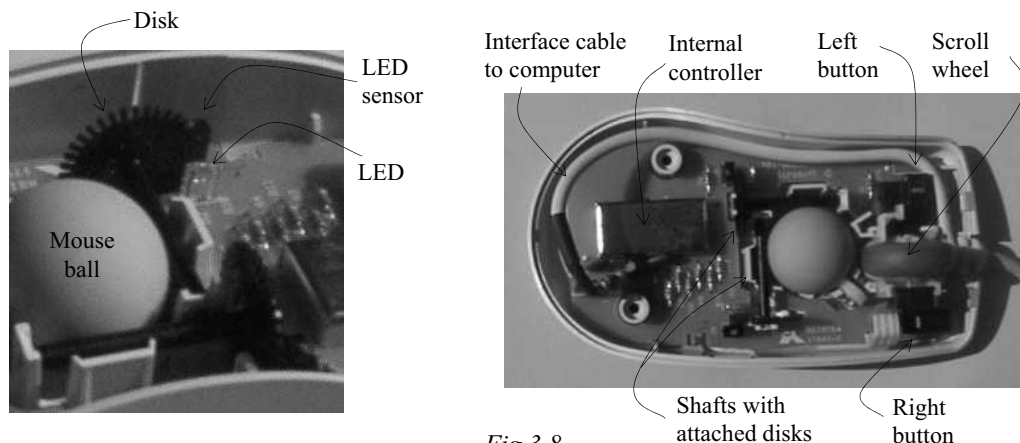


Fig 3.8

Inside a typical rolling ball mouse.

circuit each time it encounters a pulse of light. This pulsating signal is connected to the mouse's internal controller whose job is to count the number of pulses generated by each LED sensor. The controller sends this data to the computer approximately forty times every second; after each send the counters are reset back to zero. So if the mouse is moved fast to the left then maybe 200 pulses will be detected in the X direction and no pulses recorded in the Y direction; the controller sends the binary equivalent of these numbers to the computer.

You may have noticed that the above description does not explain how the mouse differentiates between left and right or forward and backward motion. In actuality the LED sensor is really two LED sensors built into a single unit (many mouse designs actually use two pairs of LED and LED sensor for each direction). This twin sensor is positioned such that when one sensor sees the light clearly the other is in a state of change. Without explaining the mathematical details, this provides sufficient information for the controller to determine the direction of movement; which is subsequently sent to the computer along with the number of light pulses detected.

The rolling ball mouse has a major drawback; the rolling action soon picks up any dust and debris present on the desktop. This rubbish rapidly accumulates on all the internal parts; in particular around each of the shafts. Once this occurs it is only a matter of time before the mouse operation is degraded and eventually stops. The obvious solution is to clean the mouse regularly, however perhaps a better solution is a mouse sealed from the outside world that has no moving parts; the purely optical mouse is one such design.

A purely optical mouse replaces all the moving and optical parts with just three components; a red LED, an image sensor and a digital signal processor (DSP). The red light from the LED is reflected off the surface of the desktop and into the lens of the image sensor (see *Fig 3.9*). The image sensor is essentially a mini digital camera; it takes a picture of the desktop some 1500 times per second. Each of these images is sent to the DSP whose primary task is to detect the direction and size of any movement by comparing features in successive images. The precision and speed of the DSP provides far more detailed information in regard to mouse movement than previous technologies; hence an optical mouse provides much smoother response and control for users.

Virtually all mouse designs include three buttons; a left and right button together with one activated by pressing down on the scroll wheel. What about the scroll wheel itself; scroll wheels do not rotate smoothly, rather they rotate in a series of clicks, each click is either in the forward direction or in the backwards direction. Consequently the data generated by the scroll wheel is represented identically to that generated by two of the other buttons; either the wheel was clicked forward or it was not, similarly it was either clicked backward or it was not. The data sent to the computer includes information in regard to the state of each of these buttons; each button is either clicked (1) or it is not (0).

Let us summarise the data collected by a typical mouse:

- Numbers representing the distance moved in both X and Y dimensions.
- The direction of the movement. Either left or right and either backwards or forward.
- The state of each button; either on or off.
- Scroll wheel events, either forward click or not, and either backward click or not.

This data, in binary form, is generated and sent approximately 40 times every second. Older mice use a PS2 port, whilst most now use a USB port for connection to the computer, hence the method of data transmission is essentially the same as that used for keyboard data.

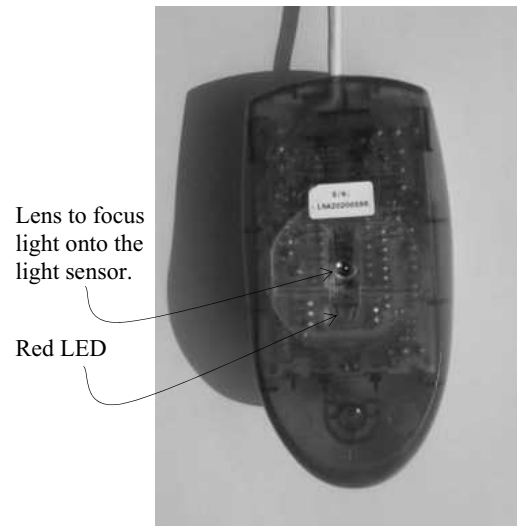


Fig 3.9
Underside of an optical mouse.



Fig 3.10
A typical mouse containing three buttons together with a scroll wheel.

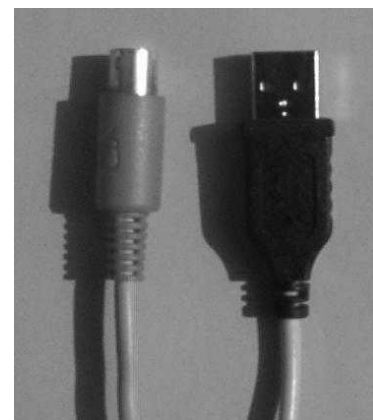


Fig 3.11
PS2 plug (left) and USB plug (right) used to connect mice.



Consider the following:

The mouse, together with most other types of pointing device, can be used to collect a variety of different media types. This is not usually their primary task, rather they are used to collect information used to initiate or facilitate the actions occurring in other information processes.



GROUP TASK Discussion

Suggest seven scenarios, one for each information process, where a mouse is used to initiate or facilitate the actions of that process.

SCANNER

There are various different types of image scanner; all collect light as their raw data and transform it into binary digital data. This digital data may then be analysed, organised and processed into numbers or text, or it may remain as image data in the form of bitmaps. Perhaps the most familiar forms of scanner are barcode scanners, used in most retail stores and flatbed scanners used to collect images in bitmap form. Let us consider the operation of common examples of each.

Barcode scanners operate by reflecting light off the barcode image; light reflects well off white and not very well off black. This is the basic principle underlying the operation of all types of scanner. A sensor is used to detect the amount of reflected light; so to read a barcode we can either progressively move the light beam from left to right across the barcode or use a strip of light in conjunction with a row of light sensors. Each of these techniques are used for different designs of barcode scanner; those based on LED, laser and CCD technologies dominate the market, *Fig 3.12* shows an example of each. Most barcode scanners incorporate a decoder to organise the data into a character representation that mimics that produced by the keyboard. This means most barcode readers can be installed between the keyboard and the computer without the need for dedicated interface software.

Barcode wands use a single light emitting diode (LED) to illuminate a small spot on the barcode. The reflected light from the LED is measured using a single photocell. As the wand is steadily moved across the barcode, areas of high and low reflection change the state of the photocell. The photocell absorbs photons (a component of light). As the intensity of photons absorbed increases so to does the current flowing through the photocell; large currents indicating white and smaller currents indicating black. This electrical current is transformed by an analog to digital converter (ADC) to produce a series of digital ones and zeros. The same LED technology is used for slot readers, where the barcode on a card is read by swiping the card through the reader.



Fig 3.12

Clockwise from top-left: LED wand, multi-directional laser and CCD based barcode scanners.

**GROUP TASK Activity**

A barcode is scanned using an LED barcode scanner and the following stream of bits is produced: 000000110011000000111111001100111111
Draw the most likely original barcode.

Lasers are high intensity beams of light and as such they can be directed very precisely. Laser barcode readers can therefore operate at greater distances from the barcode than other technologies, commonly up to about 30cm away. The reflected light from the laser is detected by the photocell using the same technique as LED scanners. There is no need to manually sweep across the barcode as the laser beam is moved using an electronically controlled mirror. Basic models continually sweep back and forth across a single path, whilst more advanced models perform multiple rotating sweeps that trace out a 'star like' pattern. These advanced models are much more effective as the user need not hold the scanner parallel to the barcode; rather the scanner rotates the scan line until a positive read is collected. Supermarkets often use this type of barcode scanner mounted within the counter top.

Charge coupled devices (CCDs) contain one or more rows of photocells built into a single microchip. CCD technology is used by many image collection devices including CCD barcode scanners, digital still and video cameras, handheld image scanners, and also flatbed scanners. For both barcode and image scanners a single row CCD is used. The light source for these scanners is typically a single row of LEDs with the light being reflected off the image back to a mirror. The mirror reflects the light onto a lens that focuses the image at the CCD. Each photocell in the CCD transforms the light into different levels of electrical current. These levels are converted into bits using a similar technique to that used in LED and laser barcode scanners. CCDs in flatbed scanners differ slightly; they convert the electrical current from each photocell into a binary number, normally between 0 and 255, using a more complex analog to digital converter (ADC).

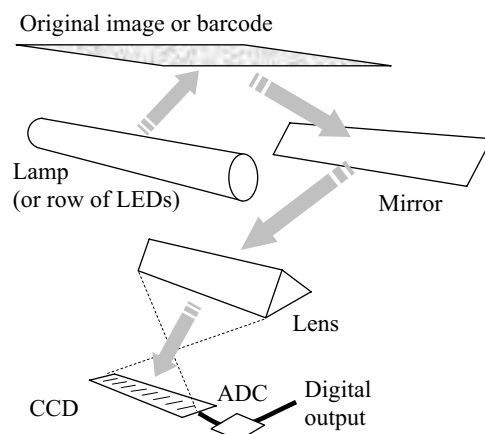


*Fig 3.13
The components and light path typical
of most CCD scanner designs.*

**GROUP TASK Investigation**

Barcode scanners are used in most retail stores and libraries. Over the next 24 hours observe closely each barcode scanner you encounter. Classify each as using either LED, laser or CCD technology. Justify your choices.

Let us now consider flatbed scanners based on CCDs in more detail. This type of flatbed scanner is by the far the most common; scanners based on other technologies are available, but currently they fall into the higher quality and price ranges. We mentioned above that the binary numbers returned from a flatbed scanner's ADC range from 0 to 255; this is the range of different numbers that can be represented using 8 bits (1 byte). If white light is used then these numbers will represent shades of grey, ranging from black (0) to white (255). So how do flatbed scanners collect colour images? Quite simply, they reflect red light off the original image to collect the red

component, green to collect the green component and blue for the blue component. Some early scanners performed this action by doing three passes over the entire image using a different coloured filter for each pass; this technique is seldom used today. Today most scanners use an LED light source that cycles through each of the colours red, green, blue; hence only a single pass is needed.

The LED lamp, mirror, lens and CCD are all mounted on a single carriage; these components are collectively known as the scan head. All the components on the scan head are the same width as the glass window onto which the original image is placed. This means a complete row of the image is scanned all at once. The number of pixels in each row of the final image is determined by the number of photosensors contained within the CCD; typical CCDs contain some 600 sensors per inch, predictably this results in images with horizontal resolutions of up to 600 dpi (dots per inch).

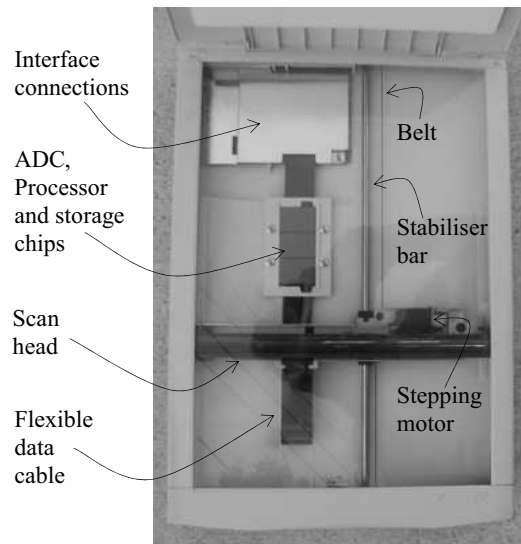


*Fig 3.14
Components of a flatbed scanner.*

So what operations occur to collect a colour image?

- The current row of the image is scanned by flashing red, then green, then blue light at the image. If you open the lid of a scanner you'll predominantly see white light, this is due to the colours alternating so rapidly that your eye merges the three colours into white. After each coloured flash the contents of the CCD is passed to the ADC and onto the scanner's main processor and storage chips.
- The scan head is attached to a stabilising bar, and is moved using a stepping motor attached to a belt and pulley system. The stepping motor rotates a precise amount each time power is applied; consequently the scan head moves step by step over the image; pausing after each step to scan a fresh row of the image. The number of times the stepping motor moves determines the vertical resolution of the final image.
- As scanning progresses the image is sent to the computer via an interface cable. The large volume of image data means faster interfaces are preferred; commonly SCSI, USB or even firewire interfaces are used to connect scanners. Once the scan is complete the scan head returns back to its starting position in preparation for the next scan.



GROUP TASK Discussion

Some scanners use 36 or even 42 bits internally to represent each pixel, yet they only output 24 bit per pixel images. Why would this be? Discuss



GROUP TASK Discussion

The packaging of a scanner implies it is able to scan at 2400dpi, you know the CCD contains just 600 sensors per inch. What is going on? Discuss

SET 3A

1. Collecting involves:
 - (A) deciding what to collect.
 - (B) locating data for collection.
 - (C) encoding the data during entry.
 - (D) All of the above.
2. Which of the following contains only input devices?
 - (A) keyboard, mouse, scanner, laser printer, microphone.
 - (B) keyboard, mouse, scanner, digital camera, microphone.
 - (C) operating system, applications, utilities, device drivers.
 - (D) keyboard, mouse, scanner, laser printer, monitor.
3. For most keyboards a single key stroke is sent to the computer as:
 - (A) an ASCII code.
 - (B) a single scan code.
 - (C) a pair of scan codes.
 - (D) text.
4. Most keyboards and mice send data to the computer via a:
 - (A) serial port.
 - (B) PS2 port.
 - (C) parallel port.
 - (D) USB port.
5. The QWERTY layout dominates because:
 - (A) it allows the word typewriter to be entered quickly.
 - (B) it slows down typists.
 - (C) the most commonly used letters are contained on the home row.
 - (D) it has been used consistently over many years.
6. The component of a flatbed scanner that progressively moves the scan head is a:
 - (A) stabiliser bar.
 - (B) belt.
 - (C) stepping motor.
 - (D) CCD.
7. A mouse that contains a ball collects movement data by:
 - (A) measuring the distance travelled in both X and Y directions.
 - (B) counting the number of light pulses reaching an LED sensor through a rotating slotted disk for both the X and Y directions.
 - (C) reflecting light off the desktop into an image sensor. The images are processed by a DSP to determine and create movement data.
 - (D) working out the distance and angle of the movement. This is accomplished using various LEDs and light sensors.
8. CCDs are used to collect:
 - (A) analog image data as different levels of electrical current.
 - (B) digital image data as a sequence of bits.
 - (C) digital image data as different levels of electrical current.
 - (D) analog image data as a sequence of bits.
9. One advantage of laser barcode readers over LED based barcode readers is:
 - (A) the whole barcode is read at the same time.
 - (B) they are generally less expensive.
 - (C) they operate at greater distances from the barcode.
 - (D) they can be installed between the keyboard and the computer.
10. Most flatbed scanners collect colour in images by:
 - (A) using three different coloured filters over three rows of CCDs.
 - (B) reflecting red, then green, then blue light off the image.
 - (C) averaging the colour value of adjacent pixels.
 - (D) Flatbed scanners cannot collect colour images.
11. Describe, in point form, the processes occurring when a single character is entered via the keyboard into a software application.
12. Describe the nature and operation of charged couple devices.
13. Explain the advantages and disadvantages of the three different types of barcode scanners discussed in the text.
14. Examine the mouse attached to your home or school computer. Describe its operation.
15. Two keyboard layouts are described in the text; QWERTY and Dvorak. There are many other layouts used for specific dedicated tasks or that are used to collect data in foreign languages. For example an ATM includes a dedicated keyboard.
Search the Internet for images of at least three different layouts and describe their purpose.

DIGITAL CAMERA

Digital cameras have completely transformed the photographic process. The traditional mechanical and chemical processes using film had been in use since the 1830s. Electronic and digital processes have largely replaced this traditional system.

Currently all digital cameras are based on either charge coupled devices (CCD) or complementary metal oxide semiconductors (CMOS). These technologies are at the heart of all digital camera designs; both are image sensing technologies, that is, they detect light and transform it into electrical currents. Currently CCDs provide better image quality, however they cost more to produce and require significantly more power to operate. CMOSs use similar production methods to other types of microchips, hence they are inexpensive to produce and have far lower power requirements. Unfortunately the quality of images produced with CMOS based cameras is currently inferior to CCD produced images. CCD technology is used in almost all dedicated digital cameras where the need for high quality output more than justifies the extra cost and power requirements. CMOS technology is currently used for applications such as security cameras and phone cameras; image quality being sacrificed to minimise critical cost and power requirements.



Fig 3.15
The Nokia 3650 contains a CMOS based digital camera.

We discussed CCD technology previously in relation to scanners; the CCDs used in digital cameras operate in precisely the same manner, they convert photons into electrical charge. At our level of discussion this is also the primary function of CMOS chips, the only significant difference being that CMOS chips combine the image sensing and ADC functions into a single integrated chip. Our remaining discussion will focus on CCD based cameras, however much of the discussion is equally true of CMOS based cameras.

Unlike scanners, who generate their own constant light source, cameras must control the amount of light used to generate the image. In a traditional film camera this is accomplished using a shutter. The shutter alters the size of the hole or aperture through which the light passes and also alters the time the aperture is open (shutter speed). Digital cameras use the same principles; many models do have mechanical shutters whilst others do away with mechanical shutters altogether. Adjusting the time taken between the CCD being reset and the data being collected can produce the equivalent process in a digital camera.



Fig 3.16
A CCD from a digital camera.

Digital cameras must be able to collect an entire image in a virtual instant. This means a two dimensional grid of photosensors is needed; the CCD shown in *Fig 3.16* contains some 2 million photosensors, or photosites, resulting in images with resolutions up to 1600 by 1200 pixels. Digital cameras are often classified according to the number of photosites on their CCDs, cameras based on the CCD in *Fig 3.16* would be classified as 2 megapixel cameras; some CCDs contain as many as 20 million photosites.



GROUP TASK Investigation

Compare the megapixel value with the resolution of the final images for a number of digital cameras. Discuss any discrepancies found.

Remember our flatbed scanner, it collected colour using red, green and blue light; the same principle is used by digital cameras. There are various ways of implementing this principle:

- Take the picture three times in quick succession, first with a red filter then a green and finally a blue filter. The three images can then be combined to produce the final full colour image. This approach is seldom used as even slight movement leads to blurred images.
- Use three CCDs where each is covered by a different coloured filter. A prism is used to reflect the light entering the camera and direct it to all three CCDs. This approach is obviously more expensive as three CCDs and various other extra components are needed, however the resulting images are of excellent quality. This technique is generally restricted to high quality professional cameras.
- By far the most common approach used is to cover each photosite with a permanently coloured filter. The most common filter pattern used is called a bayer filter; this pattern alternates a row of red and green filters with a row of blue and green filters.

The Bayer filter is the most common approach (see Fig 3.17), let us continue our discussion based on this technique. A Bayer filter has two green photosites for each red and each blue photosite. The human eye is far more sensitive to green light, hence using extra green sensors results in more true to life images. So the raw analog data from the CCD represents the intensity of either red, green or blue light in each of its photosites. This analog data is then digitised using an analog to digital converter (ADC). Earlier we discussed how 2 megapixel cameras produce final images with resolutions containing approximately the same number of full colour

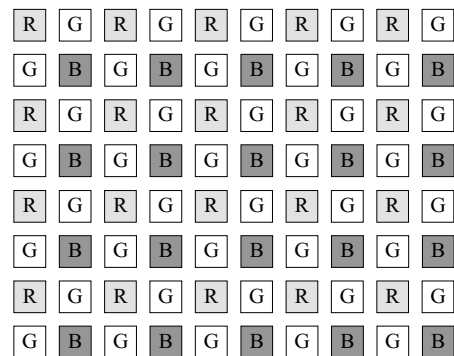


Fig 3.17

Bayer filters alternate red and green rows with blue and green rows.

pixels ($1600 \times 1200 = 1,920,000 \approx 2$ million pixels); how is this possible when the initial digital data from the ADC contains information representing the intensity of one single colour per pixel? A process known as demosaicing is used to produce the final colour values for each pixel. Examining the Bayer filter in Fig 3.17, we see that each red photosite is surrounded by four green and four blue photosites, averaging the four green values gives us a very accurate approximation of the likely actual green value, similarly averaging the blue values gives us the most likely blue value. Combining the original 8 bit red value with the calculated 8 bit green and blue values give us the final 24-bit colour value for the pixel. This processing occurs for every pixel, resulting in the output of a complete 24 bits per pixel image with a resolution similar to the number of photosites on the CCD.

The resulting image is usually compressed, to reduce its size, prior to storage; commonly a lossy technique, such as JPEG, is used. The file is then stored on a removable storage device, most cameras use removable flash memory cards. A computer later reads these cards, either directly or via an interface cable, which stores the images on the computer's hard disk.



GROUP TASK Discussion

“Digital cameras can really only see in shades of grey.” Discuss the validity of this statement based on the operation of bayer filters.

MICROPHONE AND SOUND CARD

Microphones are, predictably, used to collect data in the form of sound waves. They convert these compression waves into electrical energy. In digital systems, this analog electrical energy is converted, using an analog to digital converter (ADC) into a series of digital sound samples. In this section we examine the operation of microphones and then consider the operations performed by a typical sound card to process the resulting analog electrical energy into a sequence of digital sound samples.

There are a variety of different microphone designs, the most popular being dynamic microphones and condenser microphones. All these designs contain a diaphragm which vibrates in response to incoming sound waves. If you hold your hand close to your mouth whilst talking you can feel the effect of the sound waves; the skin on your hand vibrates in response to the sound waves in exactly the same way as the diaphragm in a microphone vibrates.

A dynamic microphone has its diaphragm attached to a coil of wire; as the diaphragm vibrates so too does the coil of wire (see Fig 3.19). The coil of wire surrounds, or is surrounded by, a stationary magnet. As the coil moves in and out the interaction of the coil with the magnetic field causes current to flow through the coil of wire. This electrical current varies according to the movement of the wire coil, hence it represents the changes in the original sound wave.

Condenser microphones alter the distance between two plates (see Fig 3.20). The diaphragm is the front plate; it vibrates in response to the incoming soundwaves, whereas the backplate remains stationary. Therefore the distance between the diaphragm and the stationary backplate varies; when the two plates are close together electrical current flows more freely and as they move further apart the current decreases, hence the level of current flowing represents the changes in the original sound waves. Condenser microphones require a source of power to operate; this can be provided from an external source via the microphone's lead or by using a permanent magnetically charged diaphragm.

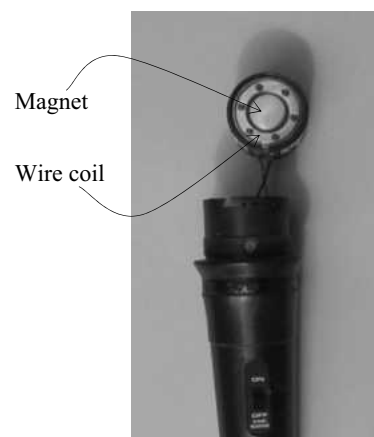


Fig 3.18

*A dynamic microphone element.
This one has the magnet
mounted within the wire coil.*

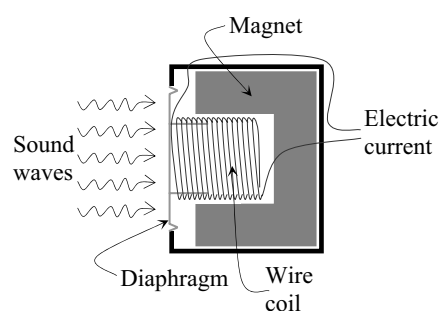


Fig 3.19

Detail of a dynamic microphone.

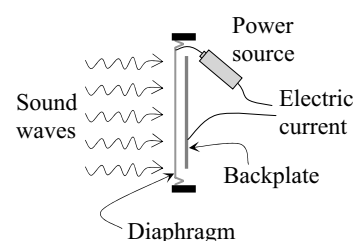


Fig 3.20

Detail of a condenser microphone.



GROUP TASK Investigation

Make a list of all the microphones you see each day. Can you determine whether these microphones are dynamic, condenser or some other design?



Consider the following:

The varying electrical current produced by a microphone is essentially the same as the raw analog signal output from all types of audio devices. Therefore it is possible to connect any of these different audio sources to one of the analog input ports on a computer's sound card; just be careful to connect to a port designed for the level of signal produced by the device. There are usually a number of input ports on most sound cards suited to different levels of analog input signal.

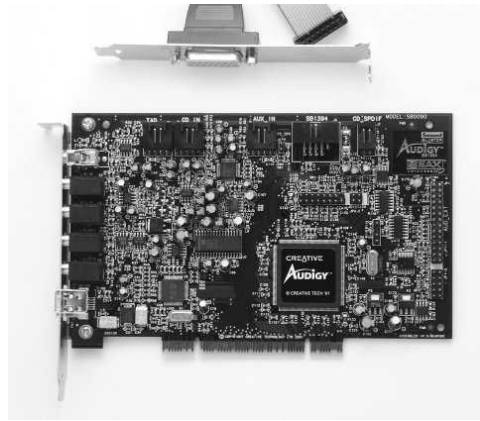


Fig 3.21
Creative's Audigy sound card.



GROUP TASK Investigation

Examine the ports, and accompanying documentation, for either your school or home computer's sound card. Describe the difference between each of the input ports and list suitable audio sources that could be connected to each port.

Let us now consider the processes taking place once the analog signal from the microphone reaches the computer's sound card. The signal is fed through an analog to digital converter (ADC), which predictably converts the signal to a sequence of binary ones and zeros. The output from the ADC is then fed into the digital signal processor (DSP), whose task is to clean up any abnormalities in the samples. The final sound samples are then placed on the computer's data bus. The data bus feeds the samples to the main CPU, where they are generally sent to a storage device.

The major components involved in processing the audio data are the analog to digital converter (ADC) and the digital signal processor (DSP). Let us consider each of these components in more detail.

Analog to digital converters (ADCs) repeatedly sample the magnitude of the incoming electrical current and convert these samples to binary digital numbers; for audio data the size of the incoming current directly mirrors the shape of the original sound wave, hence the digital samples also represent the original wave. The ADCs used in many other devices, including scanners and digital cameras, are essentially the same as those found on sound cards; the CCDs in image collection devices produce varying levels of electrical current that represent the intensity of light detected at each photosite. The electrical signal is much the same as that produced by audio collection devices.



GROUP TASK Activity

Brainstorm a list of hardware devices that would likely include an analog to digital converter. Indicate the media type collected by each device and how different levels of electrical current could be used to represent the identified media type.

Most analog to digital converters contain a digital to analog converter (DAC). On the surface this seems somewhat strange, however the digital to analog conversion process is significantly simpler than the corresponding analog to digital conversion process.

The components and data connections within a typical ADC are shown in *Fig 3.22*; this ADC performs its conversion using the following steps:

- At precise intervals the incoming analog signal is fed into a capacitor; a capacitor is a device that is able to hold a particular electrical current for a set period of time, this allows the ADC to examine the same current repeatedly over time.
- An integrated circuit, called a successive approximation register (SAR), repeatedly produces digital numbers in descending order. For 8-bit samples it would start at 255 (11111111 in binary) and progressively count down to 0.
- The DAC receives the digital numbers from the SAR and repeatedly produces the corresponding analog signal. The analog signals will therefore be produced with decreasing levels of electrical current.
- The electrical current output from the DAC is compared to the electrical current held in the capacitor using a device called a comparator. The comparator signals the SAR as soon as it detects that the current from the DAC is less than the current in the capacitor.
- The SAR responds to the signal from the comparator by storing its current binary number. This number is one of the digital sound samples and hence is output to the DSP. The SAR then resets its counter and the whole process is repeated.

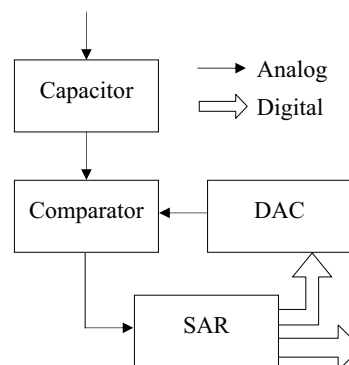


Fig 3.22
Components and data
connections for a typical ADC.

So what happens to these sound samples once they reach the DSP? The DSP's task, in regard to collected audio data, is to filter and compress the sound samples in an attempt to better represent the original sound waves in a more efficient form. The DSP is itself a powerful processing chip; most have numerous settings that can be altered using software. Most DSPs perform wave shaping, a process that smooths the transitions between sound samples. Music has different characteristics to speech, so the DSP is able to filter music samples to improve the musical qualities of the recording whilst removing noise. The DSP uses the sound samples surrounding a particular sample to estimate its likely value, if these estimates do not agree then the sample can be adjusted accordingly. Once the sound samples have been filtered the DSP compresses the samples to reduce their size. Some less expensive sound cards do not contain a dedicated DSP, these cards use the computer's main processor to perform the functions of the DSP.



GROUP TASK Discussion

The processing performed by a sound card when collecting audio data involves most of the seven syllabus information processes. Describe the operation of a sound card in terms of these information processes.

VIDEO CAMERA

Most video cameras combine image collection with audio collection; the result being a sequence of images that includes a sound track. The term ‘video camera’ is commonly used to describe devices that combine a video camera and microphone for collecting, with a video/audio recorder/player for storage and retrieval; perhaps the alternate ‘camcorder’ term better describes such devices. Analog video cameras, or camcorders, have been available for more than twenty years, however digital versions now dominate the market. There are also PC cameras or web cameras that really are just cameras, their sole task being to collect image data and send it to the computer via an interface port.

Most quality analog and digital camcorders use CCDs to capture light and microphones to capture sound. CCDs and microphones both collect analog data; they convert light and sound waves into electrical current. Digital video cameras convert these electrical signals into digital within the camera, whereas the output from an analog video camera must be converted to digital before a computer can process it. *Fig 3.23* shows a video capture device that converts the analog video and audio data from an analog source, such as an analog camcorder, into digital and sends the result to the computer via a USB port.



Fig 3.23
A digital camcorder, web camera
and a video capture device.

PC or web cameras, in most cases, use complementary metal oxide semiconductor (CMOS) chips. CMOS chips are inexpensive to produce and various functions can be combined within a single chip. The single CMOS chip in a web camera contains photosensors, and all the circuitry necessary to communicate and transmit images to the computer's port. As these cameras are designed to collect images and video for display over the Internet, the lower image quality derived from most CMOS photosensors is not significant.

Let us consider the operation of a typical camcorder in more detail. To collect video effectively it is crucial to control the changing nature of the light entering the lens. As the camera and/or subject moves the camcorder needs to respond by altering the amount of light entering the lens and also by refocussing this light onto the CCD. The CCD provides a perfect indicator of the amount of light entering the lens; if most of the photosites on the CCD record strong light intensities then too much light is entering the lens, so the diameter of the aperture is reduced; conversely if the light intensities are weak then the aperture is opened. Focussing is not so simple, the camcorder needs to know the distance to the subject of the current frame. Some camcorders bounce an infrared beam off an object in the centre of the frame; the time taken for this beam to reflect back to the camera is used to calculate the distance to the object. The camcorder uses a small motor to move the lens in or out to focus the image onto the CCD based on the calculated distance. Other camcorders compare the intensity of light detected at adjacent photosites within a rectangle of pixels in the middle of the frame; gradual changes indicating blurred images and larger differences indicating the image is focussed. The lens is then moved slightly in or out and the intensities are again compared; the process repeats until the maximum difference in intensities is achieved.

**GROUP TASK Discussion**

Virtually all camcorders provide autofocus and automatic aperture control. However, most PC or web video cameras do not contain any sort of focus or aperture control, and neither do cheaper digital still cameras. Suggest and discuss reasons why this is the case.

Each photosite in a camcorder CDD and in a digital still camera CCD collect light in precisely the same way; however a video camcorder must be able to collect some 25 to 30 images or frames every second. To accomplish this task the CCD in most camcorders has two layers of sensors, one behind the other; the front layer collects the light and then transfers the electrical current to the lower layer. Whilst the lower layer is being read, the upper layer is collecting the next image.

In all analog camcorders, and in many older digital camcorders, the lower layer of the CCD is split into two distinct fields; the first field being the odd numbered rows and the second being the even numbered rows. The data from one of these fields is read for each frame, the fields being alternated for each successive read; in effect only half the total image is retained. The images are collected in this way to reduce the amount of data and also to mirror the operation of older televisions. These older televisions display video by alternately painting the odd rows and then the even rows; this process is known as interlacing. Most digital camcorders now use 'progressive scan CCDs', this somewhat obscure term means the contents of the entire CCD is read as a single complete image. Newer televisions (and monitors) also use progressive scan to paint an entire image with each screen refresh. Camcorders using progressive scan CCDs require faster processors to manipulate the extra data, however they produce higher quality video. In addition, they are also able to collect high quality still images.

In analog camcorders each frame, or half frame, is sent to the VCR component of the camcorder where it is magnetically stored on tape. These analog tapes can then be used directly within conventional VCRs or the analog signal can be sent from the camcorder via cables to a television or other device capable of receiving analog video. For digital camcorders the data passes through an ADC; the resulting digital data is then compressed into a format suitable for storage. Currently most digital camcorders use an internal hard disk drive, magnetic tape or recordable DVD media for storage; in Chapter 6 we examine how digital data is stored these mediums. Models using a hard disk or tape require connection to the computer via an interface cable; most connect using USB or firewire ports. Models using DVD storage also include ports to connect to computers. However DVDs are convenient as their contents can be played directly using DVD players or the data can be accessed via the DVD drive on a computer. Most digital camcorders also include analog outputs and inputs; this allows transfer of video data to and from analog sources.



Fig 3.24

The Hitachi DV-MV100 camcorder stores video on recordable DVDs.

**GROUP TASK Investigation**

Most camcorders store video data in a format known as Y/C or S-Video. Investigate how this format is used to represent video data.

VEHICLE COUNTING AND MONITORING

Various government departments and commercial organisations require information in regard to motor vehicle activity; this includes statistics on traffic movements as well as information on particular vehicles. This information is used to plan future road works, monitor car parks, adjust the timing of traffic lights or even to identify and determine the speed and behaviour of individual vehicles. The data requirements vary according to the particular needs of the information system, as do the technologies used by these systems. Let us examine examples of systems that collect data using air pressure, magnetic fields and video.

Systems based on air pressure sensors are used to count the number of vehicles passing a given point. These systems include three major components; hollow rubber tubing, air pressure switches and a control unit, each of these components are visible in *Fig 3.25*. The hollow tube is placed across the roadway and is connected to an air pressure switch on the control unit; the other end of the tube is sealed. As a vehicle's tyre runs over the tube, the air pressure within the tube increases; this increase in air pressure activates the air pressure switch. The controller detects the change of state of the air pressure switch and increments a counter.



Fig 3.25
Air pressure based vehicle counter,
with inset air pressure switch.

Obviously a car will activate the pressure switch twice, and large trucks more than twice. Most controllers merely divide the number of activations by two, this is sufficiently accurate for the majority of applications. Many controllers contain multiple air pressure switches; this allows more than one traffic lane to be monitored using a single controller. Commonly the controller contains a timer, so each activation actually results in a time being recorded by the controller. The data held in the controller is transferred to a computer for further processing.

Other systems use induction coils to detect changes in magnetic fields; commonly these systems are used to control traffic lights at intersections. A loop of wire (the induction coil) is installed permanently within the surface of the road (see *Fig 3.26*). As a vehicle enters the loop, the vehicle's magnetic field alters the current flowing through the wire loop; this change is detected by the controller. These systems, rather than counting vehicles, are generally used to determine if a vehicle has arrived at a red traffic light. In practice this means the lights can remain green in the major traffic directions and only change when a vehicle is detected on the minor road.

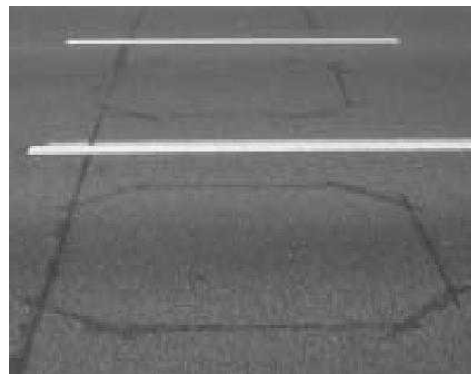


Fig 3.26
The dark lines on the above roadway
indicate an induction loop is installed.



GROUP TASK Discussion

Both of the collection devices above have limitations. Discuss scenarios where each device may fail to correctly identify a vehicle.

Systems utilising video are able to do far more than just count the number of vehicles; they can also identify individual vehicles, monitor driving behaviour and calculate speed. Examples of such systems currently in operation in Australia include speed cameras, bus lane cameras and heavy vehicle monitoring cameras.

Let us consider the Safe-T-Cam system, developed by the CSIRO, and used in NSW to monitor heavy vehicle movements. The aim of this system is to record the registration number and time as each heavy vehicle passes each Safe-T-Cam. As there is a network of more than twenty Safe-T-Cams around the state this data can be used to track the movement of individual vehicles.

Each Safe-T-Cam uses a video camera and a still camera as collection devices. Frames from the video being used to track individual vehicles and detect when they are in the correct position for the still camera to take a photograph.

The system identifies an individual vehicle by comparing the differences between the current video frame and an image of the stationary background; these differences are used to produce a difference image (see Fig 3.27). An individual vehicle can be tracked by examining the difference images produced from a sequence of frames. The example frame shown in Fig 3.27 contains just one vehicle, obviously it is more likely that multiple vehicles will be contained within each frame, the system must be smart enough to identify the same vehicle as it appears in subsequent frames. This is accomplished by comparing various attributes of the vehicles on each difference frame; if the attributes are similar then presumably it is the same vehicle.

When each identified vehicle reaches a set point on one of the video frames the digital still camera takes a photograph. These photographs are of sufficiently high resolution to allow the registration number on the vehicle to be read using optical character recognition (OCR) software. Finally the registration number and time of detection is transmitted to a central monitoring site.

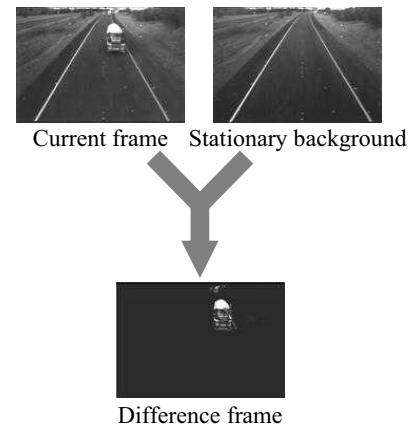


Fig 3.27

Safe-T-Cam compares the current frame with a background image to produce a difference image.



Fig 3.28

Sample photo, typical of those collected by Safe-T-Cam. The registration number can be clearly determined.



GROUP TASK Discussion

Traffic authorities aim to ensure heavy vehicles are registered and that drivers only drive for a specified number of hours each day. How can the data collected by Safe-T-Cam assist in achieving these aims? Discuss



GROUP TASK Research

There are various systems being developed that not only monitor vehicle activity but are actually able to control vehicles. Use the Internet to research these new technologies.



HSC style question:

- (a) Outline the technology and processes occurring as a digital camera collects and digitises image data.
- (b) Outline the technology and processes occurring as a flatbed scanner collects an image.
- (c) Outline the technology and processes occurring as sound is recorded using a microphone, sound card and computer to produce a sampled digital file.

Suggested Solution

- (a) At the back of a digital camera is a CCD (or a CMOS chip) which contain a photocell for each pixel. The photocells respond to the amount of light falling on them through the lens of the camera. As the photocells in a CCD are hit by incoming light, they emit a current relative to the brightness of the light. Different colours are detected through the use of a red, green and blue bayer filter covering the photocells. The electrical current is converted using an analog to digital convertor into an equivalent bit pattern for each pixel. The digital data is then processed to generate a red, green and blue colour value based on the values of the adjoining pixels. Finally the digital image data is compressed and stored on the camera's flash card.
- (b) A light source and row of sensors are mounted on a head assembly that moves in precise steps over the entire image being scanned. The more light that is reflected, then the lighter that part of the image must be. A series of focusing mirrors and lenses focus the reflected light onto the row of light sensors, which is typically a CCD. As the reflected light hits the sensors a current is produced – the more light, the stronger the current. An ADC (analog to digital converter) then converts this electrical signal into an equivalent series of bits for that section of the image. For colour scanning, there are 3 coloured lights, one for each of red, green and blue light. Each coloured light flashes in sequence over each line of the image, hence the CCD collects the intensity of red, green and blue light. The equivalent bit patterns representing the colour of each pixel of the image can then be determined.
- (c) Sound waves are detected by the microphone as a sequence of compressions and decompressions in the air. These movements in the air move a diaphragm inside the microphone which in turn generates an equivalent current. This analogue current is sent along the connecting cable through to the sound card. At the sound card, the current is sampled many thousands of times a second and put through a ADC. Each sample represents the height of the original sound and is converted by the ADC into a binary number. This effectively converts the sound wave into a long sequence of binary numbers. Most sound card also contains a digital signal processor (DSP) which smooths the transitions and filters out noise. The resulting sound samples are then transferred via the system bus to the CPU where they are directed to a storage device.

SET 3B

1. The main difference between CCDs used in digital cameras and those used in flatbed scanners is:
 - (A) digital camera CCDs have a single row of photosites, flatbed scanners have a two dimensional grid of photosites.
 - (B) flatbed scanner CCDs have a single row of photosites, digital cameras have a two dimensional grid of photosites.
 - (C) flatbed scanners generate their own light, digital cameras use natural light.
 - (D) digital cameras generate their own light, flatbed scanners use natural light.
2. A bayer filter:
 - (A) is used to remove unwanted detail during image collection.
 - (B) converts analog data into digital data.
 - (C) alternates red and green rows with blue and green rows.
 - (D) is used to compress image data.
3. Microphones are used to:
 - (A) convert sound waves into bits.
 - (B) convert sound waves into electrical energy.
 - (C) convert electrical energy into bits.
 - (D) All of the above.
4. The component on a sound card that filters and compresses the audio data is known as a:
 - (A) ADC
 - (B) CCD
 - (C) SAR
 - (D) DSP
5. CCDs that contain two layers of sensors are commonly used in:
 - (A) analog video cameras.
 - (B) digital video cameras.
 - (C) digital cameras.
 - (D) Both (A) and (B).
6. The main components of a dynamic microphone include:
 - (A) two plates and a power source.
 - (B) a diaphragm, wire coil and magnet.
 - (C) a capacitor, comparator, DAC and SAR.
 - (D) a wire coil, ADC and DSP.
7. Which of the following is true for progressive scan CCDs?
 - (A) Every second line of each image is retained.
 - (B) The data is collected to suit normal analog television.
 - (C) The entire contents of the CCD is read for each image collected.
 - (D) They use interlacing to reduce the amount of data.
8. Mechanisms on traditional cameras to control the amount of light entering the camera include:
 - (A) altering the size of the aperture and the time shutter is open.
 - (B) altering the size of the shutter and the time the aperture is open.
 - (C) the use of different types of film that have varying sensitivities to light.
 - (D) moving the lens in and out to focus the light more accurately.
9. One issue to consider when using vehicle counters based on a single air pressure switch is:
 - (A) the road temperature and air temperature commonly cause false readings.
 - (B) they are unable to produce digital data.
 - (C) they cannot differentiate between vehicles with two axles and larger vehicles with more than two axles.
 - (D) the vehicle must be stationary if the magnetic field is too influence the induction loop.
10. The Safe-T-Cam system uses a video camera and a still camera. Why is this?
 - (A) Still images are used to isolate individual vehicles from the background, video is used to determine the registration.
 - (B) Video is used to isolate individual vehicles from the background, still images are used to determine the registration.
 - (C) The still images are used as a backup should the video camera fail.
 - (D) This is not true, Safe-T-Cam only uses a video camera..
11. Describe the operation of a digital camera.
12. Describe the operation of a condenser microphone.
13. Explain how an ADC performs its conversion using the services of a DAC.
14. Explain the differences between CCDs used in digital still cameras and those used in video cameras.
15. Research and describe how speed cameras operate.

SOFTWARE USED FOR COLLECTION

Software is comprised of instructions that control the hardware and direct its operation. There are two general types of software present in all computer systems; system software and application software. System software includes the operating system and device drivers for each hardware device. In regard to collection, these software components communicate with each hardware collection device and with the current software application. That is, they provide the interface between application software and hardware collection devices. Application software, such as word processors, spreadsheets and databases, receive the collected data and process it during other information processes.



Software

The instructions that control the hardware and direct its operation.

In many cases the source of the data is the user, and in virtually all software applications the user controls or initiates the collection process; the application software must display data entry forms to enable the efficient collection of data. Therefore software performs two essential tasks during the collection process; it provides an interface with the collection hardware and it provides a mechanism for data entry.

Existing data is often moved from other systems rather than being collected directly. For instance, a large variety of different data is available from the Australian Bureau of Statistics. This data can be converted into a form suitable for use by a variety of different applications. We consider a variety of examples where participants can move data between different applications.

In this section we consider:

- Device drivers that allow hardware to interface with the operating system.
- Software that allows participants to enter data.
- Software that allows participants to import data.

DEVICE DRIVERS THAT ALLOW HARDWARE TO INTERFACE WITH THE OPERATING SYSTEM

Most hardware collection devices, and most other devices connected to computers, are used by a variety of different software applications. It would be most inefficient for each software application to contain its own set of instructions for communicating with each device. It makes more sense to store all the instructions required to control and communicate with a particular device separately. These software programs are called 'device drivers'. Most hardware manufacturers develop and distribute device drivers specific to each of their hardware devices. For example, if you buy a new printer, the packaged software accompanying the device will contain a device driver from the manufacturer. Most operating systems also include various different device drivers that are capable of communicating and controlling common hardware devices, however advanced features of specific devices may not be supported. For example the generic scanner device drivers supplied with an operating system are unlikely to support advanced features such as document feeders.



Device driver

A program that provides the interface between the operating system and a peripheral device.

When a hardware collection device, and most other peripheral hardware devices, wish to send data to the system they communicate via their device driver; the device driver provides the software interface between the hardware device and the operating system and application software. Messages to control the process are sent by the hardware to the device driver which in turn communicates with the operating system. The operating system determines which software application requires the data and informs the software application. The software application then receives the data from the hardware collection device via the device driver. The operating system controls the whole process by communicating with the device driver and the software application; *Fig 3.29* describes this process.

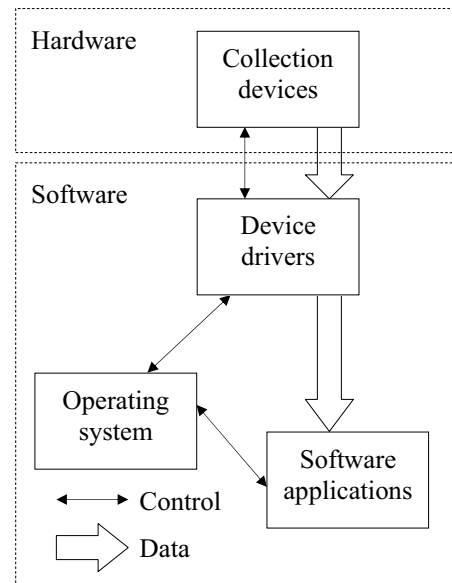


Fig 3.29

The software interface between collection devices and software applications.



Consider the following:

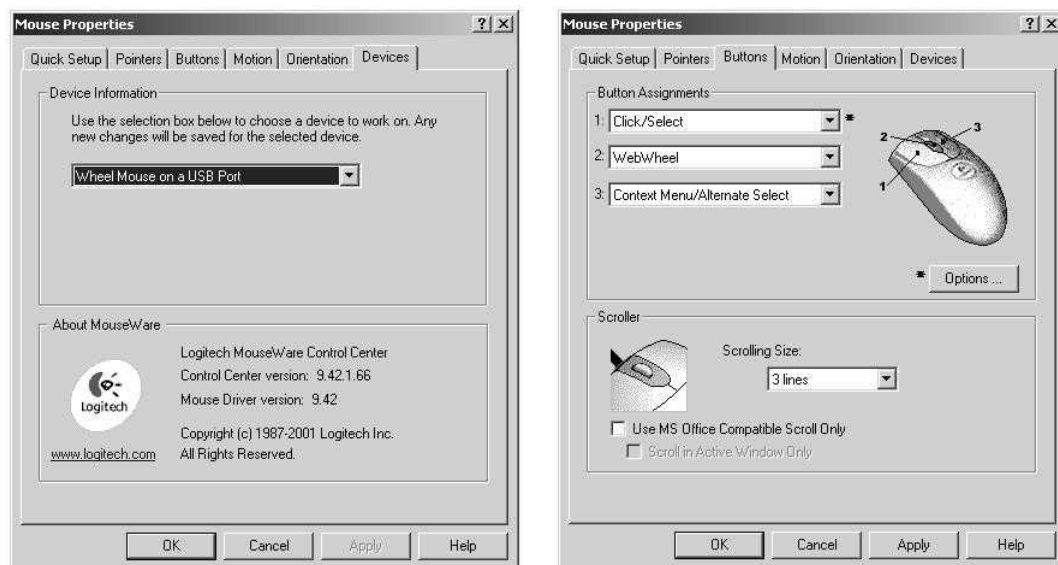


Fig 3.30

Screen shots from the user interface of a Logitech mouse driver used in Windows XP.

The screen shots in *Fig 3.30* above form part of the user interface of a Logitech mouse driver designed for use with the Windows XP operating system. These screens allow the user to alter characteristics of the mouse's device driver which in turn affects the operation of the mouse within all software applications.



GROUP TASK Discussion

Discuss how changes made to the above screens cause the mouse to operate in the same way within all software applications installed on the computer. Consider the flow of data from the mouse until it reaches the application as part of your discussion.



Consider the following:

The screen shots in *Fig 3.31* are for a keyboard installed in Windows XP via a USB port. This keyboard is described as an 'HID Keyboard Device'; HID is an acronym for 'Human Interface Device'. HID devices include most common hardware devices for collecting data from the user; this includes barcode scanners, mice, joysticks and keyboards. HID is a standard that forms part of the USB standard. HID compliant devices do not require their own device drivers; rather they use the HID device driver included with the operating system.

When an HID compliant device is first plugged in, the HID device driver accesses data from the device in regard to its operation. This data provides information specific to the particular device in a similar way that a dedicated device driver provides specific information about a device. As a consequence specific functionality available in the new device can be used without the need to install a dedicated device driver.

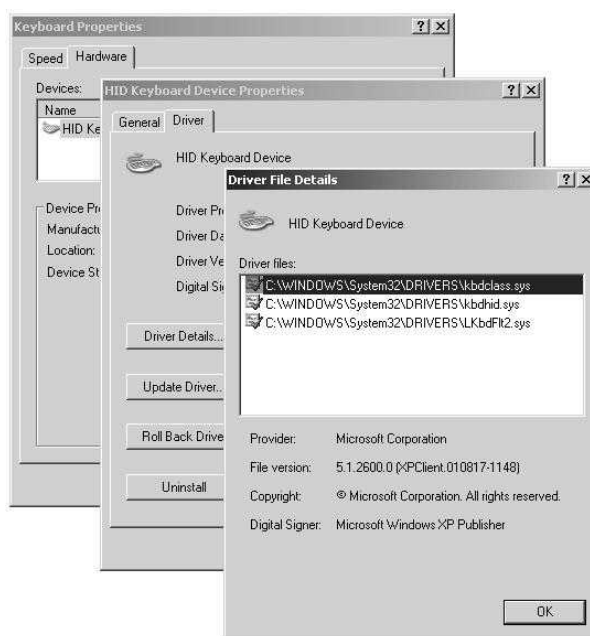


Fig 3.31
Driver details for a standard HID
keyboard in Windows XP.



GROUP TASK Investigation

Examine the device drivers for each hardware collection device installed on either your school or home computer. Do any of these devices utilise the HID standard?



GROUP TASK Discussion

List and describe advantages and disadvantages of standards, such as the HID standard, for both users and hardware manufacturers.

SOFTWARE THAT ALLOWS PARTICIPANTS TO ENTER DATA

Virtually all software applications collect data directly; hence a means for data entry must be provided. Most applications use the keyboard and mouse in conjunction with data entry forms displayed on the screen to enter data; this is particularly the case for text and numbers, together with computer generated image, audio and video data. Even image, audio and video data collected using various other hardware collection devices use the keyboard and mouse to collect data to control the operation of the device.



Application software

Software that performs a specific set of tasks to solve specific types of problems.

The aim of all data entry screens is to collect data in the most accurate and efficient manner. The accuracy of data within a system is known as 'data integrity'. Data integrity is the aim of all information systems; the data needs to be correct and accurately reflect its source at all times. High quality data has high levels of data integrity. Achieving high levels of data integrity is a time consuming and difficult task for most information systems. It is an ongoing process whereby new and existing data is repeatedly checked for accuracy, not just at the time of collection but throughout the life of the data.



Data Integrity

Occurs when data is correct and accurately reflects its source. The quality of the data.

Ensuring the integrity of data during collection is accomplished using data validation checks as well as data verification checks. The efficiency of data entry processes is largely determined by the design and behaviour of the data entry screens; the user interface. Let us consider data validation, data verification and some basic principles in regard to user interface design.

Data Validation

The computer performs data validation as each data item is entered. Data validation ensures each data item is reasonable. For example, when entering the cost of a product, data validation criteria would likely include checks for a number and ensuring that the number is positive. Data entry screens often use self-validating components that ensure only valid data can be entered. For example sets, of radio buttons restrict the range of data that can be entered to one of the available choices, hence radio buttons are said to be self-validating.



Data Validation

A check, at the time of data collection, to ensure the data is reasonable and meets certain criteria.



GROUP TASK Activity

Examine data entry screens from software applications installed on your school or home computer. Describe the different types of data validation used for each component on these screens.

Data Verification

Data verification is a more difficult task than validating the reasonableness of the data. For example, the computer can quite easily check that a phone number contains the correct number of digits however verifying that these digits are indeed the persons phone number is a more difficult task. Furthermore, people often change their phone numbers therefore data verification must be ongoing. Data verification includes all the procedures that are used to verify the correctness of the data within an information system. In regard to data entry into application software, data verification is often implemented as a procedure whereby the user compares the source data to the data just entered. For example, when taking a credit card order over the phone, the operator verifies the credit card number entered by reading it back to the customer.



Data Verification

A check to ensure the data collected matches the source of the data.

User Interface Design

The aim of the user interface is to guide the user through the collection process in such a way that the data is collected accurately and efficiently. The user interface is more than just the placement of components on the screens; rather it provides the total interaction between the user and the software application. In regard to collecting data the user interface displays information to the user to guide them through the collecting process.

There are numerous design factors that influence the efficiency and accuracy of user interfaces. The study of user interface design is itself a complete discipline; nevertheless let us consider some basic principles that could be used when assessing the quality of user interfaces.



User Interface

Part of a software application that displays information for the user. The user interface provides the means by which users interact with software.

- Know who the users are. What are their goals, skills, experience and needs? Answers to these questions are required before an accurate assessment of the user interface can be made. For example, a data entry screen that will be used every day by data entry operators will be quite different to one used infrequently by unskilled users, such as members of the public.
- Consistency with known software and also consistency within the application. Users expect certain components to operate in similar ways and to be located in similar locations. For example, the file menu is located in the top left hand corner of the screen and placing it elsewhere would be inconsistent and confusing. Consistency allows users to utilise their existing skills when learning new software applications.
- Components on data entry screens should be readable. This includes the words used as well as the logical placement and grouping of components. The interface should include blank areas (white space) to visually imply grouping and to rest the eye. Colour and graphics should be used with caution and only when they convey information more efficiently than other means.
- Clearly show what functions are available. Users like to explore the user interface; this is how most people learn new applications, therefore functions should not be hidden too deeply. If a particular function is not relevant then it is better for it to be dulled or greyed out than for it to be hidden, this allows users to absorb all possibilities. At the same time the user interface should not be overly complex.
- Every action by a user should cause a reaction in the user interface. This is called feedback; without feedback that something is occurring, or has occurred, users will either feel insecure or will reinitiate the task in the belief that nothing has happened. Feedback can be provided in simple ways; such as the cursor moving to the next field, a command button depressing or the mouse pointer changing. Tasks that take some time to complete should provide more obvious feedback indicating the likely time for the task to complete.
- User actions that perform potentially dangerous changes should provide a way out. Many modern software applications include an 'undo' feature, whilst others provide warning messages prior to such dangerous tasks commencing. In either case the user is given a method to reverse their action.



Consider the following data entry screen:

Mr Window Cleaner: Clients

Client No: 4016

Name: Company: Parramatta Education Centre
Surname:
First Name:
Title:
Address: Street: 116 Great Western Highway
Suburb: Mays Hill
Postcode: 2145
Phone/Email: Home: 96355517
Work:
Fax: 96355518
Mobile:
Email: pedc@pedc.com.au
Jobs: Latest job: 20/10/03
Client Contacts: Latest contact: 22/09/03
View All Jobs
Add/Edit Contacts

Initial: Quote: \$300.00
Time taken: 8 hours
Includes: ☒ Windows
☐ Frames
☒ Screens
Details:
Schedule

Maintenance: Quote: \$100.00
Regularity: 4 months
Time taken: 2 hours
Includes: ☒ Windows
☐ Frames
☐ Screens
Details:
Schedule

Half Service: Quote: \$59.00
Regularity: 3 months
Time taken: 2 hours
Includes: ☒ Windows
☐ Frames
☐ Screens
Details:
Schedule

Job Notes: General:
Be Aware:
Rating:
☐ Poor
☐ Average
☒ Very Good
Employee: Danny

Select Client to find: Parramatta Education Centre
Client No to find: 4016
Address to find: 116 Great Western Highway
Record: 1 of 31

Fig 3.32

A data entry screen used to collect and display client information.

The above data entry screen is used to collect and display client information for a window cleaning company. The screen is linked to the company's phone system; as the phone rings the information system retrieves the 'Caller ID' and this data is then compared to the phone number details held in the database. If a match is found then the appropriate client details are displayed. At some stage during the call the operator confirms that the data held in the first column on the screen is correct. If no match to the Caller ID is found, then a blank client screen is displayed and in this case the operator must first perform a search operation using one or more of the find components at the bottom of the screen. If the client cannot be located then a new record is created.

Only the left hand column of data is entered during a phone call for a new client, the remaining data is collected via a paper form by the window cleaner assigned to complete the first job or quote. The paper form from the window cleaner provides the source of the data for the remaining data items on the screen.



GROUP TASK Discussion

List and describe techniques used on the above screen and in the above scenario to improve the data integrity of the client data.



GROUP TASK Discussion

Critically evaluate the above data entry screen based on the user interface design principles outlined on the previous page.

Collection via data entry web pages hosted on a web server.

Data entry screens in the form of web pages can be created to collect data from users visiting the web site. These data entry screens perform in much the same way as other data entry screens. One major restriction is the speed with which the data can be validated. Simple data validation can be carried out by the browser within the downloaded web page (often using JavaScript) whereas any validation that requires examining the data source held on the web server will take time to occur and is often implemented as a batch process after a complete page of data has been entered.

**Browser**

A software application that interprets HTML code into the words, graphics and other elements seen when viewing a web page from a web server.

There are various server-side technologies available for generating data entry web pages. At the time of writing many data entry web pages are implemented using server-side scripting languages such as PHP (Hypertext Pre-processor), PERL (Practical Extraction and Report Language) or ASP (Active Server Pages). All these server-side technologies use programming code executing on the web server to generate HTML files for transmission to the user's web browser. Server-side scripts can contain instructions that retrieve data from a database and place it into the page prior to its delivery; similarly the data entered by the user is returned to the web server where it is stored in the database. These server-side technologies mean that specific data can be displayed for the specific user and collected from the user interactively, yet no additional software is required on the user's machine.



Consider the following:

Google is a popular search engine used to search the World Wide Web. The screen shot in *Fig 3.33* shows Google's advanced search data entry screen displayed within Microsoft's Internet Explorer browser.

Self-validating screen components, namely drop down boxes and radio buttons, are used for data validation. When a user clicks the 'Google Search' command button the data entered by the user is sent to Google for processing and the results of the search are then returned for display in the user's browser.



Fig 3.33

Google's advanced search data screen displayed in Microsoft's Internet Explorer browser.

**GROUP TASK Discussion**

Identify the source of the data from the perspective of Google and then from the perspective of a user collecting data from a website found via Google's search engine. Discuss your responses.

SOFTWARE THAT ALLOWS PARTICIPANTS TO IMPORT DATA

Many software applications use data created or originally collected by other applications. The data is moved across for further processing in some other software application. For example, images are collected using a digital camera and then moved to a computer where they are edited using a graphics application. These images are then imported into a broad range of software applications, such as word processors, desktop publishing applications and HTML editors. Data held in databases is often moved to or imported into other applications for specialised processing. For example, word processors use databases when mail merging to create personalised letters. Spreadsheets import data from databases so they can perform statistical analysis or create charts. Web browsers in combination with search engines are used to locate and download data over the Internet. This data can then be used by other software applications. In essence, the destination software application is collecting existing data created by another application.

Importing data often, but not always, involves altering the organisation of the data to suit the needs of the destination software application. Therefore, although importing is essentially a collecting process, it often includes organising and perhaps other information processes as sub-processes. For example, if a photograph is to be used on a web page then the image would likely be resized, compressed and saved using the JPEG format. In this case the conversion changes the data and its file format so both processing and organising processes are performed by the graphics application. The HTML editor imports the resulting JPEG image file without altering its content or organisation. In other cases the source application exports the data in a format understood by both applications. In this case both the source and destination applications reorganise the data. For instance, many database management systems (DBMSs) are able to output delimited text files. These text files can be imported into spreadsheets. In this case, the DBMS converts or reorganises the data from its native format to create the delimited text file. The spreadsheet performs a conversion from the text format to reorganise the data into its own format. There are also scenarios where just the destination or importing application performs the conversion. For example, most word processors include the ability to import files that are in the native format of other word processors – for instance, Microsoft Word can import files in WordPerfect format. In this case Microsoft Word performs the reorganisation as it imports or collects the WordPerfect file.

Web browser and search engines used to locate and download Internet data.

Web browsers are software applications used to collect data and information from web servers. The problem is not so much actually getting the information from the web server to your browser, rather the problem is to locate the web server that holds the information you require. Search engines are websites dedicated to assisting in the task of locating information on the World Wide Web. When using a search engine you are not directly searching the web, rather the search engine queries its own database for possible URLs of relevant web sites. When suitable data or information is located on a website it can be collected by copying and pasting, downloading files or by saving the HTML code. When collecting data in this way it is difficult to control the format of the data and even more difficult to reliably assess its integrity.



GROUP TASK Discussion

Why is it difficult to control the format and assess the integrity of data collected directly from the web using a web browser? Discuss.

SET 3C

1. Software is:
 - (A) all the components of an information.
 - (B) instructions that collect data from the environment.
 - (C) the instructions that control and direct the operation of the hardware.
 - (D) programs that solve a specific problem.
2. A program that provides an interface between a peripheral device and the operating system is called:
 - (A) a HID.
 - (B) a utility program.
 - (C) application software.
 - (D) a device driver.
3. Which of the following statements is FALSE?
 - (A) Data usually travels from the collection device through the device driver and then to the software application.
 - (B) The operating system directs data from the device driver to the software application.
 - (C) Software applications usually communicate directly with collection devices.
 - (D) Most hardware can only communicate with the operating system via their device driver.
4. Data that accurately reflects its source, is said to have high levels of:
 - (A) data validation.
 - (B) data integrity.
 - (C) data security.
 - (D) data verification.
5. The phrase 'self-validating screen component' means:
 - (A) the user is unable to enter inaccurate data.
 - (B) the data entered can never cause an error.
 - (C) it is only possible to enter reasonable and valid data items.
 - (D) the data entered will always be correct.
6. After completing a web-based form the user is presented with a screen asking them to confirm that the data entered is correct. This is an example of:
 - (A) data validation.
 - (B) data integrity.
 - (C) data verification.
 - (D) data accuracy.
7. The user interface can be best described as:
 - (A) all of the various screens displayed during execution of software.
 - (B) the instructions used to control the collection of data into software.
 - (C) the means by which users interact with software. This includes the display of information for the user.
 - (D) The design and placement of components on screens.
8. To import data from a source application to a destination application requires:
 - (A) the data to be in a format that both applications can understand.
 - (B) the data to be reorganised by the importing application.
 - (C) requires the source application to reorganise the data into a format the destination application can read.
 - (D) both applications to be installed on the user's system.
9. PHP is used because:
 - (A) it ensures all users view identical data in an identical format.
 - (B) it allows web pages to be adjusted automatically to suit the individual user.
 - (C) users do not require any extra software, apart from a browser to view web pages.
 - (D) Both (B) and (C).
10. A user interface should include all of the following features EXCEPT?
 - (A) Consistency of design.
 - (B) Feedback after every user action.
 - (C) Bright colours and fonts to attract attention.
 - (D) A method of reversing dangerous actions.

Refer to the following screen when answering questions 11, 12 and 13.

Add/Edit Courses

Course ID: *Must be unique*

Name:

Sort Order: *Number from 0 to 100*

2 classes use this course.

☒ Result 1

Result Type: ☐ A,B,C,D,E, I or N ☒ Numeric

After:

Include on report: ☐ Result ☐ Average ☒ Rank

Stats over: ☐ Class ☒ Course ☐ Calculate rank

☒ Result 2

Result Type: ☐ A,B,C,D,E, I or N ☒ Numeric

After:

Include on report: ☒ Result ☐ Average ☒ Rank

Stats over: ☐ Class ☒ Course ☒ Calculate rank

☒ Application

Assessment Rank: 24/73 Exam Mark: 50% Exam Rank: 24/73 Application: 2

Select Course to find:

Select Course ID to find:

Record: of 137

11. Identify and describe any self-validating components on the above screen.
12. Identify and describe aspects of the above screen that provide feedback to the user.
13. Critically evaluate the above screen. As part of your evaluation include practical suggestions that would improve the design.
14. Describe how software interfaces with hardware collection devices.
15. Collecting data using the Internet involves many of the seven syllabus information processes.
Describe the process of collecting data using the services of an Internet search engine in terms of the seven information processes.

NON-COMPUTER PROCEDURES IN COLLECTING

Commonly the source of data for collection into an information system is a non-computer source. As a consequence non-computer procedures are required as part of the collection process. For example, businesses in Australia are required to complete a BAS (Business Activity Statement) at various times throughout the year. Most businesses complete the BAS on the paper form provided and forward it to the tax office where it is input into their computer system. Various types of procedures are used depending on the nature of the data source and the nature of the data to be collected.

In this section we consider the following non-computer procedures that are commonly used as part of the collecting process:

- Literature searches
- Surveys and interviews
- Form design for data collection
- Manual recording of events
- Existing non-computer data

LITERATURE SEARCHES

Literature encompasses all published works, both in electronic and hardcopy forms. Locating the desired information from traditional hardcopy forms of literature requires manual searches using catalogues to locate possible publications, then contents pages and indexes to locate particular information within each publication. Most libraries maintain a computerised catalogue of all their publications that allow users to search using keywords; it is still necessary to physically locate the publication on the shelves to view the actual contents.

Libraries maintain collections of formally published works; in general the integrity of such literary works is higher than those found on the Internet. The effort of editing and publishing a literary work in printed form is significant and hence accuracy of the contents is likely to be higher. Contrast this with works available on the Internet where anybody can express their ideas with little effort and without the critical eye of an editor.

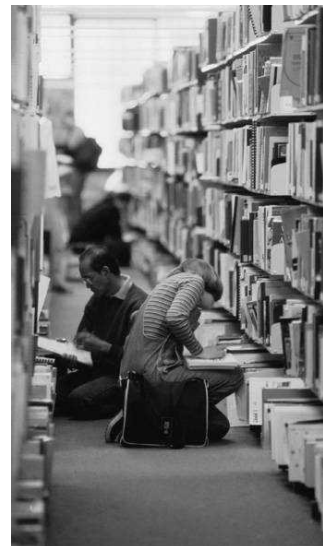


*Fig 3.34
Literature searches are
primarily a manual
procedure.*



Consider the following:

A medical researcher is attempting collect data on the incidence and causes of a particular disease within different populations throughout the world. The results of most medical research are published in medical journals. The data collected via his literature searches will be entered into a computerised information system for processing.



GROUP TASK Discussion

Each block of data found by his literature search is presented in a different format. Discuss problems this situation presents for the researcher.

SURVEYS AND INTERVIEWS

Both surveys and interviews are conducted to collect data directly from people. The basis of both surveys and interviews is questions. Questions can be highly structured, for example, ‘What is your age in years?’ or unstructured, for example, ‘Do you have any further comments?’ If the data collected is to be input into an information system then the analytical nature of the system can cause questions to be biased. For example, it is difficult for computers to analyse free responses, so it is common for questions to be somewhat artificially formulated so that they have only a set number of possible responses. This forces each respondent to choose one of the provided answers when in fact they may wish to answer differently or further elaborate on the reason for their choice. This situation commonly occurs with surveys, however often interviewers are required to make similar judgements whilst listening to responses from interviewees.

Surveys are generally conducted using paper-based forms comprised of various questions, although electronic surveys using the Internet are also common. The purpose of most surveys is to collect data from a large number of individuals and combine the results using various statistical analysis techniques. To allow the results to be efficiently combined means that the responses must be fairly structured, for example, the responses or answers given must be within a particular range.

Interviews, on the other hand, are conducted in person, or by telephone, by an interviewer. The purpose of interviews is generally more individual; the interviewer, or the organisation with which they are attached, is interested in the responses given by individuals rather than the statistics generated from the combined responses of many individuals. For example, in a job interview the purpose is to examine each interviewee’s responses whereas most surveys do not even require individual respondents to be identified. The questions used during interviews can be less structured to allow respondents the opportunity for open-ended answers.



Consider the following:

A survey that includes all members of a population is called a census. The Australian Bureau of Statistics (ABS) performs a census on the entire population of Australia every five years; at the time of writing the last such census was conducted in 2006; *Fig 3.35* is an extract from the 2006 census survey.

The extract in *Fig 3.35* includes various design features to facilitate its completion by individuals and also to enable the efficient input of the data into the ABS’s computer system. For example different shades of grey are used to differentiate various components on the survey.

The image shows a portion of the 2006 Australian Census form. It is divided into two main columns for 'Person 1' and 'Person 2'. The form includes instructions at the top to use CAPITAL letters only. Questions 2 through 8 are visible, covering topics such as names, date of birth, relationship to the household head, marital status, and usual residence. The form uses a grid-like layout with various input fields, checkboxes, and radio buttons. Different shades of grey are used to group related questions and instructions.

Fig 3.35

Extract from the 2006 census from the Australian Bureau of Statistics.



GROUP TASK Activity

Identify design features present on the extract in *Fig 3.35* that facilitate its completion by individuals and also to enable the efficient input of the data into the ABS’s computer system.



Consider the following:

There are many factors that influence the success of the interview process. Most of these factors revolve around the way the interviewer conducts himself or herself during the interview. Following are lists of positive and negative attributes for consideration when conducting interviews:

Positive interviewer attributes:

- Well-prepared questions.
- Attention and careful listening.
- Personal warmth and an engaging manner.
- The ability to sell ideas and communicate enthusiasm.
- Putting the subject at ease.
- Politeness and generosity.
- Focus on the topics that need to be covered.

Negative interviewer attributes:

- Lack of preparation.
- Not allowing enough time for the interview.
- Talking too much.
- Losing focus.
- Letting the candidate direct the conversation.
- Biased towards people with similar ideas and styles to their own.
- The tendency to remember most positively the person last interviewed.



GROUP TASK Discussion

Many of the above attributes of the interviewer are to do with their personal communication skills. Why are such skills so important for successful interviews? Discuss.

FORM DESIGN FOR DATA COLLECTION

In this section we are concerned with the design of paper forms used for collecting data. The design of paper forms shares many similarities to the design of user interfaces for software applications; you must understand the users of the form, there should be consistency throughout the form's design, and each component on the form must be readable. Let us examine some design considerations particular to paper-based form design:

- Often a paper form is used to collect data that will subsequently be input into a computer system; in this case the paper form and the data entry screen need to be structured to assist the data entry process as well as the manual completion of the paper form by the user. Paper forms should not merely be a printout of the corresponding data entry screen; rather both versions should use the strengths of their respective mediums whilst maintaining consistency in terms of the order of data elements.
- Paper-based forms cannot react to a user's responses therefore instructions must be available and clearly stated. General instructions relevant to the whole form should be placed before the questions commence, whereas instructions for particular items should be present at the point on the form where they are needed. For example, if a certain answer means the person must jump to question 9 then this needs to be stated clearly; on a data entry screen the questions that are not needed can be dulled or simply not displayed at all.

- Colour, texture, fonts and the paper itself cannot be altered when using paper forms. Paper forms therefore should be designed so that these elements will work for all, or at least the majority, of users. The paper should be thick enough that type cannot be seen through the page. Consider having large print versions available for sight-impaired users.
- Appropriate space for answers. The space provided for answers on a paper form cannot increase or decrease. Most people use the space provided as an indicator of the amount of information they need to supply. On data entry screens it is possible for such space to grow as needed whereas on paper forms such space needs to be more carefully considered.



Consider the following:

Fig 3.36
Sample Business Activity Statement (BAS) from the Australian Taxation Office.

The above example Business Activity Statement (BAS) is used to collect data in regard to the goods and service tax (GST) and pay as you go (PAYG) tax from all Australian businesses. Each business operating in Australia is required by law to complete a BAS either every month or every three months.



GROUP TASK Activity

Assume you have just started a business and this is your first BAS statement. Critically evaluate the design of the BAS form.



GROUP TASK Discussion

Instructions for completing the BAS are provided in a separate leaflet. Is this an appropriate method for providing such instruction? Discuss.

MANUAL RECORDING OF EVENTS

An event is something that happens at some particular place at a particular time. For example, each time a courier delivers a package an event is occurring, the courier manually records the time they deliver the parcel and they obtain a signature from the person receiving the parcel. Manually recording of events is common in most industries; it is used to monitor activities and to collect data for later analysis.

Often the place in which the data is collected makes it difficult, or inappropriate, for the data to be collected directly into a computer-based information system. For example, nurses spend much of their time manually recording the blood pressure, temperature and various other data for each patient in their care. To collect this data directly into a computer-based information system would require hardware and software at every bed in the hospital. Currently most hospitals have such facilities within their intensive care units but rely on manual recording for other beds.



Consider the following:

It is common in many small businesses for telephone messages and internal memos to be recorded on slips of paper and then to be manually placed on the appropriate employee's desk. This system is often maintained even in offices where each employee has their own computer.



GROUP TASK Discussion

Discuss advantages and disadvantages of using a manual telephone message and memo system over a computer-based messaging system.

EXISTING NON-COMPUTER DATA

Many government and private organisations have extensive records that pre-date the advent of computer technologies. It is a daunting prospect to digitise such large collections; this is particularly the case in regard to image, audio and video data. As a result those wishing to access such records must use manual collection procedures. For example, it is common for researchers of all types to examine newspapers from the period under examination. If the period of time is prior to digital newspapers then either microfiche records or even manually searching through actual hardcopy of each paper must be performed.

Perhaps the most common example of existing non-computer data can be found in libraries. All libraries contain a vast number of books, journals and newspapers; these publications are included in the libraries catalogue, however the actual data itself must be accessed manually. Other examples include historical records held by the land titles office, water board and various other government departments. When data held in such records is required it is necessary for one of the department's employees to manually locate the relevant record and copy it, this process involves significant labour costs which are generally charged to the organisation or individual requesting the information.



GROUP TASK Discussion

It is likely that your teachers maintain a manual markbook of all the results for each of their classes; only assessment marks being entered into a computerised markbook. Discuss reasons why teachers use manual markbooks rather than recording all results on computer.

SOCIAL AND ETHICAL ISSUES IN COLLECTING

In Chapter 1 we discussed general social and ethical issues relevant to information systems, in this section we focus on issues of particular importance during the collecting process.

Social and ethical issues of particular importance during the collecting process are commonly related to:

- Bias within the collection process.
- Inaccuracy of the collected data.
- Failure to acknowledge the source of data.
- Privacy concerns of individuals.
- Ergonomics for data entry participants.

BIAS WITHIN THE COLLECTION PROCESS

Bias is an inclination or preference that influences most aspects of the collection process. The result of bias during collection is inaccurate data leading to inaccurate outputs from the system. Those involved in collecting data must aim to minimise the amount of bias present.



Bias

An inclination or preference towards an outcome. Bias unfairly influences the outcome.

When deciding on the data to collect bias can be introduced. Often incomplete data is collected with the aim of simplifying the system. For example, it is common for loan applications to collect data on a person's income based entirely on their last few tax returns. This data is used to assess each individual's ability to repay the loan; the assumption being that an individual's income is likely to remain relatively constant over time. In fact many people, particularly those who own or operate businesses, are able to adjust their income to suit their expenses. By simply collecting past income data the success of each loan application is biased in favour of salary and wage earners at the expense of business owners.

Locating or identifying a suitable source of data for collection is another potential area where bias can occur. Often efficiency of data collection means that the cheapest or most available source of data is used rather than the best source of data. Consider surveys; the source of data for all surveys should aim to be a representative sample of the entire population. However for ease of collection many organisations now collect survey data from users over the Internet. Internet users, in most cases, are not a representative sample of the population. In general, these users are younger, have higher incomes and possess higher technology skills than the general population. Consequently results derived from such surveys will not accurately reflect the entire population.

The collecting process itself should take into account the likely perceptions held by those on whom the data is collected. People answer questions and fill out forms differently based on their perception of how the data will be used. For example, a survey conducted by the Australian Taxation Office is likely to yield different results to a similar survey conducted by the Australian Bureau of statistics. Individuals would likely perceive the tax office as being interested in their individual responses whereas a survey conducted by the Australian Bureau of Statistics would be viewed as truly anonymous.



Consider the following questions:

1. Have you ever cheated on an examination?
2. Do you enjoy IPT?
3. How many hours each week do you spend studying?
4. Have you ever stolen anything?



GROUP TASK Discussion

Think about your answers to the above questions if firstly they were asked by your IPT teacher and secondly if they were asked by one of your close friends. Discuss the reasons for any differences in your answers. What techniques could be used to minimise these differences?

INACCURACY OF THE COLLECTED DATA

Previously in this chapter we considered techniques for checking the accuracy to improve the integrity of data as it is collected, in particular we considered data validation and data verification checks. In the previous section we considered bias as the cause of various inaccuracies throughout the collecting process. In this section we consider some possible consequences of inaccuracies in collected data.



Consider the following:

1. Despite repeated requests to be removed from a mail order company's database, Fred continues to receive catalogues.
2. Mary receives her monthly credit card statement and finds an entry for a product that she did not purchase. Naturally she contacts the bank, which advises her to sort the problem out directly with the business that sold the product. She tracks down the business; they claim that her credit card number was indeed used to purchase the product over the telephone.
3. A popular current affairs television program conducts a telephone survey on people's voting preferences. Each phone call is charged 55 cents. Once the election is complete it is found that the results of the survey were significantly inaccurate.
4. John is stopped at immigration on his way back from an overseas holiday. He is questioned at length in regard to his activities whilst away from Australia. Eventually immigration determines that the information that led to his questioning concerns a different individual.
5. Julie's pay is credited each week directly into her bank account. Normally Julie has not bothered to check the amount deposited is correct, however this week she does. Julie finds an error, and so she goes back and checks the amount of her previous pays. She finds similar errors have been occurring for a number of months.



GROUP TASK Discussion

For each of the above scenarios: identify likely problems that caused the inaccurate data and suggest techniques that could be used to minimise the chance of such issues occurring in the future.

FAILURE TO ACKNOWLEDGE THE SOURCE OF DATA

In Chapter 1 we examined the *Copyright Act 1968* and its implications when using or copying software applications and databases of information. We found that the laws governing copyright do not apply to the actual information within a database but rather to the work and expense used to gather the information together. This means copyright is breached if the data within an existing database is copied without permission and acknowledgement.

There are various other reasons, apart from copyright, for acknowledging the source of data. Some of these reasons include:

- Justification of outputs. For example, the results from surveys will only be accepted if the source of the data can be shown to be accurate. Describing and acknowledging the data source assists in this process.
- Providing a mechanism for tracking and auditing data. If the source of data is unknown then it is difficult to track and determine the accuracy of the data. For example, audits of financial transactions must be able to determine the precise source of each transaction to check its authenticity.
- Requirements of the source organisation. Secondary data sources are those that provide previously collected data; such sources often require, or at least request, that they be acknowledged when others use their data.



GROUP TASK Discussion

Why would organisations that allow their data to be used as a secondary source wish themselves to be acknowledged as the original source of the data? Discuss.

PRIVACY CONCERNS OF INDIVIDUALS

In Chapter 1, we discussed privacy as being about protecting an individual's personal information. As the collecting process is where personal information first enters an information system then it makes sense to consider specific concerns related to the collection of personal information. Such concerns are addressed directly as part of the first and tenth National Privacy Principles.

NPP1: Collection deals with issues in regard to the collection of any personal information and *NPP10: Sensitive Information* deals specifically with the collection of sensitive personal information such as medical records, criminal records, disabilities, sexual preferences, etc.

Let us examine some practical guidelines organisations should consider to comply with NPP1 and NPP10:

- Collection must be necessary. It is not acceptable for organisations to collect personal data with an expectation that it may prove useful sometime in the future; rather the data must be necessary for the organisation to carry out one of its functions.
- Collection must be fair and lawful and individuals must be informed that collection is occurring. Often it is obvious that personal data is being collected, for example, completing a form. However this is not always the case. For example, personal details on a competition entry form could be used to create a mailing list, this is unlawful with respect to NPP1.

- Individuals should be informed about the purpose of data collection. This includes any legal obligations, such as taxation office requirements, as well as any possible consequences of not providing personal details. For example, “if an ABN is not provided then you will not be considered for this contract.”
- Whenever possible, personal details should be collected directly from the individual; the aim is for individuals to have a clear understanding of who holds their personal details. If secondary sources are used to collect personal information then individuals should be informed of any organisations that will later use the information.
- Sensitive personal information should not be collected without individuals giving their specific consent. Convincing reasons for collecting such information must also exist and be made clear to the individual. For example, a blood bank is required by law to collect data about each donor’s sexual preferences.



Consider the following:

1. You have just purchased a new pair of jeans and the shop assistant asks for your name, address and phone number.
2. On a job application one of the questions asks if you have a criminal record.
3. You subscribe to an Internet newsgroup, which involves entering your email address. Subsequently you begin receiving various marketing emails from other businesses.



GROUP TASK Discussion

Identify privacy issues present in each of the above scenarios. Discuss suitable techniques that should be used during the collecting process to help resolve each of the issues you have identified.

ERGONOMICS FOR DATA ENTRY PARTICIPANTS

Extended periods of time entering data magnifies the possible effects of any ergonomic inadequacies. In Chapter 1, we listed a number of broad ergonomic issues; it may be worth reviewing page 27 to familiarise yourself with these issues.

It is generally accepted that participants who spend more than two hours of their day at a computer workstation are susceptible to health problems including repetitive strain injury (RSI), vision problems and general muscle strain. Most data entry operators spend far in excess of two hours per day at their workstations, so the risk of such health problems occurring is significant.

Vision problems and muscle strains experienced by data entry participants are commonly caused by muscles being held in a single strained position for an extended length of time. In relation to vision problems, the muscles controlling the eyes are focussed at a set distance on the screen. Similarly muscles in the back, shoulders, neck and arms are held still during data entry. Muscles within the body are designed to expand and contract; this movement causes blood to flow freely to each muscle. When a muscle is held in a static position the blood does not flow freely, hence oxygen supply to the muscle is reduced and waste products are not efficiently removed; the result being the pain experienced by the operator. Such problems are rarely long term and can be corrected by improved ergonomics.

Repetitive strain injury, which is also known as occupational overuse syndrome (OOS) is a much more serious problem. RSI is caused by continually performing the same task; muscles and tendons are not designed for such repetitive tasks. Almost any part of the body is susceptible to RSI, however in relation to computer users; arms, wrists and fingers are the most likely victims. The most common type of RSI caused by repetitive keyboard use is called “Carpal Tunnel Syndrome”. The carpal

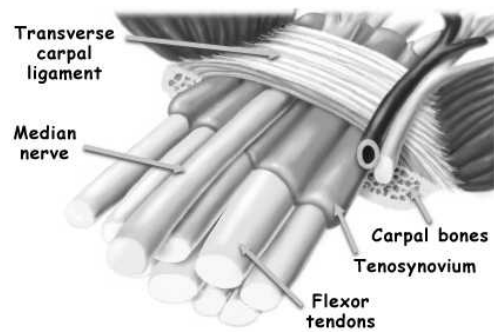


Fig 3.37

Detail of the carpal tunnel showing tendons, median nerve and tenosynovium.

tunnel (see Fig 3.37) is within the wrist and is surrounded by the transverse carpal ligament; this ligament surrounds most of the tendons that operate the fingers. When these tendons are overused the lubricating sheath (tenosynovium) around each tendon swells causing restrictions within the carpal tunnel. Such restrictions result in pressure being applied to the median nerve within the carpal tunnel; the result being the characteristic numbness of the fingers experienced by sufferers. RSI, and in particular carpal tunnel syndrome, can result in long-term damage, therefore it is vital to prevent such problems occurring using sound ergonomics within the workplace.



GROUP TASK Discussion

Sufferers of carpal tunnel syndrome are often prescribed wrist braces or anti-inflammatory drugs. In severe cases surgery is used to cut the transverse carpal ligament. How do you think each of these techniques could help relieve the symptoms of carpal tunnel syndrome? Discuss.

The two most significant ergonomic considerations for preventing vision, muscle strain and RSI problems are the design of the work routine and the design and adjustment of equipment. Let us consider both of these in more detail.

Design of the work routine

The aim is to design a work routine that allows data entry operators the ability to change their physical position regularly and to design tasks that do not require repetitive actions for extended periods of time. To accomplish this aim each data entry operator should be assigned a variety of different tasks and they should have control over the order in which they complete these tasks. Tasks that do require significant time at the keyboard should be interspersed with other tasks or with rest breaks.

Everyone has different needs in regard to their most suitable work routine; therefore it is not appropriate to insist that all operators complete tasks in the same order and for the same length of time, rather each operator should have the freedom to complete their tasks in the order that best suits their needs. Structuring the work routine to suit each individual not only assists in directly preventing injuries but it also increases job satisfaction for each operator. Improved levels of job satisfaction lead to increased productivity, resulting in increased profits for the business.



GROUP TASK Discussion

Studies have shown that employees with higher levels of job satisfaction experience a far lower number of workplace injuries. Suggest likely reasons why this is the case.

Design and adjustment of equipment

The importance of ergonomically designed and adjusted equipment increases with the amount of time spent at the computer. As data entry operators spend more than two hours at a computer workstation they should take particular care in regard to the design, placement and adjustment of the equipment they use.

Chairs, desks (or keyboard) and monitors should all be height adjustable. Often desks do not allow their height to be adjusted, if this is the case then a footrest of variable height may be required. Many desks designed for computer usage incorporate a height adjustable keyboard panel.

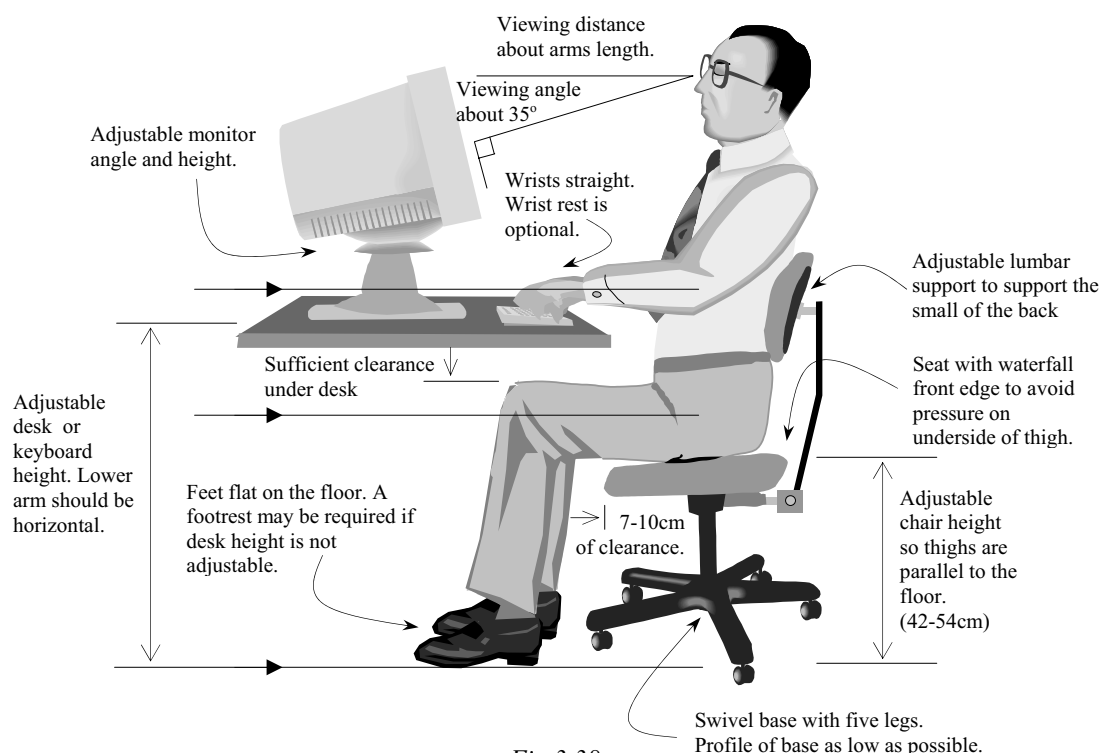


Fig 3.38

Features of an ergonomically sound workstation.

Fig 3.38 shows many of the recommended adjustments and features that should be present in the design of an ergonomically sound workstation. The chair height should first be adjusted so that the thighs are parallel to the floor when the feet are flat on the floor. Next the height of the desk (or keyboard) is adjusted so the forearms are parallel with the floor when using the keyboard. Finally the monitor height is adjusted so the centre of the screen is viewed approximately 35 degrees below the horizontal. Various minor adjustments can then be made to each individual piece of equipment to ensure all muscles are maintained in a relaxed state.



GROUP TASK Activity

Try to adjust the equipment for either your school or home computer workstation to comply with Fig 3.38. List any items that are not present or adjustments that are not possible.



GROUP TASK Discussion

In Chapter 1 on page 27 we listed other ergonomic considerations apart from work routine and equipment design and adjustment. Discuss the significance of these other considerations for data entry participants.



HSC style question:

A market research company has developed a paper-based survey asking people to rate specific products they use on a ten-point scale from poor to excellent. Surveys will be answered anonymously. However, each survey includes a unique Survey ID number. An extract from an example survey is reproduced below – actual surveys will typically contain hundreds of products. It is anticipated that hundreds or even thousands of people will complete surveys.

Survey ID: 345289509		I use this product	Rating Poor Excellent									
Code	Product		1	2	3	4	5	6	7	8	9	10
1436	Heinz baked beans											
2845	Black and Gold baked beans											
1865	Homebrand baked beans											

- (a) The intention of the survey was for people to rate only the products they use. During the collection process it is found that some people have rated products they do not use and others have not rated products that they do use. Explain how such issues could have been avoided using an online survey form.
- (b) During data collection using the paper forms it is noticed that many individuals do not use the full range of possible ratings. For example, some people rate all products as either poor (1) or excellent (10), others use a small range of ratings such as from 3 to 7 and many use just three ratings, often 2, 5 and 9.

Describe likely effects on the survey results and describe possible techniques for minimising these effects.

Suggested Solution

- (a) Validation during online entry would avoid these issues. The software could dull the rating field so that users cannot proceed to enter the rating until a product has been selected as one they use - perhaps from a list box. Equally, once a product has been selected, the user cannot proceed to select another product until a rating has been clicked on or entered for the current product.

The problem with the anonymous paper-based form is that once the form is sent in, it is impossible to check back with the person what their intentions were, if data is not entered correctly.

- (b) The effect on the results of the survey is that any findings will not necessarily be valid. The problem is likely due to the fact that 'poor' and 'excellent' are qualitative measures that do not have the same meaning for all people. This is made more complex by providing a scale of 1 to 10, when perhaps a smaller range, say from 1 to 5, along with descriptions for each rating might be easier for people to intuitively use. Currently it seems people are using different numbers to mean the same thing, for example, 10, 9 and 7 are all maximums for some people. It is therefore unclear if averaging the ratings for each product will accurately reflect people's overall satisfaction with each product.

Another possible strategy could be to statistically adjust each individual's ratings so they cover a more typical range. This may increase the accuracy of the results, however it may also have the opposite effect, causing the results to be skewed.

SET 3D

1. The data and information contained in formally published books is, in general, more accurate than data and information sourced using Internet. One reason for this is:
 - (A) The Internet is susceptible to viruses that can easily corrupt data.
 - (B) In general, formally publishing a book requires significantly more effort.
 - (C) It is often difficult to determine the source of data available via the Internet.
 - (D) Once a book has been published its contents cannot be altered.
2. Surveys and interviews are used to:
 - (A) collect data from secondary sources.
 - (B) collect data directly from people.
 - (C) collect data from all members of a population.
 - (D) collect data from a random sample of the population.
3. A census can be best described as:
 - (A) a survey that is completed by a random sample of the population.
 - (B) a survey conducted every four years by the Australian Bureau of Statistics.
 - (C) a survey that is completed by all members of a population.
 - (D) a statistical analysis technique that summarises the results of a survey.
4. Selling a database containing personal information on individuals could be allowed if:
 - (A) the company selling the database actually collected the data directly from the individuals.
 - (B) none of the data is of a sensitive nature.
 - (C) the data is not necessary for the purchasing company to carry out its functions.
 - (D) the individuals, whose personal information is in the database, have been informed of any organisations who will purchase the data.
5. It is generally accepted practice to include instructions for paper-based forms:
 - (A) at the start of the form.
 - (B) where they are needed within the form.
 - (C) as a separate reference document
 - (D) (A) for general instructions and (B) for specific instructions.
6. The amount of space left for answers on paper-based forms:
 - (A) should be the same for all questions.
 - (B) should reflect the amount of information required.
 - (C) can be changed as an individual completes the form.
 - (D) should be adjusted to enhance the overall look of the form.
7. Carpal Tunnel Syndrome:
 - (A) is another name for RSI.
 - (B) is the most common form of RSI experienced by data entry operators.
 - (C) is caused by muscles being held in a static position.
 - (D) can be easily corrected by improving the design of workstation furniture.
8. When developing surveys many researchers have a theory that they wish to be supported using evidence from the survey. Surveys created for such purposes:
 - (A) are likely to be biased if the researcher designs the survey.
 - (B) should be designed by individuals who do not have an expectation that one outcome is more likely than another.
 - (C) should collect data from a random sample of the population or from the entire population.
 - (D) All of the above.
9. Which of the following contains only positive interviewer characteristics?
 - (A) Remembering the last person interviewed more positively, letting the candidate direct the interview.
 - (B) Well-prepared questions, talking too much, putting the subject at ease.
 - (C) Careful listening, politeness and generosity, focusing on the topics to be covered.
 - (D) Personal warmth and engaging manner, not allowing enough time for the interview.
10. The main aim of adjusting furniture and equipment correctly is to:
 - (A) reduce the amount of stress experienced by users.
 - (B) ensure all muscles are maintained in a relaxed state.
 - (C) reduce the number of repetitive movements performed by users.
 - (D) increase the amount of time users can spend at the keyboard.

11. The design of paper forms for data collection shares many aspects common to the design of user interfaces, however there are significant differences. Describe differences in the way paper forms should be designed compared to computerised user interface forms.
12. Imagine you are conducting interviews to fill a position for a data entry operator. Devise a list of questions suitable for such an interview.
13. Explain how the symptoms of carpal tunnel syndrome result from repetitive overuse of the fingers.
14. Most libraries maintain a computerised catalogue of each of their resources, however the resources themselves are not held in digital form.
 - (a) Explain why most libraries do not digitise all their resources.
 - (b) Describe how data held in non-computer based library resources can be located.
15. List and describe any social and ethical issues apparent in each of the following scenarios:
 - (a) A researcher is conducting a survey to determine the current population distribution of an endangered species of bird. The researcher sends out a survey form to each landholder within the region in which the bird has previously been encountered. The landowners are requested to note the number of individual birds of the species they encounter, together with other details in regard to each sighting.
 - (b) A credit card company sends out letters offering to increase the credit limit for a selected number of their cardholders. The cardholders who are offered the increased credit are selected based on their income, past purchasing history and poor payment history; these are the most profitable customers for the credit card company.
 - (c) Mary works for a telephone sales company. She is required to work 10 hour shifts, after every 2 hours she is scheduled a 10 minute rest break. Mary's job entails making phone calls and recording the result of each call into a database, she is only permitted to talk to her supervisor during each shift.

CHAPTER 3 REVIEW

1. Hardware devices for collecting image data include:
 - (A) scanners, digital cameras and camcorders.
 - (B) keyboard, mouse and barcode readers.
 - (C) barcode readers, microphones and digital camcorders.
 - (D) CCDs, pressure switches and USB ports.
2. LEDs are used to assist the collecting process in many:
 - (A) keyboards, mice and scanners.
 - (B) mice, barcode scanners and flatbed scanners.
 - (C) digital still cameras, camcorders and web cameras.
 - (D) analog to digital conversion processes.
3. Microphones collect audio data and organise it into:
 - (A) analog sound samples.
 - (B) digital sound samples.
 - (C) digital electrical energy.
 - (D) analog electrical energy.
4. CCDs output:
 - (A) digital electrical energy.
 - (B) analog electrical energy.
 - (C) digital light.
 - (D) analog light.
5. Flatbed scanners do not require aperture and shutter speed controls because:
 - (A) they use a Bayer filter to control the light collected, hence modification of the light is simply not needed.
 - (B) images collected using flatbed scanners do not emit light.
 - (C) only a single image is being collected at a time so it is not necessary to change these settings.
 - (D) flatbed scanners produce their own light source and the image is always a constant distance from the CCD.
6. Application software:
 - (A) provides the interface between hardware devices and software.
 - (B) performs a specific set of tasks to solve specific types of problems.
 - (C) is used to control and coordinate the functions of a computer system.
 - (D) is used to manage the hardware and software resources of the system.
7. Users interact with computer systems via:
 - (A) collection devices.
 - (B) the user interface.
 - (C) the keyboard.
 - (D) application software.
8. Collecting data from Internet users via a web page:
 - (A) requires that each user's computer must be a web server.
 - (B) requires a data entry form to be stored on or created by a web server.
 - (C) means that each user must install the appropriate data entry software on their machine.
 - (D) is not possible as web pages are only able to display information.
9. Data integrity is a measure of:
 - (A) the accuracy of the data.
 - (B) the validity of the data.
 - (C) the ability of the system to update its data.
 - (D) how often the data needs to be analysed for errors.
10. Health concerns for participants entering large volumes of data include:
 - (A) vision problems, general muscle strain and RSI.
 - (B) work routine and design and adjustment of equipment.
 - (C) lack of job satisfaction leading to various workplace injuries.
 - (D) privacy, copyright and security issues.
11. You have been assigned the task of collecting an image of each student attending your school for inclusion in the school's database. Describe suitable collection hardware and software necessary to achieve this task.
12. Text data is commonly collected using the keyboard, however it can also be collected using voice recognition and optical character recognition (OCR). Research and describe the hardware and software needed to collect text using voice recognition and OCR.
13. Draw a diagram to illustrate the essential features of an ergonomically sound workstation.
14. List and describe possible reasons for inaccuracies in data as a consequence of the collecting process.
15. 'The collecting information process does not operate in isolation. Many of the other syllabus information processes must occur as part of the collection of data'.
Do you agree? Justify your answer using at least three specific examples.

In this chapter you will learn to:

- choose the most appropriate format for a given set of data and identify and describe the most appropriate software and method to organise it
- describe how different types of data are digitised by the hardware that collects it
- compare and contrast different methods of organising the same set of data using existing software applications
- use software to combine data organised in different formats
- use data dictionaries to describe the organisation of data within a given system
- assess future implications when making decisions about the way data is organised

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Organising – the process by which data is structured into a form appropriate for use by other information processes

How different methods of organising affect processing, for example:

- letters of the alphabet represented as images rather than text
- numbers represented as text rather than numeric

The way in which the hardware used for collection organises data by digitising image, audio, video, numeric and text

Software for organisation, including:

- paint and draw software that allows image manipulation
- mixing software for audio manipulation
- video processing software that allows arrangement of video and audio clips on a timeline
- word processors and desk top publishing for the arrangement of text, images and numbers for display
- spreadsheets for the arrangement of numeric data for processing
- website creation software that uses hyperlinks to organise data to be displayed in web pages
- presentation software allowing data to be arranged on slides, providing control over the sequence in which information is displayed

Non-computer tools for organising

- hard copy systems such as phone books, card catalogues and pen and paper forms
- pen and paper methods for organising data

Social and ethical issues associated with organising, including:

- current trends in organising data, such as:
 - the increase in hypermedia as a result of the World Wide Web
 - the ability of software to access different types of data
 - a greater variety of ways to organise resulting from advances in display technology
- the cost of poorly organised data, such as redundant data in a database used for mail-outs

TOOLS FOR INFORMATION PROCESSES: ORGANISING

The organising information process prepares the data for use by other information processes. It does this by structuring and representing the data in a form suited to the needs of the subsequent information process. We discussed the meaning of structuring and representing as it applies to the organising process in Chapter 2; it may be worthwhile reviewing that section. The organising process does not alter the actual data, rather it modifies the way it is structured and represented; the data itself is still the same. For example, when text is entered into a word processor it is organised by structuring it into a string of individual characters where each character is represented using a binary code (commonly an extension of the ASCII system). The data is still the same characters, words, sentences and paragraphs that the user entered; it has just been organised into a suitable form for the word processor.

Organising takes place just before, just after or even as an integral part of other information processes. This is particularly the case in regard to collecting, displaying, storing and retrieving, and transmitting and receiving. During and after collecting data must be organised to modify its format to suit the requirements of subsequent information processes. During and prior to displaying information it must be organised into a form that can be used and understood by the display device. When storing data it is first organised into a format suitable for storage and subsequent retrieval. Transmitting involves the reorganisation of data to conform to the rules required for communication; receiving data involves reversing this organisation process.

The ability to analyse and process data efficiently depends on the way the data is organised. For example, when a page of text is scanned each character is part of an image. This method of organising text is inappropriate if you wish to subsequently edit the text itself. The scanned image of the page of text needs to be reorganised into individual characters in preparation for editing. In this example optical character recognition (OCR) software could be used to organise the image into a series of characters that can be used by a word processor. The selection of software that is able to organise data appropriately is critical to the success of all computer-based information systems. If the data is organised appropriately then vital analysing and processing tasks can be completed more efficiently.

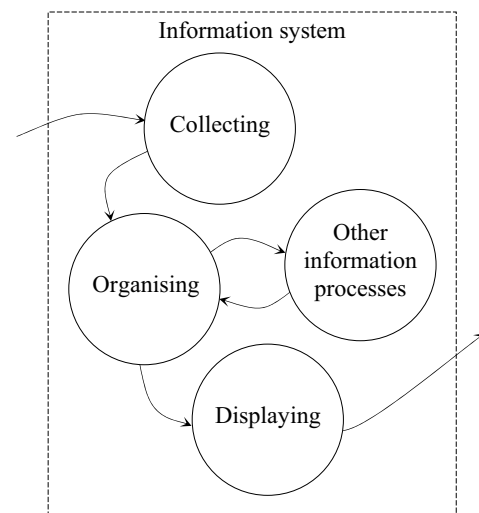


Fig 4.1

Organising prepares data for use by other information processes.

Our focus in this chapter is on the strengths and weaknesses of various different software tools, the aim being to make informed decisions when selecting software tools to use within an information system. The method a software application uses to organise data determines the type and efficiency of processing that can take place. We therefore need to understand how different software applications organise data.

We shall examine examples of the following types of software:

- paint and draw software for images.
- mixing software for audio.
- video editing software for video and audio.
- word processors and desktop publishing for text, images and numbers.
- spreadsheets for numeric.
- database software that organises data into tables.
- website creation software that uses hyperlinks to organise data for web pages.
- presentation software that arranges data on slides.

We then consider some tools used to organise non-computer data and finally consider a number of social and ethical issues associated with the organising process.



Consider the following:

The collection of analog data into information systems involves organising the data into an appropriate digital format. Analog to digital conversion, although an integral part of the collecting process is primarily an organising process, the data is being structured and represented in digital form suitable for use by subsequent information processes. The organising process is not supposed to alter the data, rather it should just modify the way in which it is arranged and represented. Analog to digital conversion in most instances does alter the data, hence more than just organising is occurring. For example, audio data is sampled at precise intervals meaning that the detail of the sound between each sample is lost.

In Chapter 3, we discussed techniques used by various hardware collection devices to convert or organise data into a digital format suitable for use by computer-based information systems. Hardware devices examined included:

- Barcode readers connected between the keyboard and the computer for text.
- Flatbed scanners for images.
- Digital cameras for images.
- Sound cards for audio.
- Digital camcorders for video.



GROUP TASK Activity

For each device above, describe the nature of the analog data input into the device and then describe the nature of the digital data output from the device. How has the data been structured and represented?



GROUP TASK Discussion

Compare the original analog data with the digital data. Has the actual data been altered? Justify your answers by describing the differences or explaining why no differences exist.

THE EFFECT OF ORGANISATION ON SOFTWARE APPLICATIONS

In this section we examine various different types of software application. To structure our discussion we first examine how each type of software application organises its data, we then discuss the types of processes that can be performed as a consequence of this organisation.

Software applications are available to perform almost any task or combination of tasks; it is impossible to include all possibilities. Rather we restrict our discussion to general categories of software commonly used for processing each of the different media types. Within each category there are numerous applications available, each with their own strengths and weaknesses; we restrict our discussion to the major functionality and types of organisation present in most applications within each category.

PAINT AND DRAW SOFTWARE FOR IMAGES

In Chapter 2, we described two different techniques for organising image data, namely as a bitmap or as vectors. Bitmap images are processed using paint software applications and vector images are processed using draw applications. Be aware that many software applications are available that combine the functions of both paint and draw applications; generally such applications are able to include a bitmap image as an object within a larger vector image.

Paint Software Applications

All bitmap images, regardless of their storage format, are processed within paint applications as uncompressed bitmaps. When a compressed bitmap image is opened it is decompressed and organised into a two dimensional arrangement of individual pixels, each pixel representing a particular colour. The processing performed by paint software alters the colour values of individual pixels. When the image is later saved it is first reorganised into the desired storage format. The organisation of the stored data is often quite different to the organisation of the data whilst it is being processed by the application; this is the case for most software applications not just paint applications. For example, JPEG images for use on the web can be stored in such a way that when downloaded, first a low resolution version of the entire image is received, followed by progressively more pixels until the image appears at its full resolution. This organisation of the pixels reflects the often slow and varying speed of Internet connections.

Bitmaps with a colour depth of 24 bits in most cases use the RGB system where 8 bits are used to represent the intensity of red, 8 bits for the intensity of green and 8 bits for the intensity of blue. The RGB system is used as it corresponds directly to the red, green and blue light used by monitors to display images. As 8 bits are able to represent decimal integers from 0 to 255, each colour has a range of intensity from 0 to 255. Fig 4.2 is a screen from Microsoft Paint where the decimal red, green and blue values can be altered directly. The screen

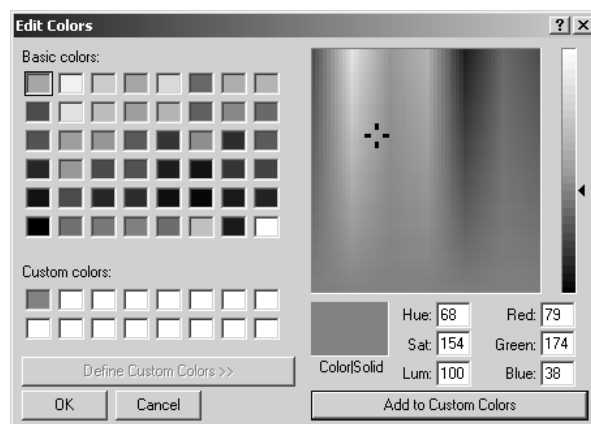


Fig 4.2

If a bitmap image has a colour depth of 24 bits then each pixel is represented using intensities ranging from 0 to 255 for red, green and blue.

also provides the facility to alter colours using hue, saturation and luminance values. These values can be entered directly as decimal integers or the mouse can be used on the colour swatch and luminance bar. Moving horizontally across the swatch alters the hue. Hue is the pure colour within the spectrum of light; it ranges from red through yellow, green, blue and then violet. The saturation is changed by moving vertically up and down within the swatch. Saturation is a measure of the dilution of a hue. Luminance is controlled using the luminance bar and it alters the brightness of the colour. Regardless of the method used to edit a colour it is the red, green and blue RGB values that are used by the majority of paint applications, including Microsoft Paint, to represent the colour of each pixel. Some specialised paint applications use other methods of representation such as the hue, saturation and luminance values (HSL) or CMYK. CMYK is an acronym for cyan, magenta, yellow and key; key really means black, K is used to prevent confusion with the B used in RGB representations. The CMYK system is used within professional printing software applications as cyan, magenta, yellow and black correspond to the primary pigments used on commercial four colour printing presses.

Let us consider the functions included in most paint software applications. For ease of discussion let us assume an RGB system of representation is used. Even moderately sized bitmap images contain thousands of pixels; it would be tedious to alter the colour of each pixel individually. This is what most functions within paint applications do; they alter the colour values of multiple pixels. The processes within paint applications can be broadly split into those that operate without reference to other pixels and those that do consider the colour values of other pixels.

Let us consider examples of processes that operate without reference to other pixels. Most paint applications include a 'negative', 'invert' or 'inverse' function. If this function is applied to an entire bitmap image then every colour value is reversed, this process merely subtracts the current colour value from 255. If a pixel in a 24-bit image has the value 0, 0, 255 then the negative function alters these values to 255, 255, 0. This would alter the pixel from being full intensity blue with no red or green to being no blue with full intensity red and green, hence the pixel would appear yellow. This negative process could equally be applied to a grayscale image (see Fig 4.3), in this case 8 bits are used to represent the intensity of black for each pixel, again subtracting the current value of each pixel from 255 produces the negative image. Because bitmaps are organised as a two dimensional arrangement of pixels it is a relatively simple task to perform this operation on all pixels or even on a selected group of pixels. Many processes within paint applications operate in this manner. Consider a fill operation; a group of pixels is selected, a colour is chosen and the fill tool is selected. When the cursor is clicked within the selected area, all pixels within the area are changed to the RGB values of the chosen colour. Similarly pen, brush and shape tools are used to draw a line or shape on the image; once complete the pixels beneath the line or shape are altered to the same RGB value as the currently selected colour.



*Fig 4.3
The effect of a negative function
being applied to an image.*

Processes that do consider the colour values of adjacent pixels require more involved processing. Consider a process that is used to blur or sharpen the edges of objects within an image. Such a process first needs to identify the edges; this involves comparing the colour values of adjacent pixels. If the colour values vary significantly then it is reasonable to assume the edge of an object has been found. To blur these edges requires the colour values to progressively change over a larger number of pixels than is present in the existing image. To sharpen the edge the number of pixels over which the change occurs is lessened, hence the edge of the object becomes more distinct. *Fig 4.4* shows the effect of these processes; on the left is the original line, the middle line is the result after the edges have been blurred and the right hand image shows the result after edge sharpening. Notice the effect on the pixels near the edge of the line compared to the original.



Fig 4.4
The effect of blurring and sharpening the edges of a line.

Resizing, stretching or skewing bitmap images involves processes that either increase or decrease the total number of pixels in an image. Such processes require the paint software to estimate new colour values for each pixel. For example, if an original bitmap image has a resolution of 200 by 200 pixels and is enlarged to 400 by 400 pixels then instead of 40,000 pixels we now have 160,000; the number of pixels has increased by a factor of 4. Simple enlargement could just produce a block of four pixels with identical colour values to each of the original pixels or it could average the colour values of adjacent pixels when producing the new pixels. Similarly if a bitmap is reduced in size by a factor of 4 the software averages the colour values of each block of four adjacent pixels to produce each new pixel. In either case the image will lose some of its original clarity.



GROUP TASK Discussion

How does the organisation of bitmap data assist paint applications to perform the processes described above? Discuss.



GROUP TASK Activity

Open an image in a paint software application. Try using some of the tools and functions available in the application. List tools and functions that consider the colour of adjacent pixels and those that do not.

Draw Software Applications

Draw applications are used to process vector images. Vector images are composed of a series of different shapes, or objects, whose individual attributes or properties can be changed independently. For example, a circle has attributes such as the colour and thickness of the line, the fill colour used and its position and size relative to other objects. When a new shape is drawn a new instance of an object is created, this object does not affect any existing objects. For example, in *Fig 4.5* a square has been drawn on top of a circle, the circle remains unaltered underneath the square. Vector images are therefore structured as an arrangement of different objects where each object is represented as a particular type with particular attributes.



Fig 4.5
The complete circle exists under the square.

As vector images are represented mathematically they can be resized without loss of clarity. Resizing a vector image does not alter the organisation of the vector image data, rather it alters the size of the bitmap used to display the image. Monitors and printers can only display bitmap images, hence a vector image must be organised into a bitmap of the desired resolution prior to display. The mathematical description of each object allows bitmaps to be created at any resolution without loss of clarity; hence a single vector image can be displayed at the maximum resolution available on the display device.

Processing in draw applications alters the attributes of objects that make up a vector image. Attributes of most objects include; line thickness, fill pattern and colour. Many processes are able to alter the attributes of multiple objects within an image. If a number of objects are selected then it is common for just the attributes present in all the selected objects to be available. Alterations to an attribute are applied to all the selected objects. To simplify this process most draw applications include the ability to group objects together, once grouped then only the attributes common to all the objects in the group can be edited. Furthermore, grouped objects can be repositioned within the image as if they were a single object.

Resizing and reshaping of objects is commonly implemented using handles or nodes. These nodes are significant points on the object used to determine its shape and size. On a rectangle or square the four corner points are sufficient to determine its shape and size, these four nodes need only determine the position of each point, joining the points with straight lines creates the shape. More complex shapes use nodes that contain further information used to determine the shape of the line running through the node. Bezier curves are common objects used in draw applications; in fact many clipart images are entirely composed of Bezier curves.

The shape of a Bezier curve is determined by the attributes of each node. In simple terms each node contains two points, an anchor point and a control point. In most draw applications each of these points can be selected and moved using the mouse. The anchor point lies on the curve, whereas the control point is used to define a straight line to the anchor point. This

line is always a tangent to the curve; it just touches the curve at the anchor point. Longer lines have more influence over the curve; they appear to attract the curve to them, shorter lines tend to repel the curve. The curve shown in *Fig 4.6* is actually two Bezier curves, one from point A to B and another from B to C. The node at B contains three points, an anchor point and two control points. If the curve is to be smooth at B then these three points must lie on a straight line, if they do not lie on a straight line then a sharp corner would be produced. The control point to the left of B determines the shape of the curve as it is produced from B to A and the control point to the right of B determines the curve as it is produced from B to C.

Within many images it is not the actual Bezier curve itself that forms the images rather it is the fill colour applied to the curve that creates the image. Notice that in *Fig 4.6* the actual Bezier curve is of a uniform thickness from point A to point C, lines that occur in nature or that have been drawn by an artist are rarely so uniform. On the other hand the filled section of the curve in *Fig 4.6* changes more naturally. The original brain image used as an icon for many of the group tasks in this text is entirely composed of filled Bezier curves.

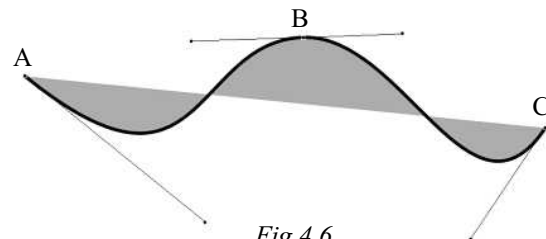


Fig 4.6
Bezier curves are common objects available in most draw software applications.



GROUP TASK Discussion

How does the organisation of vector image data assist draw applications to perform the processes described above? Discuss.



Consider the following:

Right clicking on a Bezier curve drawn within Microsoft Word opens a menu containing an 'Edit Points' function. If this function is selected then nodes appear on the curve, if one of these nodes is then right clicked the screen shown in *Fig 4.7* appears. This screen includes settings that can be applied to the selected node (or point). Selecting 'straight point' ensures both control points and the anchor point lie on a straight line (point D in *Fig 4.7*), a smooth point is a straight point where the anchor point is always in the middle of the two control points (point E) and a corner point allows the control points to be moved to any position (point F). Various other attributes of the curve can be edited by selecting 'Format AutoShape'.

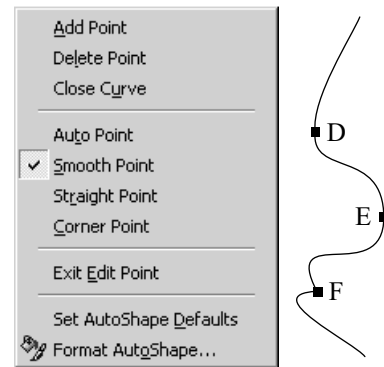


Fig 4.7

MS-Word provides various options for editing nodes on Bezier curves.



GROUP TASK Activity

Open Microsoft Word, or some draw software application. Try to reproduce the Bezier curve shown in *Fig 4.7*. Make a list of the attributes of the Bezier curve and briefly describe their effect.

MIXING SOFTWARE FOR AUDIO

Audio data is organised as either a series of sound samples or as descriptions of various attributes of each individual note. In this section we restrict our discussion to software applications used to process sampled audio data. In Chapter 2, we discussed sounds samples as representing the instantaneous amplitude of a sound wave recorded at precise time intervals; this is how sampled audio data is organised. It is structured as a sequence of separate samples where each sample represents the amplitude of the sound wave at a particular point in time. For example, every second of CD quality sound contains 44100 samples for both the left and right channels, each of these samples is 16 bits long, in effect an integer in the range 0 to 65535. Software applications that process audio data are able to analyse and alter these sound samples; that is they change the integers used to represent each amplitude sample. As is the case with other software applications, most mixing software is able to organise the data into various formats in preparation for storage and subsequent retrieval, however during processing mixing applications operate on the raw sound samples.

Mixing software is used to alter a sound sample and also to combine sound samples from multiple sources. The processes available within mixing applications operate by automating the alteration of multiple sound samples. Sequences of individual sound samples, in almost all cases, need to change progressively; this ensures that a smooth rise or fall in amplitude is maintained. It would be tiresome to manually edit individual sound samples directly; maintaining appropriate differences between each sample would be near impossible. Most mixing software applications display the sound data graphically as a wave. Remember the amplitude, or height of the wave

determines the volume or level of the sound whereas the frequency or number of waves per second determines the pitch of the sound. *Fig 4.8* below shows two screen shots from 'Cool Edit', a mixing software application written by Syntrillium Software Corporation. The left hand screen shows an entire wave, indicating that the level or volume increases at the start of the sound and steadily decreases as the sound finishes. The right hand screen has been zoomed to display just 0.001 seconds of this sound so that the individual sound samples can be seen. The sound displayed on these screens is in stereo (two channels); the top wave represents the sound played through the left speaker and the bottom represents the sound played through the right speaker.

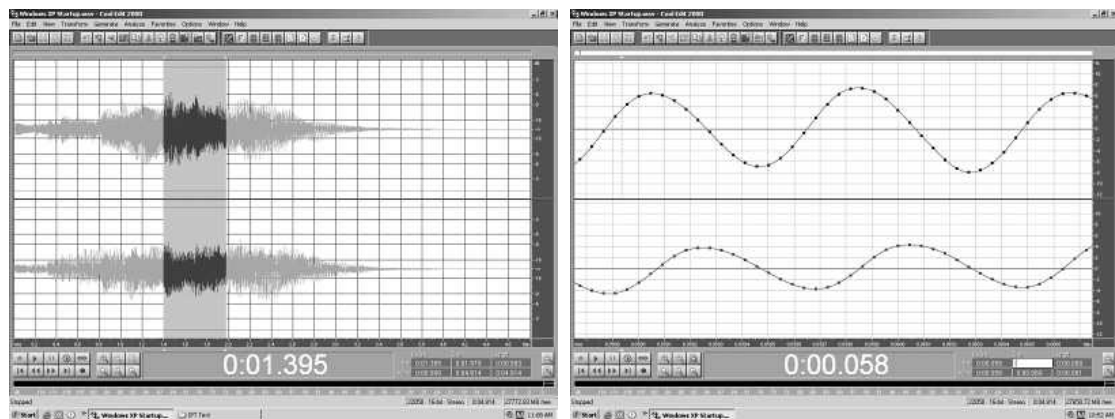


Fig 4.8

The screen at left shows an entire 4.814 second stereo wave form. The right hand screen shows just 0.001 seconds of this wave so the individual sound samples are visible. Screenshots courtesy of Cool Edit by Syntrillium Software Corporation.

The level of a sound is a measure of the relative differences in amplitude. To maintain the fidelity or detail of digital audio it should be collected (recorded) using the widest possible range of amplitudes. For example, if 16 bit samples are used then the loudest sounds in the recording should ideally have a value of 65535. If the level is set low such that the loudest sample is represented as only 32000, then all the sound samples recorded will be compressed to be within a range from 0 to 32000. In effect much of the detail of the original sound will be lost. Often mixing software is used to adjust the average levels of different tracks within a music compilation; this process is called normalisation. Normalised recordings allow listeners to set the volume on their amplifiers once in the knowledge that each track on the compilation will on average play at the same volume.

Some processes in mixing applications alter the sound without analysis of the existing sample e.g. trimming silence, fading in or out, and combining two sounds, whilst others first analyse the sound to determine the nature of the alterations to be made e.g. filtering out noise. Let us consider these processes applied to a single channel and discuss how the raw sound samples are altered.

Trimming is a process similar to cutting, where parts of the sound are removed; in fact the familiar cut, copy and paste functions are also present in most mixing software. Commonly when sound samples are collected they contain initial periods of near silence and they also conclude with a period of near silence. Such samples will appear on the display as areas with low or zero amplitude. In most applications the user highlights the required section of the wave pattern and then initiates the trim function. The trim function removes all the sound samples that do not lie within the selected range; hence the trim function reduces the total number of raw sound samples.

Fades progressively alter the level or amplitude of a sound; a fade-in occurs when the sound level progressively increases and a fade-out occurs when the sound level progressively decreases. Fades do not alter the frequency of the wave so the pitch remains constant, just the volume changes. In *Fig 4.9*, the top wave, which has a constant frequency, has first had fade-in applied and then fade-out applied. Notice that the wavelength and hence the frequency of each wave is the same. In most mixing software it is possible to adjust the nature of the fade using various envelopes – an envelope describes the change in a wave's shape over time. The graphs in *Fig 4.9* show the effect of a simple straight-line envelope being applied. When an envelope is applied each new sound sample is calculated as a percentage of the corresponding old sample amplitude. In the fade-in example in *Fig 4.9* the envelope is a straight line so the percentages used increase constantly from 0 to 100. Envelopes that are not straight lines will vary the percentages used to mirror the shape of the envelope.

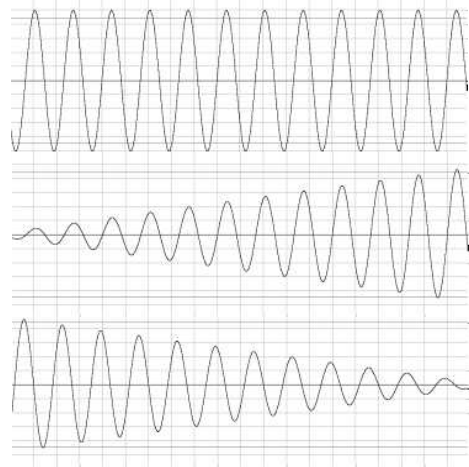


Fig 4.9
Fades progressively alter the amplitude of a wave without affecting its frequency.

Combining two sounds so they play simultaneously is commonly used to add additional instruments or vocals to an existing audio track. Such functions are implemented in mixing software using a special case of the familiar paste function, however instead of inserting the new sound samples into the existing data, the new sound samples are combined or mixed with the existing samples. For example, in 'Cool Edit' the edit menu contains a 'Mix Paste...' function. When this menu item is selected various criteria, including the file to be pasted are specified prior to the mixing process commencing. So how are the raw sound samples altered when one sound is combined with another? Let us consider a simple example where a 200Hz and a 400Hz sound sample are combined (see *Fig 4.10*). Each sound sample represents a particular amplitude at a particular point in time. When two samples are

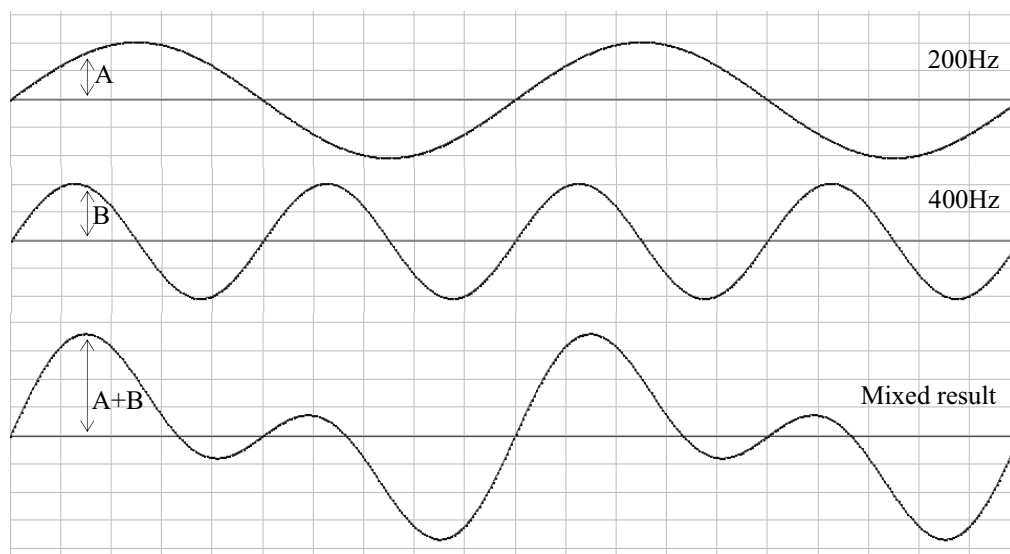


Fig 4.10
Mixing simply adds the raw amplitudes of each sample. If any results are greater than the range the sample size allows, then all the final samples are scaled proportionally.

combined each pair of amplitudes occurring at the same time are added. For example, in Fig 4.10 amplitude A and amplitude B occur at the same time, hence the resulting amplitude is $A+B$. If the sum $A+B$ is greater than the range of values that can be represented (i.e. greater than 65535 for 16-bit samples) then all the resulting values are scaled proportionally. This scaling affects just the amplitude not the frequency of the wave at each point, hence the average level of the resulting sound will change but the pitch will be correct.



GROUP TASK Discussion

How does the organisation of audio data into sound samples assist mixing applications to perform the processes described above? Discuss.

Let us now consider the processes involved to filter out background noise from a sound. Background noise is any unwanted sounds present throughout an audio clip; it commonly includes noise from the environment or generated by the recording equipment. To remove background noise from a clip involves first analysing a section of the clip that should be silent, that is a series of sound samples is examined that represent the noise that is to be removed. In Fig 4.11 suitable groups of sound samples for noise analysis are shown at A and B. These noise samples are analysed to determine the frequencies and levels (amplitudes) present. For example, the analysis may find that the noise contains a frequency of 100Hz that is 2% of the maximum level. Finally the original wave is analysed to find and remove occurrences of these frequencies that occur at the level determined in the noise sample. The processing occurring during noise reduction is complex, as the raw data is not organised into different frequencies and their corresponding levels; these properties must be determined from the raw sound samples. However once these properties have been determined the final alteration of each raw data item is a simple subtraction process. Fig 4.11 shows an original waveform before noise reduction and then the resulting waveform after noise reduction; notice that the sections corresponding to A and B on the original have essentially become straight lines indicating silence.

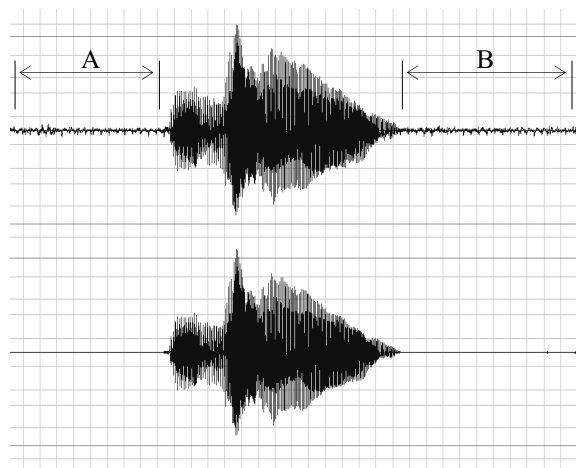


Fig 4.11
A wave form before and after background noise reduction.



GROUP TASK Activity

Record two sounds; say your voice saying hello and goodbye. Use a mixing software application, and your recorded sounds, to perform examples of each of the processes discussed above.



GROUP TASK Research

There are various formats used to organise audio data in preparation for playback (display). Make a list of as many different audio formats as you can and specify the advantages and disadvantages of each format.

SET 4A

1. Organising involves:
 - (A) deciding what to organise.
 - (B) structuring and representing data.
 - (C) altering the data.
 - (D) permanently storing data.
2. Which of the following is true for all examples of the organising process?
 - (A) It formats data in preparation for display.
 - (B) It prepares data for use by other information processes.
 - (C) It determines the storage format for files.
 - (D) It alters the data so it can be understood by humans.
3. Analog to digital conversion:
 - (A) is purely an organising process.
 - (B) is primarily an organising process.
 - (C) is a collecting information process.
 - (D) is a processing information process.
4. In terms of data organisation, the essential difference between a bitmap and vector image is:
 - (A) Bitmaps are composed of individual pixels; vector images describe each shape mathematically.
 - (B) Vector images are composed of individual pixels; bitmaps describe each shape mathematically.
 - (C) Bitmaps require greater storage than vector images.
 - (D) Vector images can be scaled without loss of quality; this is not true of bitmaps.
5. The nodes on an object within a draw application are used to:
 - (A) alter the fill colour, line colour and thickness of the object's outline.
 - (B) move the object relative to other objects within the image.
 - (C) determine the storage format used when the image is saved.
 - (D) determine and alter the shape and size of the object.
6. A 24-bit RGB colour is represented:
 - (A) using 24 bits for each of the red, green and blue components.
 - (B) as a sequence of pixels, where each pixel value is compressed.
 - (C) using a colour table that includes at least 256 different combinations.
 - (D) using 8 bits for each of the red, green and blue components.
7. A line is drawn on top of an existing image, if this line is later selected and moved behind the image then the image must be a:
 - (A) bitmap image.
 - (B) vector image.
 - (C) sequence of pixels.
 - (D) Window's Metafile.
8. A wave displayed in an audio mixing application is really:
 - (A) a sequence of amplitude values joined to form a smooth curve.
 - (B) a sequence of frequency values joined to form a smooth curve.
 - (C) a representation of the raw analog data.
 - (D) a combination of volume samples taken at various periods in time.
9. Combining or mixing two sounds involves:
 - (A) simple subtraction of each corresponding sound sample.
 - (B) simple addition of each corresponding sound sample.
 - (C) multiplying each corresponding sound sample.
 - (D) playing both sounds at the same time whilst collecting new sound samples.
10. Evenly reducing the volume of a sequence of sound samples involves:
 - (A) altering each sound sample to the same amplitude.
 - (B) decreasing the frequency of each wave form.
 - (C) multiplying all sound samples by a number between 0 and 1.
 - (D) multiplying all sound samples by a number greater than 1.
11. Describe an example of the organising process occurring before, after or during each of the other six information processes.
12. In terms of the organising process, what does structuring and representing mean? Provide examples as part of your answer.
13. Describe the organisation of both bitmap and vector images and describe the nature of images suitable for each method of organisation.
14. Describe the organisation of CD quality stereo sampled audio data.
15. "The organisation of bitmap images and sampled audio data is a compromise between quality and storage size." Do you agree? Justify your answer.

VIDEO EDITING SOFTWARE FOR VIDEO AND AUDIO

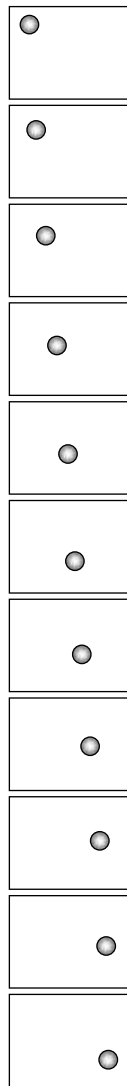


Fig 4.12
Video is essentially a sequence of bitmaps.

Video data, when displayed, must first be organised into a series of bitmap images. By displaying these images in sequence the illusion of motion is created, for example, the filmstrip at right in Fig 4.12. These images may or may not be accompanied by an audio track. In Chapter 2, we discussed techniques for reducing the size of video data to organise it in such a way that it can be more efficiently stored and communicated. This compression and decompression is a common technique for reducing the size of video data files; commonly it is used with video data collected from the real world. The compression techniques discussed in Chapter 2 certainly reduce the amount of data; however the primary method of organising the data is still as a sequence of bitmap images.

**GROUP TASK Discussion**

The overriding aim when organising video data is to reduce the total amount of data required. Why is this important? Discuss.

Advances in hardware design, in particular large fast hard disks, DVDs and fast CPUs, mean that video editing can now be performed on personal computers. Currently both Windows and Macintosh operating systems come pre-packaged with video editing software; Windows computers include Windows Movie Maker and Apple Macintosh computers include iMovie; Fig 4.13 shows sample screenshots from each of these applications. The purpose of these applications is to combine clips into a single video file; a clip being a video file, sound file or even an image.

Let us discuss typical processes performed within video editing software applications, namely joining video clips together, trimming clips and adding transitions between clips. The aim is to describe and justify the organisation of video data within these applications.

Typically editing a video involves joining a number of clips together. Each of these clips is stored in a separate file and potentially contains an enormous amount of raw data. During the editing process each clip remains in its original file; the video editing software knows about each file together with its location. On both screens shown in Fig 4.13 you can see a thumbnail representation of each clip known to the software. For example, on the Movie Maker screen you can see clip 1, 2, 3, 4, 5 and the sound clip called Narration. The timeline at the bottom of each screen is used to arrange these clips into the desired order. For example, on the Movie Maker screen the



Fig 4.13
Sample screenshots from iMovie (top)
and Windows Movie Maker (bottom).

timeline sequence is clip1, then 2, then 3 then 4 then 1 again and then clip 2 again; the Narration sound clip plays for the duration of these clips. At this stage the video editing application has not altered any of the original video data; rather it knows each clip exists and it knows in which order they should be displayed. In essence the video is organised as a particular arrangement of clips where each clip is represented as a file name, location and thumbnail.



GROUP TASK Discussion

The generation of a final edited video file is accomplished as a separate process. It involves specifying the format, resolution, frame rate and method of compression. Why is it necessary for this to be a separate process? Discuss in terms of hardware and data organisation.

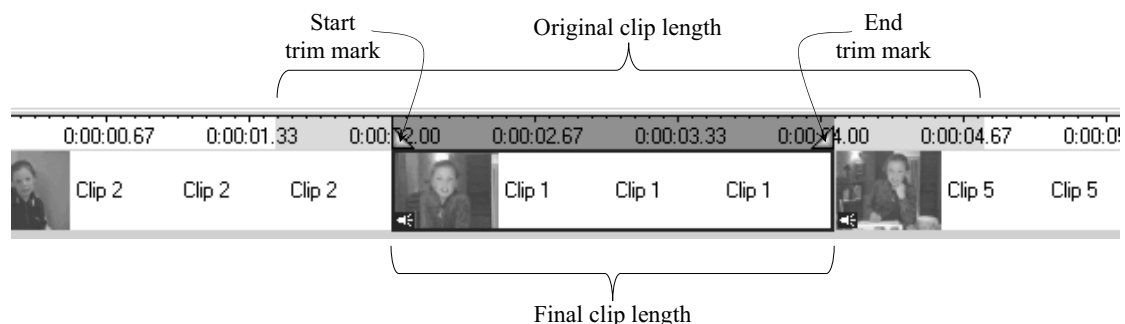


Fig 4.14

Timeline from Windows Movie Maker detailing trim settings.

Trimming is the process of removing parts of a clip, typically from its start or end. This process is implemented in most video editing applications by adding trim marks to clips on the timeline. Start and end trim marks are used to define the portion of the clip to be included in the final video. Fig 4.14 details the method used in Movie Maker to display trim information; in this application the mouse can be used to drag the trim marks to the desired location within the selected clip. Data, in regard to trim marks, forms part of the timeline. So the timeline data contains not just the sequence of clips but also the position of each clip represented as pairs of trim marks.

Transitions occur between clips, they determine how one clip ends and the next begins. With no transition the change from one clip to another is abrupt, particularly when the content of the clips is substantially different. There are many effects that are used to enhance the transition from one clip to another, perhaps the most common being a fade where the previous clip, over a number of frames, turns into the current clip. This process is similar to applying fades to a sound sample except for video a series of frames is faded rather than a series of sound samples. Also rather than fading to nothing, it is more common for one clip to fade into another clip. Fig 4.15 shows the effect of such a fade transition. Notice that each subsequent frame contains less of the first clip and more of the second clip. Each new frame is calculated pixel by pixel, for example a frame that is to be 80% first clip and 20% second clip is created by adding 80% of each colour value in the first clip's frame to 20% of the colour values



Fig 4.15
A fade transition.

in the corresponding second clip's frame. By altering these percentages a complete sequence of fade images is produced.

So how is transition data organised? A transition occurs when two clips overlap on the timeline. This overlap is already represented by the position of the trim marks, that is, transitions occur where an end trim mark for a previous clip is located after the start trim mark for the current clip. If different types of transition are available in the application, then the type of transition must also be specified and represented. The video frames that form the transition are created when the final video is produced.

In most video editing applications the clip and timeline data is saved as a project or reference file. This file contains no actual video data; rather it describes the organisation of the various different video and audio elements in preparation for the production of the final video file. This project file is specific to the video editing software. To produce the final video the user specifies the required resolution, number of frames per second and details of the file format and codec (compression and decompression) to use. The software then creates the final file. The final file could be saved on the hard disk, written to a DVD or uploaded to a website. To view the final video requires a compatible player application which contains details of the codec used to compress the file.



GROUP TASK Activity

Use a video editing application to produce a video file that combines image, video and sound clips. Does this application seem to organise data in the manner discussed above? Justify your response using examples from the application you used.



Consider Animated GIFs

GIF is an acronym for 'Graphics Interchange Format'. GIF is a protocol or set of rules owned and maintained by CompuServe Incorporated. The GIF protocol can be used freely as long as CompuServe is acknowledged as the copyright owner. As a consequence of CompuServe making its specifications freely available, GIF files are one of the most commonly used graphic formats. The GIF specifications includes the ability to store multiple bitmap images within a single file, however sound cannot be included and the number of different colours within an individual image is limited to 256. When an animated GIF file is decoded the images are displayed in sequence to create the animation. The ability to decode all types of GIF files is built into many common software applications, including most web browsers. Most other video formats require their own dedicated software (often in the form of a browser plug-in) to decompress and play video files.

There are many software applications which produce animated GIFs. Fig 4.16 shows the main screen from one such application called 'Easy GIF Animator'.

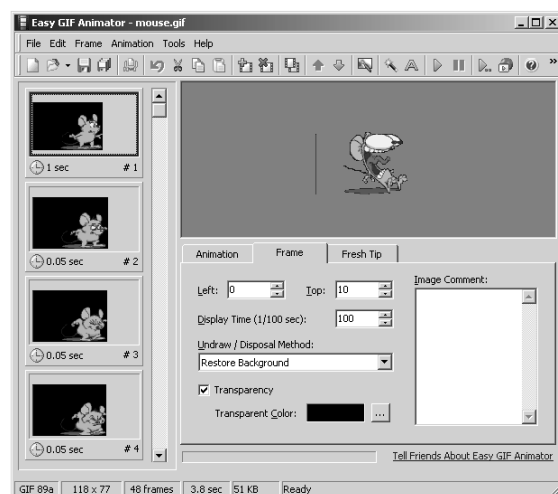


Fig 4.16

Main screen from 'Easy GIF Animator' by Blumentals Software, a Latvian company.

Notice that each frame, or bitmap, in the animation is shown as a filmstrip down the left hand side of the screen. Animation software that produces animated GIF files organise data as a sequence of bitmap images, together with colour palette, timing and various other settings. Most animated GIF software includes functions that produce new frames within the animation. For example, effects that cause an image to slide off the screen or to be progressively enlarged. These functions do not edit existing frames; rather they create a sequence of new frames by repositioning or resizing an existing image.



GROUP TASK Activity

Using an animated GIF application, examine various effects that produce new frames. Describe changes in the new frame's bitmap data compared to the existing frame's bitmap data.



Consider Flash Files

Flash is a standard developed and maintained by Macromedia. In early 2000 Macromedia released details of the flash file format (SWF files) to the public, together with details required to play these files. Flash is now an open standard, as a consequence other software development companies are beginning to produce applications that produce SWF files. For example, SWiSH is one such application developed and distributed by Sydney software company SwiSHzone.com Pty. Ltd. All files created with applications based on the flash specifications must be able to play without error in Macromedia's Flash player. Studies have shown that more than 97% of Internet users have this player installed; in fact it comes packaged with most operating systems and web browsers. With such a large audience Flash has become the de facto standard for delivering interactive animation and sound on the web.

Let us consider the organisation of flash data within SWF files, within Macromedia's flash player and finally for display. Flash or SWF files organise video data by arranging it into definition tags, control tags and actions; an SWF file is a sequence of such tags and actions. Definition tags are commands to the flash player to create and modify characters; a character is like an actor, prop or even the sound track in a movie, they are elements within the animation that are available to be displayed. Each character is maintained in a dictionary within memory. Control tags are used to place instances of these characters on the display list in memory. The order in which characters reside on the display list determines their order when displayed on a frame. A special control tag called ShowFrame is used to instruct the flash player to actually create a bitmap of the frame based on the current display list. Creating interactive flash animations involves responding to user input; in flash this is implemented using events and actions; actions occur in response to events such as clicking the mouse. For example, an action to restart the animation may occur in response to clicking a button.

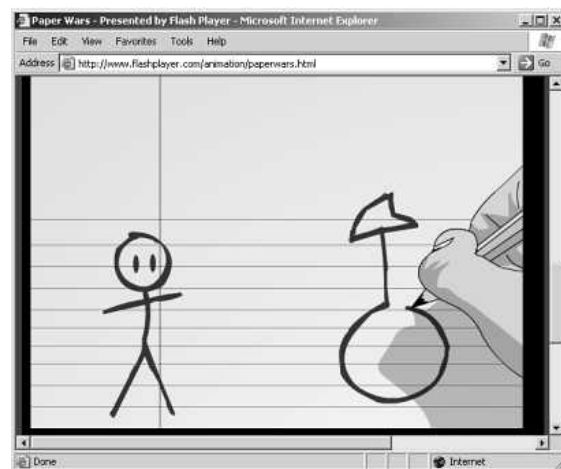


Fig 4.17
A flash animation playing inside Microsoft Internet Explorer.

Fig 4.18 below describes the organisation and processing of a small flash SWF file. The arrows on the diagram indicate the processes taking place as a consequence of each line in the file being executed. The depth values, within the display lists, indicate the order in which characters are placed on the screen: depth 1 first, then depth 2 and so on.

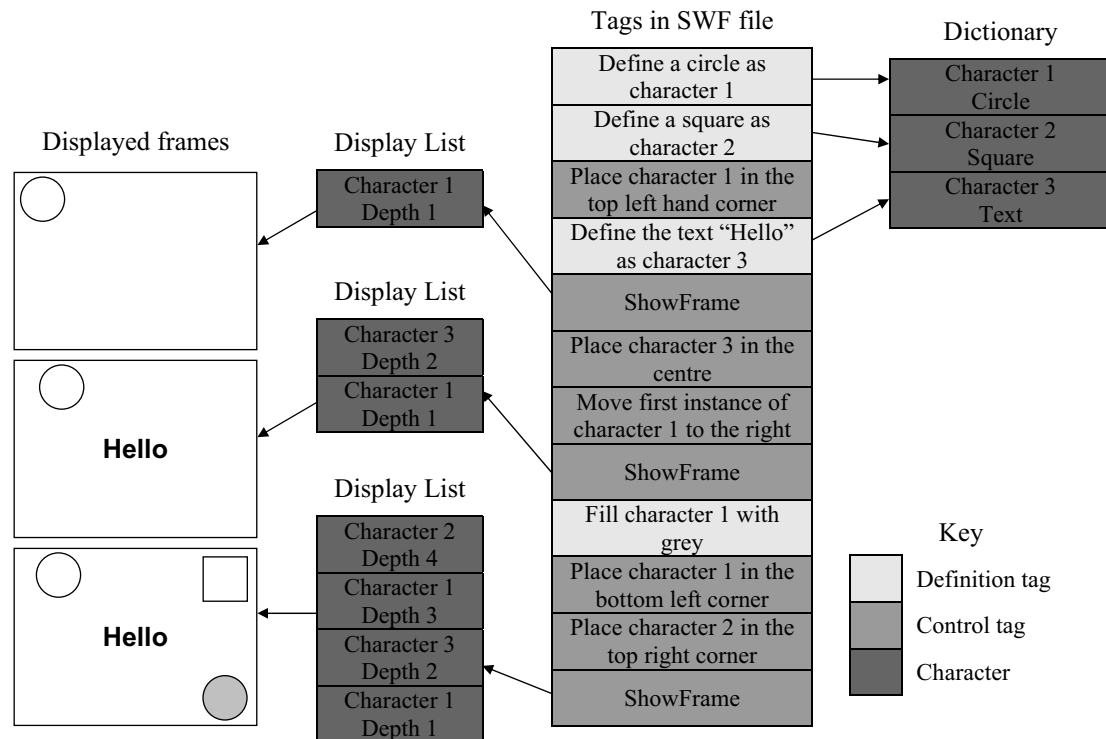


Fig 4.18

A simplified SWF file together the dictionary created and the resulting frames displayed.



GROUP TASK Discussion

Work through each tag in the above SWF file in sequence. Discuss how each tag is processed by the flash player to display the animation frames.



GROUP TASK Discussion

Flash is an example of a streaming media, this means flash files can begin playing whilst the file is still being downloaded. Discuss aspects of the organisation of flash that makes streaming possible.



GROUP TASK Activity

Use a software application that is able to create flash files. Whilst creating a simple flash animation, note any processes and functions that correspond to definition tags and those that correspond to control tags.

WORD PROCESSORS AND DESKTOP PUBLISHING FOR TEXT, IMAGES AND NUMBERS

The distinction between word processor and desktop publishing software is somewhat blurred, as a consequence much of the functionality found in a word processor is also present in a desktop publishing application and vice versa. The primary purpose of both these applications is to organise text, images and sometimes numbers, in preparation for display; usually printing. Word processors fundamentally organise text

data as a sequence of characters, these characters combine to form words and the words to form paragraphs. The output from word processors is generally printed on a personal printer, e.g. laser or inkjet printers. On the other hand desktop publishing applications organise text, together with other media elements, within frames, these frames can be positioned precisely on the page. The output from desktop publishing applications is commonly sent to a commercial printer.

Let us consider the organisation of text data and typical processing available within word processors and then within desktop publishing applications.

Word Processors

To structure our discussion we consider rich text format (RTF) files. The main purpose of the RTF format is to specify a method for organising text data so it can be transferred and used by different word processors running on a variety of different operating systems. RTF is a storage format, however in the context of the organising process, studying the arrangement of text data within an RTF file provides an insight into the organisation of text within word processors in general.

To view the raw data within an RTF file requires a text editor. Text editors do not attempt to make sense of files, rather they merely display each character within the file. Notepad is an example of a text editor that is provided with Microsoft's Windows operating system; similar text editors are available for other operating systems. *Fig 4.19* shows an RTF file as it appears when viewed in Microsoft Word and then the same file as it appears when viewed using Notepad.

The *cat* sat on the mat.

This was most pleasing for the cat.

```
{\rtf1\ansi\ansicpg1252\deff0\deflang3081
{\fonttbl
  {\f0\fswiss\fcharset0 Arial;}
  {\f1\froman\fcharset0 Times New Roman;}
}
\f0\fs32 The \i cat\i0 \b sat\b0 on the mat.\par
\f1\fs24\par
This was \ul most\ul0 pleasing for the cat.\par
}
```

Fig 4.19

*An RTF file viewed in a word processor
and the same file viewed in a text editor.*



GROUP TASK Activity

Enter the text shown in the lower frame of *Fig 4.19* into a text editor. Save the file with the name cat.rtf. Open this file using a word processor and confirm the result is the same as that shown in *Fig 4.19*.

The sample RTF file in *Fig 4.19* illustrates important aspects in regard to the organisation of text data within word processors. Let us analyse this file in more detail, the aim being to understand the organisation of the data and how this organisation assists processing. RTF files organise data into groups, which are enclosed within parentheses {}. Groups are composed of control words, which commence with a \, and the actual unformatted text data.

- The first line of the file informs the word processor that this is an RTF file and that it uses a particular character set and a particular language. ANSI code page 1252 means Western European and the defined language is Australia, represented by code 3081.
- Lines 2, 3 and 4 create a font table; font 0 being Arial and font 1 being Times New Roman. The other details within these lines provide information to the word processor to assist it to substitute a similar font should the specified font not be found.

So far no actual text data has been encountered, rather details of common elements used throughout the document have been defined; these elements form the header of an RTF file. As a consequence of these lines the word processor holds data on the language used and also details of each font to be used. It would be inefficient to specify a font each time it is used; rather each font is specified once within a font table. Each font within the font table can then be used any number of times within the document. Our example is a simple one; in reality word processors maintain a variety of structures similar to the font table. For example, colour tables, style tables, and paragraph format tables.

- Line 6 begins by specifying font 0, which within the font table is Arial. It then specifies the font size as 32; in RTF font sizes are specified in half points, hence `\fs32` sets the font size to 16 points. The actual text data follows; this data contains control words, namely `\i`, `\i0`, `\b`, and `\b0`, these control words predictably turn on and off italics and bold respectively. The line ends with the control word `\par`, which predictably signals the end of a paragraph.
- Line 7 produces an empty paragraph using a 12-point version of font 1. Line 8 specifies another paragraph of text that contains an underlined word.

These lines contain the text data together with specific details in regard to its formatting for display. Remember the main purpose of a word processor is to format text data in preparation for printing. Consider applying underlining to some text; to do this in most word processors you first select the required text and then initiate the underline function. This process adds a start underline control word prior to the selected text and an end underline control word to the end of the text.

The RTF `\par` control word is used to indicate the end of a paragraph. When this control word is encountered most word processors record various properties of the paragraph within a table. For example, in Microsoft Word a paragraph marker is really a pointer to data describing attributes of the paragraph (see Fig 4.20), hence copying a paragraph marker causes the destination to inherit the attributes of the source paragraph.

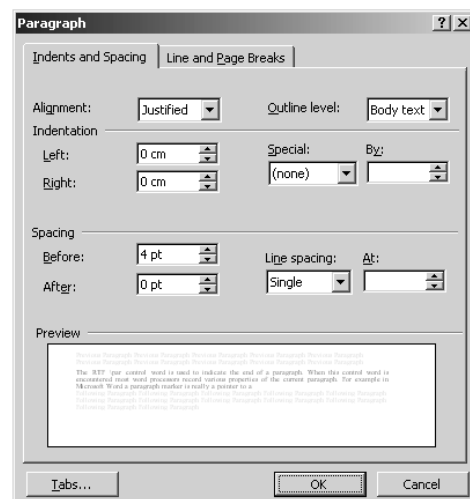


Fig 4.20
Paragraph attributes in MS-Word.

In summary, word processors organise raw text as a sequential list of characters. The formatting applied to blocks of text is arranged into tables of data that include font tables, colour tables and paragraph tables. Data in these tables can then be used multiple times to format different blocks of text. Some common formatting, such as bold, italics and underlining, is embedded directly within the text. The text itself, together with all the formatting data, is represented in binary using an extension of the ASCII system.



GROUP TASK Discussion

Copying and pasting are common processes within word processors. Discuss the changes taking place to the underlying data during such an operation. Will the contents of font and paragraph tables be altered?

Desktop Publishing Applications

The primary purpose of desktop publishing applications is to organise text and other elements in preparation for use by commercial printers. Furthermore, the design specified within the application must represent the final printed output accurately; this includes the precise colour, shape and placement of all elements. Most current desktop publishing applications also provide web page design functionality; however this is not their main purpose.

We need a basic understanding of commercial printing processes if we are to make sense of the organisation of data used by desktop publishing applications. There are two common printing processes used by commercial printers; four-colour process and spot colour. Both these techniques require that different colours be printed separately. Spot colour uses one or more different coloured inks. For example, if a book is printed using blue and red ink then the blue is printed in its entirety followed by the red. Four-colour process is used for full colour publications; it uses transparent cyan ink, magenta ink, yellow ink and black ink. Each page is printed using each of these colours in turn. These four colours, known as CMYK, when mixed are able to produce many thousands of different colours, but not the enormous range present in the RGB system. Mixing ink colours (CMYK) is a less precise process compared to mixing light (RGB).

The output from desktop publishing applications must be able to accurately separate out the required colours. *Fig 4.21* shows an example of a CMYK separation prepared for four-colour process printing. Each separation shows the portions of the publication that will be printed using each colour; these are ink colours rather than the RGB colours used to display the publication on the screen. Desktop publishing applications provide a mechanism for specifying the colour system, usually CMYK, spot colour or sometimes RGB.

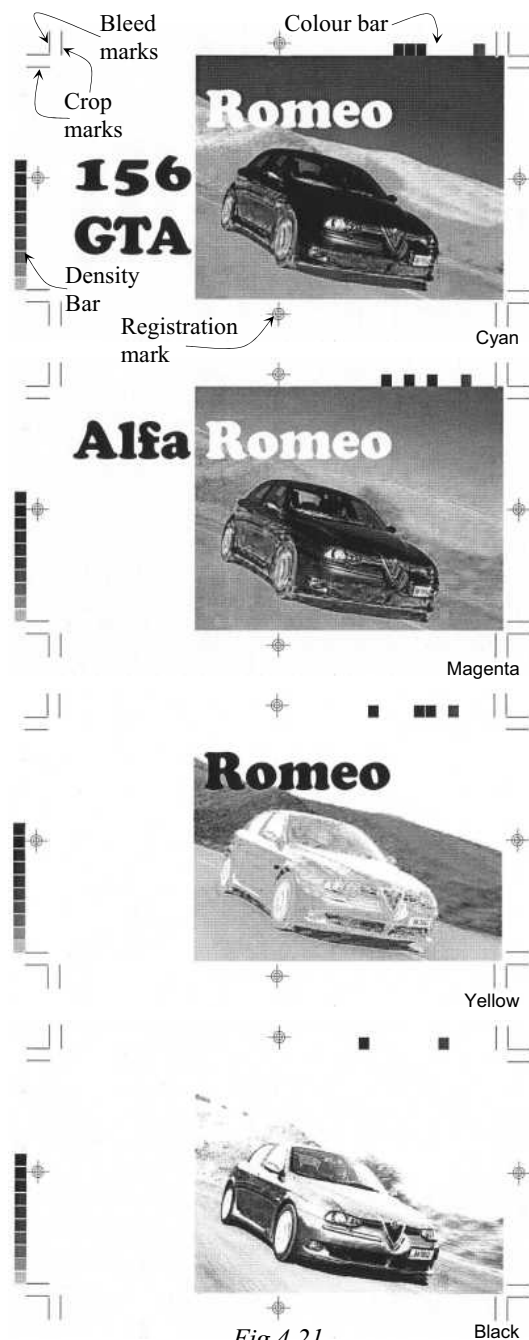


Fig 4.21
CMYK colour separations from a desktop publishing software application.



GROUP TASK Activity

Create a simple document, which includes various colours, using a desktop publishing application. Select a suitable method of colour representation. Print colour separations of your document.

As computer monitors display colour as intensities of light rather than printed colour a physical colour matching system is required; the Pantone® colour matching system being the standard (*Fig 4.22* shows a screen used to specify colours using this system). Printed colour swatches containing all the Pantone colours are used to show the precise printed colour. Once a colour has been selected on the swatch its number can be entered into the desktop publishing application.

In *Fig 4.21* various marks have been included to assist the printing process. Let us consider the purpose of such marks. As each colour is printed separately it is vital that each separation is printed in precisely the same position on the page; registration marks are used for this purpose. In most cases commercially printed publications are printed on paper larger than the final cut size; crop or trim marks are used to indicate precisely where the final product should be guillotined. Coloured areas that extend to the edge of the publication are printed slightly outside the area defined by the crop marks; this is called a bleed. Bleed marks specify the outside edge of all coloured areas within the page. The example separations in *Fig 4.21* include colour bars and density bars; these elements help the printer ensure the publication uses the precise colours intended.

Professional publications require precise control over the placement and spacing of individual characters and groups of characters within text data. Hence, within text frames, various properties of paragraphs, words or even individual characters must be precisely specified. Some examples are illustrated in *Fig 4.23*. Kerning is the process of altering the space between individual character pairs, adjusting Leading changes the vertical space between lines of text and altering Tracking adjusts the horizontal space between characters evenly within a line of text. Most word processors include such functionality, however desktop publishing applications provide far greater control over such settings.

Let us consider how desktop publishing applications organise data to reflect the commercial printing requirements discussed above:

- Frames are used to contain each publication element. The position of each frame is precisely specified relative to the final edges of the page. Each frame contains elements, such as text, images and various other graphic elements; the location of these elements within each frame is represented precisely.

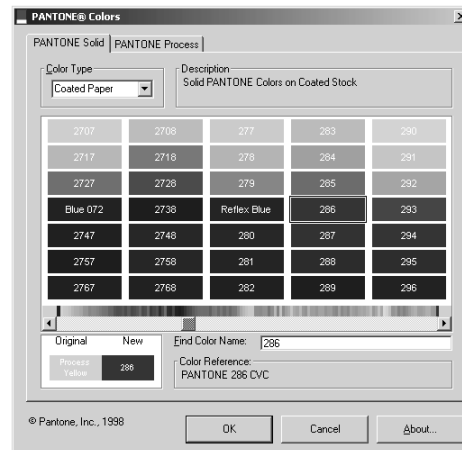


Fig 4.22

Pantone is a color matching system for both spot and four colour process.

Kerning

AW AW

Leading

The cat sat on the mat. This was most pleasing for him. The cat sat on the mat. This was most pleasing for him. The cat sat on the mat.

Tracking

The cat sat on the mat. This was most pleasing for him. The cat sat on the mat. This was most pleasing for him. The cat sat on the mat. This was most pleasing for him. The cat sat on the mat. This was most pleasing for him.

Fig 4.23

Examples of Kerning, Leading and Tracking.

- Colour is represented using the same system as that used for the final separations. This restricts the use of colours to those that can actually be printed. For example, if two spot colours are used, say Pantone® 279 and Pantone® 485, then the palette of available colours is restricted to different intensities of these two colours. The colour of each element within the publication is represented as a percentage of one or more of the available ink colours. Representing colour in this way means no colour conversion is needed when the separations are produced and consequently the final publication accurately reflects the colours within the original design.
- Desktop publishing applications must represent characters not just in terms of their sequence, font and size. They must also represent various properties in regard to the attributes of each individual character, together with attributes specified for particular groups of characters. For example, a pair of characters may have a particular amount of kerning applied.



GROUP TASK Activity

Examine the available features for controlling text spacing within a word processor and then within a desktop publishing application. Create a table comparing the features found in each application.

SPREADSHEETS FOR NUMERIC

The primary purpose of a spreadsheet is to store, analyse and process numeric data using mathematical techniques. Traditional paper-based spreadsheets were used to maintain the financial records necessary to operate a business. These paper spreadsheets required users to manually calculate totals and other statistics every time an entry was changed or a new entry was added; computer-based spreadsheets automate this process. They must be able to accomplish mathematical calculations quickly and results must accurately reflect the current data, that is, any changes to the data must be reflected in all results that use that data.

The above requirements mean that the organisation of data within spreadsheets is fundamentally different to the organisation of data within most other applications. Within both paper and computer-based spreadsheets, the input, data, processing and output are integrated within a single form or screen. Even a single cell within a spreadsheet is the basis for multiple information processes. For example, in a spreadsheet application a cell can be used to collect formula data, the data in this cell is then used to analyse other data and also display the result of the analysis. Most other applications separate collection, analysis and displaying into distinct processes.

Spreadsheets organise numeric and other data into an arrangement of columns and rows; columns are generally labelled alphabetically and rows labelled numerically. The intersection of a column and a row is called a cell. For example, the cell address B7 refers to the cell at the intersection of column B and row 7. Each cell contains a particular data item; the method used to represent each of these data items changes based on the type of data within the cell. In general, cells contain numeric, text or formula data, each being represented differently.

Consider the sample Microsoft Excel spreadsheet shown in *Fig 4.24* below. The cells in row 1 contain text data that is used as headings for the data contained in each column. In spreadsheet terminology, cells containing text data are known as labels. A label identifies and gives meaning to something, in most cases cells containing text do just that, they identify and give meaning to the numeric data. A range of cells is specified using the address of the top left hand cell, a colon, and the address of the

that multiplies 11 by the value in the cell above and then subtracts 2. Since C1 contains the value 0.2, one would expect each cell in the range C2:C26 to calculate $11 \times 0.2 - 2$ and display the result 0.2; in *Fig 4.25* this is clearly not the case.

The ability to copy formulas and have their cell references change to reflect their new location is a powerful feature present in all spreadsheet applications. In *Fig 4.25* the formula in cell C2, namely $=11 \times C1 - 2$, was copied, or filled down, into cells C3 to C26. As a consequence the formula data in cell C3 is $=11 \times C2 - 2$; the reference to C1 in the original has changed to C2. How has this happened? The reference to C1 in the original formula actually refers to the cell directly above cell C2, C1 is an example of a relative reference. When a formula is copied to a new location, all relative references will point to cells relative to the new cell's location. If this is not desirable then an absolute reference should be used in the original formula. Absolute references are specified in most spreadsheets using dollar signs. For example, in a formula the reference $\$C\$1$ always refers to cell C1, $\$C1$ always refers to column C but allows the row to change relative to the new position, and $C\$1$ always points to row 1 but allows the column to change relative to the new position.

Numeric data, including the results from formulas, can be formatted in various ways, for example as currency, dates, percentages, or fractions. Altering the format of the data does not alter the underlying raw data; rather it organises the data in preparation for display. This makes the data understandable for humans. *Fig 4.26* shows some of the various number formats included in Excel; it is also possible to define custom formats to suit specific requirements.

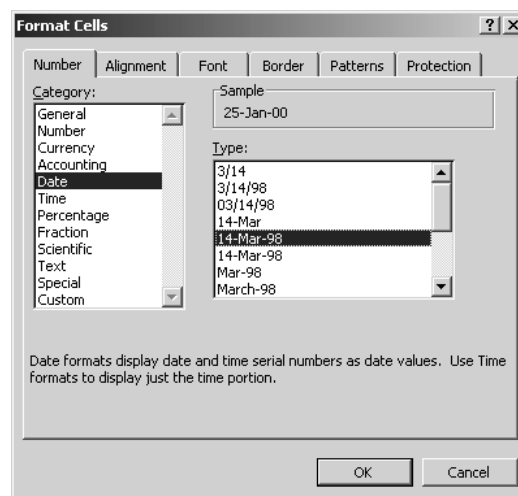


Fig 4.26
Numeric cell formats included in Excel.



GROUP TASK Activity

Reproduce the spreadsheet shown above in *Fig 4.25*. Observe the effect of changing the format of the cells to alter the number of decimal places displayed. Explain your observations in terms of data organisation.

In summary, spreadsheets organise data according to the following:

- Data is arranged in columns and rows; the intersection of a column and row determines a cell.
- Any cell can hold and represent text, numeric or formula data.
- Formulas refer to other cells using their cell address. Such references can be relative or absolute references.
- Numeric data, and the results from formulas, are represented using the double precision floating-point system.
- When formatting is applied to cells it has no effect on the actual data held within the cells.
- The organisation of data in a spreadsheet, in particular formulas within cells, allow collecting, analysing and displaying to be integrated processes.



Consider the following:

The Excel spreadsheet in *Fig 4.27* is used to determine the number of surnames commencing with each letter of the alphabet. For example, there are 11 surnames that start with the letter A. Currently the spreadsheet contains a total of 234 names sorted into alphabetical order.

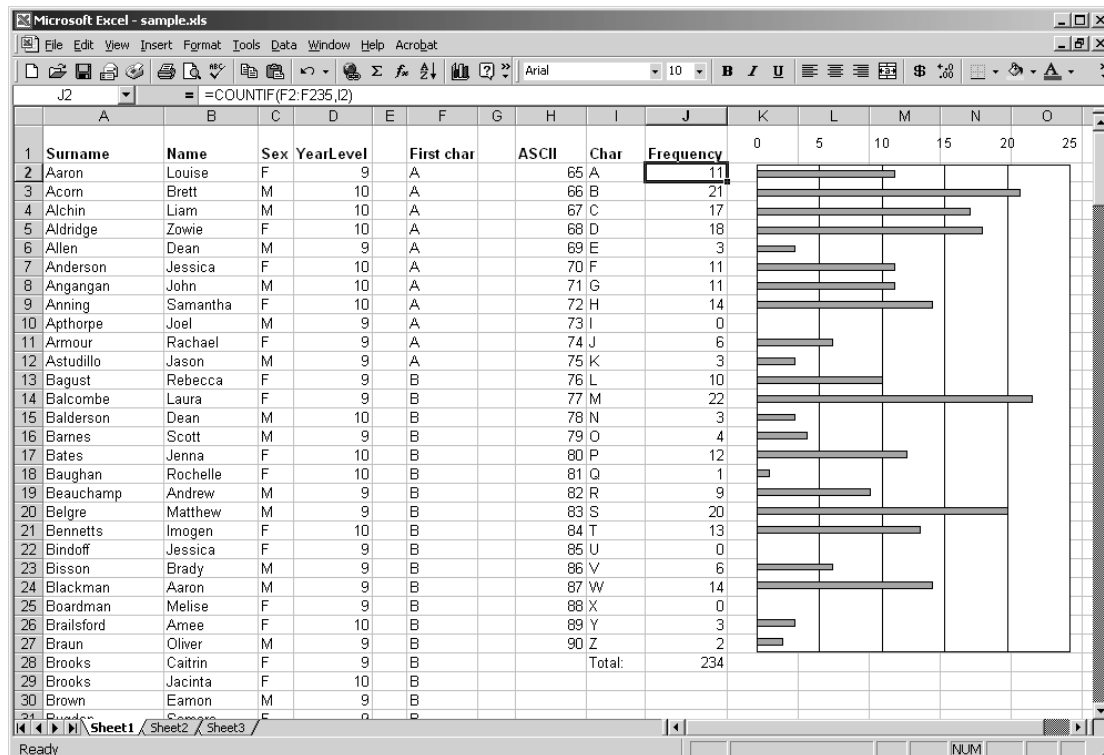


Fig 4.27

Spreadsheet analysing the frequency of surnames commencing with each letter of the alphabet.

The following formulas have been used within the spreadsheet:

- Cell F2 contains the formula =LEFT(A2,1)
- Cell H3 contains the formula =H2+1
- Cell I2 contains the formula =CHAR(H2)
- Cell J2 contains the formula =COUNTIF(F2:F235,J2)
- Cell J28 contains the formula =SUM(J2:J27)



GROUP TASK Discussion

Classify each cell on the spreadsheet as containing text, numeric or formula data.



GROUP TASK Discussion

Analyse the formulas used and determine which cell references could (or should) be absolute and which should be relative references.



GROUP TASK Activity

Reproduce the spreadsheet in *Fig 4.27* using your own data, and based on the above screen shot and information.

SET 4B

1. Which process alters the precise distance between pairs of characters?
(A) tracking
(B) kerning
(C) leading
(D) font size
2. An animation, for use on a web page, that has a small number of frames contained within a small screen area would most likely be stored as a(n):
(A) animated GIF file.
(B) MPEG file.
(C) flash file.
(D) quick time file.
3. The dictionary, maintained by a flash player, is used to:
(A) hold the meaning of each tag used within a flash file.
(B) store a sequential list of all the definition tags in the current flash file.
(C) create each frame of the animation prior to its display.
(D) maintain a description of each character that can be used in the animation.
4. Responding to user actions is possible in:
(A) animated GIFs.
(B) MPEG files.
(C) flash files.
(D) All of the above.
5. For a video file to support streaming:
(A) the data must all be received prior to the first frame being displayed.
(B) all the data for each complete frame must be received in the order the frames are to be displayed.
(C) the video data should be compressed.
(D) each frame needs to be stored as an ordered sequence of independent bitmaps.
6. Most video editing software:
(A) creates the final video data as editing takes place.
(B) organises the ordering and transitions between video clips.
(C) produce the final video data after editing has been completed.
(D) Both (B) and (C).
7. Font tables are used by word processors:
(A) to make understanding the organisation of the data difficult.
(B) to describe the precise shape of each character within the text.
(C) because they save specifying the detail of each font every time it is used.
(D) to specify the font used for each particular block of text.
8. Desktop publishing applications use the CMYK system for representing colour because:
(A) there are less colours available compared to the RGB system.
(B) CMYK represents the ink colours used during the four colour printing process.
(C) the resulting files are smaller in size.
(D) computer monitors reproduce CMYK colours more accurately.
9. Formatting a cell in a spreadsheet as a date will:
(A) alter the underlying data.
(B) not alter the underlying data.
(C) only change the way the data is displayed.
(D) Both (B) and (C).
10. If the formula =B1*\$D\$2 is copied from cell A2 to cell C3 then C3 would contain the formula:
(A) =B1*\$F\$3
(B) =D2*\$F\$3
(C) =D2*\$D\$2
(D) =B1*\$D\$2
11. Compare the organisation of data in a flash file with that in a video clip collected using a digital video camera.
12. List and describe the essential differences between the organisation of data in a word processor and a desktop publishing application.
13. "Spreadsheets integrate many information processes, including the organising process." Explain how the organisation of spreadsheet data facilitates this integration of processes.
14. Define each of the following desktop publishing terms: four-colour process, spot colour, colour separation, kerning, leading and tracking.
15. Create a spreadsheet that contains names and marks out of 100 for an assessment task. Develop a formula to convert each mark to a performance band. Band 6 for 90 and above, band 5 from 80-89, band 4 from 70-79, band 3 for 60-69, band 2 for 50-59 and band 1 for marks less than 50. Create a table to determine the number of students in each performance band.

DATABASE SOFTWARE THAT ORGANISES DATA INTO TABLES

Arranging data into tables is one of the most common methods of data organisation. Each row in a table represents all the data about a particular person or thing. Individual columns contain similar data about each person or thing. For example, a table containing personal contact data would likely include columns for surname, first name, street address, suburb, postcode and phone number (see Fig 4.28). Each row in the table contains all the data about a particular

Surname	FirstName	Street	Suburb	Postcode	Phone
Davis	Luke	16 Hilda St	Guy Lake	2123	(02) 4567 8901
Fendall	Janine	11 Foogle St	Foggleton	2222	(02) 1234 3456
Holden	Michelle	5 John Cres	Lane Cove	2066	(02) 3333 3333
Jackson	Mark	1 Henley St	Lane Cove	2066	(02) 9899 8799
Jamison	Geoffrey	22a Yellow Rd	Klementon	3456	(02) 1234 5678
Martin	Mary	66 Villiers Way	Foggleton	2345	(02) 2345 6789
Nerk	Fred	15 Timid St	Whatsville	1234	(02) 9635 5517
Willis	John	32 Harold Place	Joanstown	2355	(02) 2345 6789

Fig 4.28

Sample personal contact data arranged as a table.

person, hence the data contained in each row is related and should be represented in such a way that it remains together. Furthermore, the type of data held in each column is always the same; this makes it possible to search, sort or otherwise process the data based on particular columns. It makes sense to organise the data in such a way that the above properties of table data is preserved and even enforced.

Most software applications organise at least some of their data into tables. In many applications this table-based organisation is not obvious to the user. For example, earlier in this chapter we examined the RTF specifications; RTF uses tightly structured font tables, colour tables and paragraph tables. The most obvious applications that utilise tables of data are those based on databases. Most databases are composed entirely of tables of data and furthermore this table organisation is clearly apparent to the user. Hence in this section we concentrate on the organisation of data within simple databases and how this method of organisation assists processing. Much of what we discuss also relates to various other software applications where the organisation of the data into tables is less obvious.

The rows in a database table are known as records and each record is composed of fields. The records within a particular table can be considered to be of a particular data type; all records possess the same fields and each field is of the same type, hence each record is also of the same data type.

When creating a database table each field is created with a particular data type together with various other properties; this information is specified using a data dictionary (see Fig 4.29). The data type and all other specified properties of the field are enforced whenever data is entered or edited. Examples of data types commonly available include text, numeric types including integers and floating-point representations, Yes/No or Boolean, and even data types capable of storing images, audio and video data. Common field properties include the number of characters for text fields, default values for new records, default formatting specifications such as the number of decimal places to display, and even validation rules to restrict data entry.

Fig 4.29

Sample data dictionary created with Microsoft Access.

Let us now consider how organising data in tables assists the processing commonly performed on such data. We shall consider two common analysis processes, namely searching for a particular record and sorting records.

Certainly one of the most common processes performed on tables of data is locating a particular record based on specified search criteria. For example, when logging onto the Internet your ISP (Internet Service Provider) receives your username and password. Your username is used as the criteria to search the ISP's database for your record. Once the record is located the data in the password field is retrieved; subsequently this data is compared to the password you entered. The organisation of the ISPs database into tables greatly simplifies this process. All usernames are held in a particular field, so only this field needs to be examined. Once the username is found, all fields within this matching record, including your password, are also found. In general, when a particular data item is located within a field, it is the entire record (or row) that is retrieved not just the matching field data. This means that any of the data within the found record is readily available.

Another common analysis process performed on tables of data is ordering or sorting data in preparation for display. In a database table the order in which records are physically stored is not significant; conceptually records exist in no particular order. When a sorting or ordering process is initiated all the records must be arranged into this specified order. When sorts are applied to large tables with many thousands of records this is a potentially lengthy process. How can the data be organised to improve the efficiency of sorts? The answer is to use indexes. Indexes within database tables are similar to those in the back of a book. Think about the index in a book; it provides an alphabetical listing, where each entry points to a specific page. If, rather than the page numbers in the index, you inserted the actual information then the result would be the book sorted into alphabetical order based on the keywords within the index. Indexes within database tables operate in a similar way; they describe a particular record order without actually ordering the records. The index can then be used to quickly order the data when required. Note that indexes are also used to increase the speed of searches. Maintaining many indexes in a table decreases performance when adding and editing records, therefore a compromise must be reached; indexes should only be created for fields, or combinations of fields, that are often used as the criteria for sorts and searches.

In summary, data organised into a table is arranged in rows and columns. All the data in each column or field is of the same data type, hence the method used to represent each data item within a column is identical. The data in a single row or record holds all the data about a single person or thing, all records in a table have the same data type. Most table processing manipulates entire records. For example, a search or sort may display particular fields however in reality entire records are being processed.

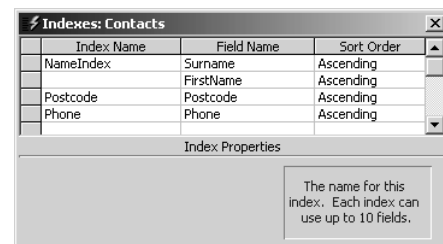


Fig 4.30
Defining indexes in Microsoft Access.



GROUP TASK Activity

Create a table of data using a database application and then copy and paste this data into a spreadsheet. Perform sorts on both sets of data. In terms of data organisation, explain why the methods of sorting differ.

WEBSITE CREATION SOFTWARE THAT USES HYPERLINKS TO ORGANISE DATA FOR WEB PAGES

There is an enormous range of website creation software applications available from basic text editors through to professional packages such as Adobe's Dreamweaver. There are specialised applications for creating surveys, news sites, discussion forums and various other interactive web sites. Whilst some web pages actually exist as files waiting to be retrieved from a web server, many others are created dynamically based on data collected from users. For example, search engines create a specific web page of hyperlinks based on the search criteria entered by each user. Most organisations now use a Content Management System (CMS) to create web pages dynamically. The CMS uses a database to store the hyperlinks, text, images and also details in regard to the page layout and formatting of web pages. When a user requests a particular web page the CMS queries the database and then creates the web page dynamically based on the query results. We cannot hope to examine the functionality of website creation software in any detail; rather we will concentrate on the nature of the hyperlinks created by all such software applications.



GROUP TASK Activity

Search the Internet to find a popular example of a Content Management System (CMS). Determine the Database Management System (DBMS) used and also the programming language used by the CMS as it generates web pages.

Consider all the web pages that can be created dynamically, together with the billions of actual pages stored as files on web servers. The total is infinite. All these pages form the World Wide Web (WWW) and all are linked together using hyperlinks. When a user clicks on a hyperlink they are taken to a related document; this new document may also contain hyperlinks to further documents. As a consequence documents are connected to each other in a highly complex and unstructured manner. Despite the unstructured organisation of hypertext, it better reflects the thought processes of the human mind than other methods of data organisation. The human mind operates largely on associations; we read a passage of text and our mind generates various related associations based on past experiences. Our thoughts move continually from one association to another; hypertext is an attempt to better reflect this behaviour.



GROUP TASK Discussion

Documents containing hyperlinks are said to “better reflect the thought processes of the human mind.” Do you agree with statement? Discuss.

Hypertext is a term used to describe bodies of text that are hyperlinked. The related term, hypermedia, is an extension of hypertext to include hyperlinks to a variety of different media types including image, sound, and video. In everyday usage, particularly in regard to the WWW, the word hypertext has taken on the same meaning as hypermedia; in our discussions we shall just use the term hypertext. Be aware that when we discuss hyperlinks to other documents, these other documents are not necessarily text; they could be any of the media types.

Documents created by website creation software and then accessed via the WWW are primarily based on HTML. HTML is an acronym for hypertext markup language and is the primary method of organising hypertext for use on the WWW. Clicking on a link within an HTML document can take you to a document stored on your local hard drive or to information stored on virtually any computer throughout the world. From

the user's point of view, an HTML document is just retrieved and displayed in their web browser; the physical location of the source document is irrelevant.

Let us consider the organisation of a typical HTML document. All HTML documents are text files, that is, a sequential list of characters where each character is represented using a coding system similar to the ASCII system. Hence, HTML files can be created and edited using text editors. *Fig 4.31* is an example of such a file viewed within a text editor. The large majority of website creation software applications automate the creation of HTML files; text editors are the most basic example.

```
<a href="http://www.pedc.com.au/"></a>
<p>This <a href="http://www.pedc.com.au/ipptext.htm">text</a> is
published by <a href="http://www.pedc.com.au/">Parramatta
Education Centre</a>.</p><p>If you wish to contact us by <a href=
"mailto:info@www.pedc.com.au">Email</a>, then please do so!</p>
```

Fig 4.31

Example HTML file viewed in a text editor.

HTML uses tags to specify formatting, hyperlinks and numerous other functions. All tags are enclosed within angled brackets `< >`; these brackets indicate to the web browser that the text enclosed is an instruction rather than text to be displayed. In most cases, pairs of tags are required; a start tag and an end tag. The function specified in the start tag is applied to the text contained between the tags. For example, in *Fig 4.31* `<p>` is the tag to start a new paragraph and `</p>` ends a paragraph; the paragraph contains the text between these two tags. A hyperlink to the Parramatta Education Centre web site is specified using:

```
<a href="http://www.pedc.com.au/">Parramatta Education Centre</a>
```

The start tag for a hyperlink commences with `<a`, followed by `href=`, then the URL to the required website within double quotes, and finally the end bracket `>`. Actually more than just the URL can be specified; you can specify a particular HTML document on the website, or even a particular position within an HTML document. Following the end of the start tag is the text to which the hyperlink is applied; in the above example the text is Parramatta Education Centre. The end tag `</a>` finalises the hyperlink. When viewed in a web browser, all text, and any other elements, contained between the start and end tags become the hyperlink. *Fig 4.32* shows the HTML file from *Fig 4.31* as it appears in a web browser. Notice that the text Parramatta Education Centre is underlined, and in reality is a different colour, to indicate its status as a hyperlink. Clicking on this text would take the user to the URL `http://www.pedc.com.au/`.

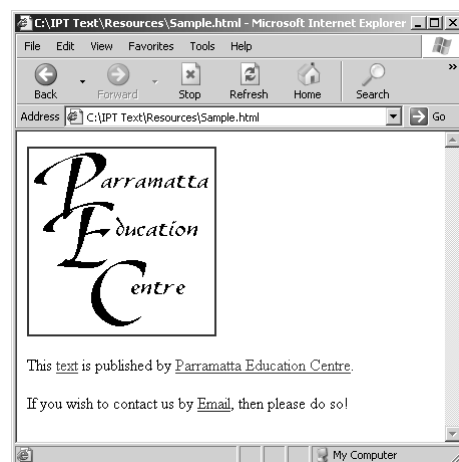


Fig 4.32

The HTML file shown in Fig 4.31 displayed in a web browser.



GROUP TASK Discussion

Compare the text data shown in *Fig 4.31* with the display in *Fig 4.32*. Based on your observations, list and describe the effect of each HTML tag within the original HTML file.

In general, HTML documents and also other documents that contain hyperlinks are organised as follows:

- All HTML documents are stored as text files. That is, they are arranged as a sequential list of characters where each character is represented using a system similar to the ASCII system.
- Pairs of tags are used to specify hyperlinks and other instructions. Pairs of tags can be nested inside each other.
- Tags are themselves strings of text, they have no meaning until they are analysed and acted upon by software.
- In HTML, tags are specified using angled brackets < >. Text contained within a pair of angled brackets is understood by web-enabled applications to be an instruction; all other text is displayed.
- Web browsers, and other web enabled software applications, understand the meaning of each HTML tag. Such applications are able to analyse tags and respond accordingly.



GROUP TASK Activity

Create a simple webpage using a website creation software application. Include three hyperlinks - one to another part of the document, one to an external document and one to an image. Examine the resulting HTML file and locate the hyperlink tags.

PRESENTATION SOFTWARE THAT ARRANGES DATA ON SLIDES

Presentation software is used to create slideshows that are commonly used to supplement a speaker's presentation. The slideshow is organised into a sequence of slides which are displayed as the speaker makes their presentation. A master slide is used to specify elements common to all slides, such as the presentation title, company logo, background colour and the location and default properties of text boxes and other elements. Each slide can contain a combination of media elements including sound, video, images and text. Different media elements on a particular slide are often displayed progressively by the speaker as the presentation progresses. A variety of different techniques can be used to animate the transition from one slide to the next and also to animate individual elements on each slide. For example, the next slide might slide in from the left to cover the previous slide or an image may expand out from a smaller thumbnail.

Prior to the widespread use of presentation software a speaker may well have used an overhead projector, whiteboard, analog video system and a separate sound system. Presentation software combines the functionality of all these systems. For example, *Fig 4.33* shows an overview of the slides within a Microsoft Power Point presentation. This particular presentation includes sound, video, image and text. The text itself is animated so that during the presentation each dot point appears as the presenter clicks their mouse. Using an overhead projector a similar effect was achieved manually as the presenter progressively uncovers each dot point.

Many presentations are designed to be delivered live where the speaker controls the timing of the slides as they speak, for example; teachers, lecturers and sales presentations. In this case a data projector is used to display the slides to the audience and the presenter commonly uses a wireless controller to move through the slides. Other presentations are available at any time. In this case the slides can be accompanied by a narrated sound track (which is often recorded during an initial live presentation of the material). As the narration is recorded the software maintains a

record of the precise moments when the narrator moves to the next slide or the next element within a slide. The timing of all slide transitions is stored within the presentation file.

Most presentation software includes a number of options for distributing and

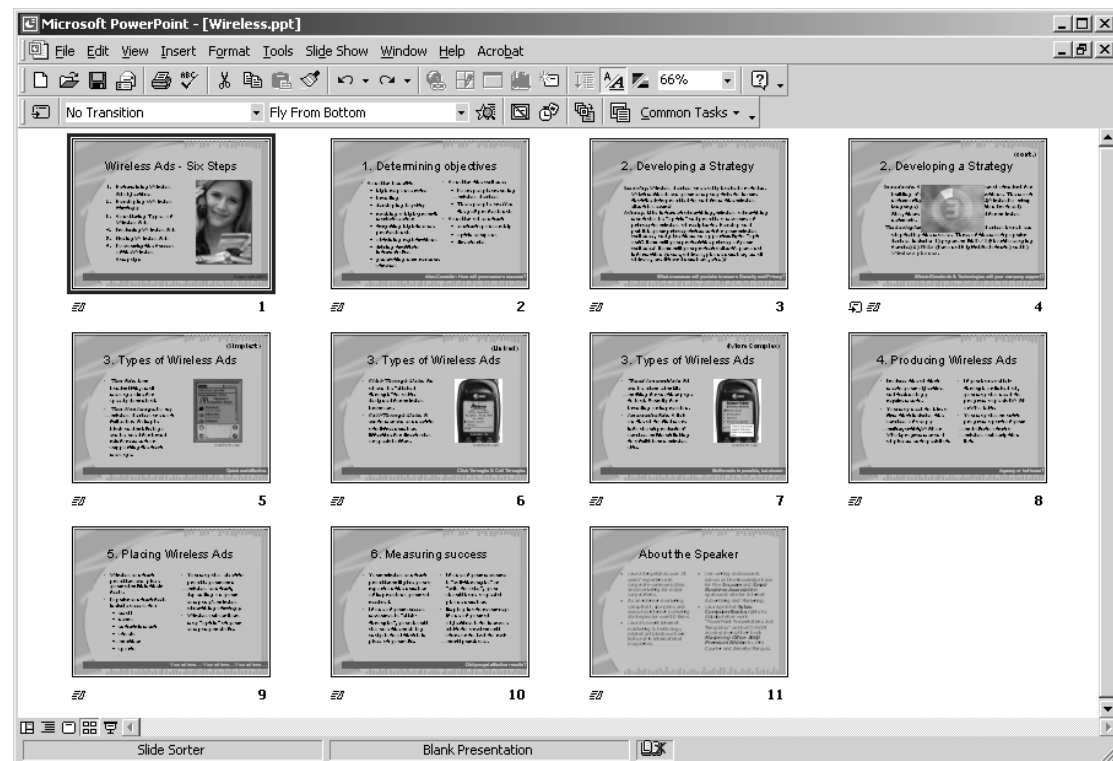


Fig 4.33

The slide sorter screen within Microsoft's PowerPoint presentation software.

displaying presentations. Common options include:

- The main presentation is contained within a single file which contains links to any larger media elements, commonly video and sound files. The slide show is viewed within the original presentation software application often using the same computer used to create the presentation. This is the usual method for live speaker presentations using a data projector.
- The slides, media elements and narration can be contained within a single file for viewing within the original presentation software application or using a simplified viewer application. In this case any media elements used on slides should be embedded wherever possible so they are stored within the single presentation file. Often larger media elements, such as video files, will need to be copied along with the main file – most presentation software includes functions to automate the copying process so links to these files are maintained. This is the usual method when distributing the presentation on CD-ROM, via email or download.
- The presentation can be exported as HTML files together with files for each of the media elements. In this case, the files are uploaded to a web server and can then be accessed and viewed using a browser.



GROUP TASK Discussion

Often audiences complain about the overuse of 'wizz bang' sound and animation within presentations. All this added fluff detracts from the actual message being delivered. Do you agree? Discuss.

NON-COMPUTER TOOLS FOR ORGANISING

For thousands of years people have been organising their data on paper or its equivalent. In the 1930's, as radio and later television emerged, the idea of the paperless office came in to being. It was thought that very soon the need for paper would disappear. The paperless office has never been realised; in fact the opposite is actually the case. Today we use more paper than ever before in history. When writing a simple letter using a word processor it is likely that multiple drafts will be printed; commonly more than half a dozen sheets of paper being used to complete the process. Prior to computerised word processors, a hand written draft would be created and edited on a single sheet of paper. The final letter was then hand written or typed on a typewriter; in total just two sheets of paper were used. Why for many tasks do we still prefer and use paper-based hardcopy and manual pen and paper methods over their electronic equivalents? To assist in answering this question let us examine examples of hard copy systems and also pen and paper systems used to organise data. Many of these non-computer tools utilise computer technology to assist the collection and/or initial organisation of the data, however once created it is the organised hardcopy that is used to communicate the information to people.



Hard copy

A copy of text or image based information produced on paper.



Consider the following:

- Telephone books use enormous amounts of paper, yet virtually every household and business throughout the world receives a new telephone book, or set of telephone books each year. In Australia two sets of telephone books are distributed; the white pages, which is arranged alphabetically by surname, and the yellow pages, which is arranged into business categories and then alphabetically within each category.
- Card catalogues are used as indexes for larger collections. For example, most graphic designers store a physical proof of each design they create. This collection of proofs is indexed using a card catalogue arranged in customer name order. Commonly the actual proofs are arranged in chronological order. Each customer card contains a reference to where each customer's proofs are stored. Similar card systems were until recently used in libraries; the books being physically arranged on the shelves by their call numbers, with at least two separate card catalogues being maintained. One catalogue being sorted by title and the other by author; when a new book was added to the collection a new card was added to each card catalogue. *Fig 4.34* shows a photograph of such a card catalogue used prior to 1989 by the Oakland Public Library in California.



Fig 4.34

Card catalogue used prior to 1989 in Oakland Public Library, California.

- Filing cabinets, folders and paper documents are still used extensively in virtually all offices. It is likely that your school maintains a filing cabinet for each year level. Within this filing cabinet is a folder for each student and within each folder are copies of each student's original enrolment form, schools reports and various other documents. Some of the documents within each folder are likely to be hardcopy generated by a computer system and others being handwritten. Most businesses and government departments maintain similar filing cabinet based databases. For example, all law firms maintain extensive paper files containing original court documents.
- Many processes are still performed using manual pen and paper techniques. Pen and paper is readily available to all and furthermore the result is more personal. Consider the following common examples: Phone messages are commonly distributed on slips of paper. At school, students handwrite their notes and teachers maintain handwritten mark books. Much of the initial planning of even computer-based information systems is done using pen and paper. The minutes of business meetings are recorded using pen and paper. Many people use hand written diaries and organisers to plan and record their work and social activities. We handwrite Christmas and birthday cards to family and friends.



GROUP TASK Discussion

Describe the organisation of the data used in each of the above scenarios and examples. In each case, how is the data structured and represented?



GROUP TASK Discussion

Is the idea of a 'paperless world' or even a 'paperless office' achievable or even desirable? Discuss using examples from the above dot points.



GROUP TASK Discussion

Computer-based information systems have certainly improved the efficiency and flexibility of data organisation. Does this efficiency and flexibility result in better information for all? Discuss.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH ORGANISING

Social issues either enhance or hinder people's interactions with each other. For example, the Internet greatly enhances people's ability to interact with others across the globe, however for those without the necessary skills and technology the consequences are quite the opposite; they become more alienated from the wider world. Ethical issues affect behaviour. Such issues alter the way we conduct ourselves, they may change our sense of morality; what is right and wrong. In this section we concentrate on social and ethical issues arising from the organising information process. Such issues commonly emerge as a result of current trends in organising data and also as a result of poorly organised data. In this section we consider each of these two areas, together with examples of social and ethical issues arising as a consequence.

CURRENT TRENDS IN ORGANISING DATA

- **The increase in hypermedia as a result of the World Wide Web.**

Currently the World Wide Web via the Internet has initiated an explosion in the use of hypertext and hypermedia; this method of organisation has revolutionised the way we communicate using all types of media. For example, hypermedia allows us to jump from reading text to viewing a video, we could then be led to a sequence of images that lead us to more text; each jump occurs in a virtual instant, furthermore on the WWW the source of the data can change with little regard for geographical boundaries. The data is organised in such a way that the end-user need not concern themselves with the physical complexities of obtaining the information; from their perspective they merely make the request and the information is displayed. The organisation of hypermedia not only simplifies the browsing process for users but also simplifies the actual creation of hypermedia documents.



Consider the following:

Traditional methods of publishing require significant effort and expense, hypertext and the World Wide Web have changed this situation. It is now possible for anybody to publish his or her ideas and thoughts on the WWW with minimal effort and at virtually no cost. Information on the WWW is then available to literally billions of people across the globe.



GROUP TASK Discussion

If anybody can publish information on the WWW then how can users determine the truth and reliability of such information? Discuss.



GROUP TASK Discussion

Does the increase in the use of hypertext and hypermedia via the WWW improve the social well being of all individuals? Discuss.

- **The ability of software to access different types of data.**

Software applications no longer operate on a single media type; they are able to access and also process data of various types. For example, text, images, sounds and even video clips can be combined using word processors and presentation software or stored within databases. The organisation of the data within such applications allows a variety of different media in a variety of different formats to be integrated and processed together.

The organisation of different media types within a single software application allows all data about an individual entity to be available. For example, insurance records are no longer limited to text; they can also include photographs of jewellery and other expensive items. Furthermore, the photographs and text are readily available from within a single software application. Individuals can use their home computer to produce multimedia presentations. For example, a student's presentation in class can easily include text formatted using a word processor, a chart produced in a spreadsheet, and even a video clip downloaded from the web. The presentation software used is able to access and display media of all types.



Consider the following:

For many years Sydney Water has maintained a database of all its customers, together with their water usage bills and payments. Sydney Water has now digitised the plans describing the location of water and sewerage services to each property in the Sydney region. These images are linked to individual customer records as well as to each adjoining property record.



GROUP TASK Discussion

The new system above has changed the nature of work for many of those working at Sydney Water. Discuss the likely nature of these changes.

- **Consequences of advances in display technology**

Advances in display technologies allow media of various types to be organised and displayed more efficiently and at a higher quality than was ever thought possible. Text, numeric, image and video data is routinely combined on a single screen; this information is synchronised with sound displayed via speakers. The sound cards in most home computers are able to reproduce music of similar or higher quality than many dedicated sound systems. The computer has taken on a new role; it has become a home entertainment centre; we play games, surf the Internet, listen to music, and even communicate using the multimedia display capabilities of our home computers. The ability to organise digital data in such a way that it can be used by new high quality display technologies has largely created this new role for computer technology.

It is not just the home that has been transformed by advances in display technologies. Many industries have also been completely revolutionised. Earlier in this chapter we considered desktop publishing software; this software together with new display technologies has totally changed the work performed in the publishing and printing industries. The music and film industries have also experienced radical change; music and film is not only produced and edited using digital techniques but this is also how the data is organised for distribution and display.



Consider the following:

Currently home theatre systems are selling at an ever-increasing rate. Such systems combine six-speaker digital surround sound with large plasma and LCD TVs. Our homes have become movie theatres; in many cases a whole room being dedicated to this new technology. Furthermore, many homes connect their computers into their home theatre systems; essentially the home theatre system is being used as a display device for the computer.



GROUP TASK Discussion

It is all well and good to have such amazing technology in our homes, however the downside is a decrease in social interactions. Technology just doesn't have a social conscience. Surely all this new technology cannot be a good thing. Do you agree? Discuss.

THE COST OF POORLY ORGANISED DATA

There are often future implications that can arise as a consequence of the method of data organisation. Businesses may grow or their focus may change; decisions in regard to the organisation of data should be made with future needs in mind. For example, when creating a customer database it would be wise to include fields for fax and mobile numbers even if current requirements do not indicate a need for such data.

Selecting a suitable method of data organisation can greatly improve the efficiency of further processes. For example, spreadsheets are not designed for organising large tables of data; each cell is independent of other cells. Storing large tables of information in a spreadsheet makes selecting specific data and validating data cumbersome; dedicated database management systems (DBMSs) arrange data into records and fields to simplify such processes. Furthermore, each field should be represented using the most appropriate data type; databases provide this facility whereas spreadsheets do not allow such detail to be specified.

Redundant data is duplicate data, that is, the same data exists multiple times. The duplicate data may be within the same table in a database or it may be in different software applications used by the same organisation. For example, the sales department for a company maintains a list of contacts whilst the ordering department maintains a separate list of customers. Both lists are likely to contain many records for the same people, hence it makes sense for this data to be organised into a central database that can be accessed and used by both departments. If an address or phone number is altered then the altered data will be reflected throughout the organisation.

Redundant data can cause problems for both participants and end-users. Data entry personnel will have difficulties deciding which record to edit or may have to edit multiple records; assuming that they are even able to ascertain that a duplicate record exists. End users will become frustrated as orders are sent to old addresses and duplicate mail-outs are received.



Consider the following:

The salesmen at a large car dealership enter the details of new and potential customers into the company's database. After each promotional mail-out the company was receiving an ever-increasing number of complaints due to customers receiving multiple mail-outs or customers receiving mail-outs when they had specifically indicated that they did not wish to receive such junk mail. The situation had steadily become more significant as the dealership expanded.

In an attempt to solve the problem the dealership removed all duplicate records where the address fields were identical. Unfortunately this caused further problems, customers complained that they were not receiving notices in regard to servicing their vehicles and others began receiving bills for work done on their flat mate or family member's vehicles.



GROUP TASK Discussion

Describe the likely organisation of the dealership's data that resulted in the above problems. Suggest a more suitable method of organisation.



HSC style question:

Explain how each of the following methods of organisation affects later information processes.

- Letters of the alphabet represented as bitmap images rather than text.
- Numbers represented as text rather than numeric.
- Video represented as a sequence of different image files rather than as one video file.

Suggested Solutions

- If the letters are represented as text, they are stored as their equivalent ASCII values. This allows individual characters to be identified and processed. If they are stored as bitmap images then it is difficult for the computer to determine which character is represented by which bitmap image. This means none of the text processing features are possible. For instance, ASCII codes are arranged in alphabetical order but there is no order to the bitmaps so sorting is difficult.

In addition, an ASCII code requires a single byte of storage whilst a bitmap requires every pixel to be stored, therefore the total storage for the bitmaps will be enormous compared to its corresponding ASCII codes. Any processes will therefore take much longer when working with the bitmap images.

- If the numbers are represented as text, the digits can be sorted, but only in alphabetic (not numeric) order. Thus, 1, 110, 130, 11580 will come before 2, 240, 260, 25678. The text digits can be formatted, but not in a mathematical sense. For instance, they cannot be formatted to a number of significant figures or number of decimal places. More significantly the text version of numbers cannot be used to perform mathematical calculations as they have no intrinsic mathematical value.

If they are stored as numeric values, they can be formatted mathematically and sorted into numeric order (1, 2, 110, 130, 240, 260, 11580, 25678) and can be used in mathematical calculations.

- Representing a video clip as a series of separate image files (as in an animated GIF sequence) means that while it is possible to display these in fast succession for a small number of images, it becomes increasingly slower to present a smooth animation in the same way for large numbers of image files.

If all of the separate image frames are combined and stored as one video file then the size of the file can be reduced enormously as only the changes from one image to the next need to be retained rather than each entire image. Also, it is possible to edit the file using video editing software adding sound effects and background music to specific places in the video clip.



GROUP TASK Discussion

Describe situations where each of the above inappropriate methods of organisation actually occurs.



GROUP TASK Discussion

For each of the above, suggest suitable software that could be used to reorganise the data into its more appropriate representation.

SET 4C

1. Which of the following is true of all records in a database table?
 - (A) Each record contains exactly the same number of characters.
 - (B) All records contain the same fields.
 - (C) Every field must contain data.
 - (D) They are stored in record number order.
2. Indexes are created for database tables to increase the speed of:
 - (A) searches and sorts.
 - (B) editing data.
 - (C) data entry.
 - (D) deleting data.
3. When viewing HTML in a web browser, HTML tags:
 - (A) are displayed exactly as stored.
 - (B) are always used to specify hyperlinks.
 - (C) are instructions that are not displayed.
 - (D) are ignored.
4. The best description of redundant data is:
 - (A) data that is accessed and used by many different users.
 - (B) data that is incorrect or is out of date.
 - (C) data that exists multiple times.
 - (D) copies of data that are maintained for security reasons.
5. Which of the following software applications organises data into a sequence of slides?
 - (A) website creation software.
 - (B) desktop publishing software.
 - (C) video processing software.
 - (D) presentation software.
6. Increases in the use of hypermedia are largely a result of which of the following?
 - (A) The different types of software now available.
 - (B) Advances in display technology.
 - (C) The increase in use of the World Wide Web.
 - (D) Different methods available for organising data.
7. A data dictionary is used:
 - (A) to describe the contents of each record within a database table.
 - (B) to specify the size of each data item held in a database table.
 - (C) as the basis for selecting and sorting data within a database table.
 - (D) to specify the data type, and various other attributes, of each field in a database table.
8. A JPEG file is opened within a word processor. The file appears on screen as gibberish. The best explanation for this is:
 - (A) The word processor has converted the file into a format it understands.
 - (B) The JPEG file was corrupted prior to it being opened.
 - (C) The word processor has opened the file as a text file.
 - (D) The JPEG filename extension is incorrect; it is really a text file.
9. Card catalogues are, or were, used:
 - (A) to store non-computer data such as books, video cassettes and audio tapes.
 - (B) to hold documents that are difficult to digitise.
 - (C) to sort larger collections of data differently to their physical order.
 - (D) to reorganise data into different formats.
10. What is the result of the following HTML code when viewed in a web browser?


```
<a href="http://www.google.com">Search</a>
```

 - (A) The browser would open the website www.google.com
 - (B) <http://www.google.com> would be displayed. Clicking on this hyperlink would open the Google website.
 - (C) A text box with a 'Search' label attached would be displayed.
 - (D) Search would be displayed, clicking on this hyperlink would open the Google website.
11. Describe the organisation of data within a database table.
12. Describe the organisation of data within an HTML document.
13. Describe TWO example scenarios where poorly organised data affects future information processes.
14. List and describe reasons why most offices still maintain paper-based filing systems.
15. Digital cameras and high quality inkjet printers have recently revolutionised the photographic industry. Research and discuss reasons why the digital organisation of image data has gained such widespread acceptance.

CHAPTER 4 REVIEW

1. Which of the following is NOT true of the organising process?
 - (A) It structures and represents data.
 - (B) It prepares data for use by other information processes.
 - (C) It alters the data in preparation for processing.
 - (D) It improves the efficiency of other information processes.
2. All data displayed on a monitor:
 - (A) must first be reorganised into one or more vector images.
 - (B) must first be reorganised into one or more bitmap images.
 - (C) must be reorganised into an analog signal.
 - (D) must pass through an ADC.
3. What is the best method of colour representation for full colour images to be included in commercial print publications?
 - (A) RGB as this best reflects the different light colours used to produce the final image.
 - (B) RGB as this best reflects the different ink colours used to produce the final image.
 - (C) CMYK as this best reflects the different light colours used to produce the final image.
 - (D) CMYK as this best reflects the different ink colours used to produce the final image.
4. Altering the size of each sample in an audio file from 16 to 8 bits would:
 - (A) reduce the volume of the sound.
 - (B) alter the frequencies within the sound.
 - (C) halve the number of sound samples.
 - (D) reduce the quality of the sound.
5. The aim of hypertext is to:
 - (A) randomly move from one idea to another.
 - (B) increase the efficiency of searches.
 - (C) better reflect the associations made by the human brain.
 - (D) structure the arrangement of data.
6. Resizing an uncompressed bitmap from a resolution of 640 by 480 to a resolution of 320 by 240 would:
 - (A) result in $\frac{3}{4}$ of the image being cropped.
 - (B) approximately halve the size of the file.
 - (C) reduce the file size by a factor of four.
 - (D) make each pixel one quarter of its original size.
7. The timeline within a video processing software application is used to:
 - (A) specify the sequence and timing of video and audio clips.
 - (B) generate the final compressed video file.
 - (C) ensure the final video file will be of the required size and resolution.
 - (D) detail the nature of the transitions between clips.
8. The best description of the organisation of data within a word processor is:
 - (A) A sequential list of characters, together with tables specifying various formatting options applied to the text.
 - (B) A two dimensional table of ASCII values.
 - (C) A sequence of paragraphs, where each paragraph contains words and each word contains characters.
 - (D) A series of objects, where each object is represented as a collection of attributes describing the object.
9. In a spreadsheet each row:
 - (A) contains data of the same type.
 - (B) contains cells which are independent of each other.
 - (C) can contain labels, formulas, or values.
 - (D) defines a record.
10. Cell A1 contains the formula $=B\$2+\$C\$3 + D4$. What would the formula become when copied into cell B4?
 - (A) $=B\$5+\$D\$7+D4$
 - (B) $=C\$5+\$D\$7+E7$
 - (C) $=C\$2+\$C\$3+E7$
 - (D) $=B\$2+\$C\$3+D7$
11. Determine the most appropriate type of software application for each of the following. In each case, justify your choice in terms of data organisation.
 - (a) Creating a company logo.
 - (b) Designing a full-colour advertising leaflet.
 - (c) Preparing a budget.
 - (d) Preparing photographs for use on a website.
12. Compare and contrast alternative methods for organising image data.
13. Compare and contrast the organisation of data in a spreadsheet with that in a database table.
14. List and describe the advantages of paper-based methods of organisation compared to computer-based methods.
15. "The method used to organise data has a profound effect on the efficiency of other information processes." Do you agree? Justify your answer using examples.

In this chapter you will learn to:

- identify hardware requirements to carry out a particular type of analysis
- describe the best organisation for data for a particular type of analysis
- use software analysis features in a range of software applications to analyse image, audio, video, text and numeric data
- compare and contrast computer and non-computer tools for analysis on the basis of speed, volume of data that can be analysed, and cost
- analyse data on individuals for the purpose it was collected

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Analysing – the process by which data can be represented and summarized so humans can better understand it

Hardware requirements for analysing, including:

- large amounts of primary and secondary storage allowing for fast processing
- fast processes allowing many rapid calculations

Software features for analysis, including:

- searching/selecting data
- sorting
- modelling/simulations
- what-if scenarios
- charts and graphs to identify trends
- file comparison

Non-computer tools, for analysing, including:

- searching manual filing systems
- non-computer models and simulations with these

Social and ethical issues associated with analysis, including:

- unauthorised analysis of data
- data incorrectly analysed
- erosion of privacy from linking databases for analysis

TOOLS FOR INFORMATION PROCESSES: ANALYSING

Analysing is the information process that transforms data into information; it makes sense of the data, changing it into a form or representation that can be understood by humans and used to obtain knowledge. Transforming data into information is a central aim of all information systems. They all collect data, transform it somehow into information, and finally this information is used to increase the knowledge of the system's users. Hence the analysing information process is central to achieving the purpose of all information systems.

The analysing process does not alter the data; rather it makes use of the data to create information. Data is the raw material of analysis; however unlike production processes, analysis does not change or alter its raw materials. Rather the analysing process uses various techniques to examine and summarise, but not change, the raw data; information is created as a result of applying these techniques to the data. Common analysis techniques include searching, selecting, sorting, charting and comparing data. The purpose of such analysis is often to summarise the data, make predictions, identify trends or to simulate some real life situation. The information resulting from the analysis process will only be valid if the data is both valid and complete and the method of analysis is also valid.

Each of the other information processes exists primarily to support the analysing process. Collecting gathers data for subsequent analysis. Storing and retrieving allows data to be maintained for later analysis. Transmitting and receiving provides a method for different systems to share and analyse each other's data. Processing alters data in preparation for analysis. Organising prepares the data for each information process; in terms of the analysing process it organises the data prior to analysis and then after analysis the resulting information is organised in preparation for display. Finally the information is actually displayed.

The ability of computers to examine vast quantities of data with incredible speed and accuracy make them excellent tools for analysis; manually performing such processes is tedious and prone to errors. In this chapter we examine hardware requirements necessary for various types of analysis. Software controls hardware, hence we examine various software features of particular use during the analysing process. We then consider non-computer analysis tools and finally examine social and ethical issues associated with the analysing process.

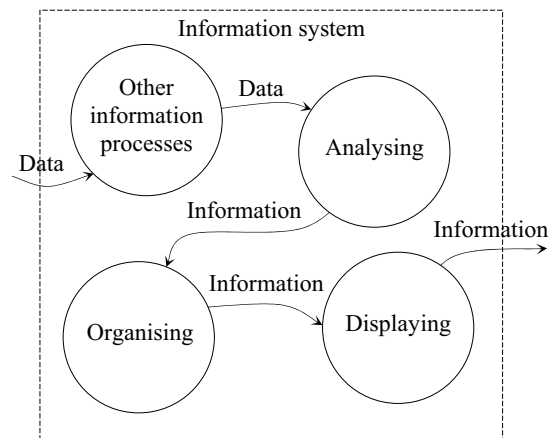


Fig 5.1

Analysing transforms data into information.

HARDWARE REQUIREMENTS FOR ANALYSING

Hardware requirements for analysing data are determined by the quantity of data and the type of analysis being performed on this data. For example, the hardware required for a teacher to calculate the average of a set of test marks is clearly quite different to that required for many hundreds of bank customers to access their account balances simultaneously. Obviously large quantities of data, such as that held by a bank, require larger amounts of secondary storage compared to the rather meagre requirements for storing the test marks for a class of students. More critically, this data needs to be available for analysis by the CPU in sufficient quantity, meaning it must be retrieved from storage at sufficient speed. The efficiency of the analysing process is a function of the speed of data access and processing, together with the quantity of data that can be accessed and processed as a single unit.

Let us consider three essential hardware components influencing the efficiency of the analysing process: secondary storage, primary storage and the CPU. We discuss the characteristics and operation of secondary storage devices in Chapter 6, and then in Chapter 7 we discuss the operation of the CPU; primary storage being discussed in both these chapters. In this chapter we are concerned with how these components affect the performance of the analysing process.

Secondary storage is permanent storage, such as hard disks; it does not require power to maintain data. Once data has been retrieved from secondary storage it is held in primary storage, known as RAM, prior to actual processing. RAM is an acronym for random access memory; it is volatile or non-permanent memory. RAM requires power to maintain its contents, however it operates at far greater speeds than secondary storage. The CPU analyses the data it retrieves from RAM, the result being information. Therefore if the analysing process is to progress efficiently then secondary storage must provide data to RAM at sufficient speed and in sufficient quantity to meet the demands of the CPU. Modern CPUs can execute millions of instructions per second, however for this to occur the data needs to be retrieved from secondary storage into RAM at the maximum speed possible. In essence, secondary storage is the weakest link in the chain, followed by RAM and then finally the CPU.

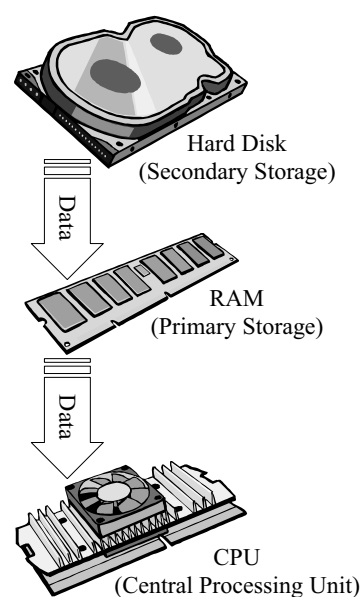


Fig 5.2
The CPU analyses data retrieved from RAM, which in turn is retrieved from secondary storage.



GROUP TASK Activity

Most computers include a light to indicate when the hard disk is operating. Observe this light whilst carrying out various different processes. Suggest reasons for your observations.



GROUP TASK Discussion

Hard disk, RAM and CPU performance are certainly vital considerations, however there are various other hardware components whose performance should be considered. Brainstorm a list of such components and discuss their role.

Different models and design of hard disk, RAM and CPU have different strengths and weaknesses; various measures are often quoted to assist in selecting the most suitable components. Let us consider common measures used to determine the performance of secondary storage (in particular hard disks), primary storage (in particular RAM) and the CPU.

HARD DISKS

In terms of the analysing process the best measure of secondary storage performance would be one that records the time taken for the actual data required by a particular analysing process to be read and successfully placed into RAM. Unfortunately trying out different models of hard disk for a particular scenario is generally impractical. Also, in the real world most hard disks are used to store data supporting many different applications, which include many different types of analysing processes. For example, hard disks on a file server need to move quickly from reading one area on the disk surface to another as different users retrieve different files; seek and latency times measure such performance. Seek time is the average time taken for the read head to move in or out to a given track and then latency time is the average wait time for the particular data to arrive under the head.

Other analysis processes require a single large file to be retrieved. For example, a graphic designer is likely to retrieve single large sized image files, the critical requirement in this case is the speed of data transfer; the time taken to locate the file being relatively insignificant. Measures to determine speed of data transfer include spindle speed and areal density. Spindle speed is the speed at which the disk rotates and is commonly expressed in revolutions per minute (rpm). Higher spindle speeds mean more data passes under the read head in a given time period, hence higher data transfer speeds. Note that higher spindle speeds also improve latency times. Areal density is a measure of the maximum number of bits that can be stored on each square inch of the disk surface. In general, higher areal density means more data passes under the heads and hence the higher the data transfer speed will be. The incredible increases in hard drive capacities and data transfer speeds is largely a result of increases in areal density together with the technology to read such tightly packed data.



*Fig 5.3
Internal view of a hard disk. The
spinning disk platters and read/write
head arms can be clearly seen.*

Most hard disks include a fast memory area called cache; during read operations data passes into cache, this includes the required data together with data the system predicts may soon be needed. Cache is fast chip-based storage; in a hard disk it is included on the hard disk's circuit board. Data within the hard drive's cache can be retrieved many times faster than data on the actual hard disk. If the prediction is accurate, meaning the required data is found to be in cache, then access times will be considerably faster. For example, the hard disk on the machine used to write this book recorded a read access speed of approximately 26 MB per second when none of the data resides in cache, however if all the data is currently in cache then the read access speed is closer to 370 MB per second. Therefore, the amount of cache contained within a hard drive can have a significant effect on data access speeds.

**GROUP TASK Activity**

There are numerous software utilities available for analysing the performance of hard disks. Download such a utility from the Internet, or otherwise, and use it to assess the performance of your hard disk.

RAM (RANDOM ACCESS MEMORY)

In terms of the analysing process the total amount of RAM installed is the most critical measure. RAM holds both the software and the data used by the CPU during processing. If there is insufficient space available in RAM then required instructions or data must be repeatedly written to and retrieved from secondary storage. As secondary storage is many thousands of times slower than RAM, a noticeable drop in performance will certainly result. To put this into perspective, a retrieving process that would take seconds using RAM will take hours using a typical hard disk. Often the cheapest and most effective means for improving performance is to add extra RAM. At the time of writing most personal computers contain a minimum of 128MB of RAM and many contain as much as 1Gb; it is likely that these figures will continue to increase.

The speed at which data held in RAM can be accessed is important for analysing processes; different types of RAM are able to operate at different speeds. The speed at which RAM chips are able to deliver data is determined by the amount of data transferred as a single unit together with the speed at which data can be stored and retrieved. Both these factors must correspond to the specifications of the CPU and also the motherboard onto which the RAM module and CPU are installed.



*Fig 5.4
DDR-RAM module
containing 256MB
of memory.*

**GROUP TASK Research**

Research, using the Internet, different types of RAM module. Classify each according to the amount of data transferred as a unit, together with the speed of transfer.

CPU (CENTRAL PROCESSING UNIT)

The number of bits that can be processed simultaneously, the speed at which instructions are executed and the nature of the instructions are just some of the factors determining the performance of the CPU; there are many others. Different CPU designs are suited to different types of processes. If all other specifications are equal then a CPU capable of processing 64 bits at a time will be twice as fast as one that processes 32 bits; similarly a CPU with a clock speed of 2GHz processes at twice the speed of one with a clock speed of 1GHz. Such measures are only reliable for processors of the same design. Different CPU designs use different sets of instructions and different techniques for executing these instructions. Executing a particular analysing process, such as averaging a set of numbers, will likely require quite a different number of clock cycles on different CPU designs. The instruction set for each family of CPU is different therefore on different CPUs a given process is likely to require the



*Fig 5.5
An Intel Pentium 4 central
processing unit.*

execution of a different number of CPU instructions. Furthermore, different processor designs are able to execute a different number of instructions at the same time, these instructions being executed in parallel and also being at different stages of execution. It seems that comparing the performance of different CPUs is an impossible task.

So how can we determine the best CPU for a particular information system's analysing processes? As was the case with secondary storage, the best method is to execute the same real analysing processes on various types of CPU and record the time taken. This is not practical for all but the largest systems and furthermore the CPU does not operate in isolation, hence other hardware components will affect the outcomes. There are various companies, including various magazines that perform benchmark tests designed specifically for this purpose. Although such tests are unlikely to replicate precise analysing tasks, at least they do provide results that are unbiased. Often these benchmark tests are performed using a variety of software and data scenarios, this makes it possible to select a scenario that best emulates the type of analysing processes relevant to the current information system.



Consider the following:

- If the total amount of data is such that it can be held in RAM during analysis then the speed of the CPU and data retrieval from secondary storage is really not significant.
- For large stores of data it is impossible to retrieve all the data prior to analysis; hence the access speed of secondary storage hardware is critical.
- If the hard disk lights on a file server are continually flashing then that's a good indicator that more RAM is required.
- The most valid means of comparing the performance of different CPUs is to consider their clock speed together with the number of bits processed at any one time.
- Increasing the amount of RAM is the most cost effective means of improving the processing performance of most computer systems.
- Faster access to secondary storage devices results in higher performance compared to increasing the amount of RAM or installing a faster CPU.
- Creating large digital videos is an intensive CPU process, hence upgrading to a faster CPU would be the best method of improving performance.



GROUP TASK Discussion

Each of the above statements is partially correct and partially incorrect, it depends on the individual processes taking place. Describe scenarios where each statement is true and scenarios where each statement is false.



GROUP TASK Discussion

"Hardware should be selected according to the individual needs of each particular information system. Often this is just not possible, so compromises are made." Why is it that compromises are so often made? Discuss.

SOFTWARE FEATURES FOR ANALYSING

Virtually all software performs analysing tasks; if this were not the case then the software application would not produce any useful information. There are a variety of different analysis features that are present in numerous software applications; our aim in this section is to examine common examples of such features and discuss how they are, or can be used to transform data into information.

In terms of software, the efficiency of analysing processes is largely determined by the organisation of the underlying data therefore it is important to consider the type of analysis to be performed when choosing a method of organisation.

In this section we examine the following software features used for analysis:

- Searching/selecting data
- Sorting
- Modelling/simulations
- What-if scenarios
- Charts and graphs
- File comparison

SEARCHING/SELECTING DATA

Most software applications search and select data based on some criteria. In many software applications, the user can directly initiate a search to find all occurrences of a particular data item. For example, the find dialogue from Microsoft



Search

To look through a collection of data in order to locate a required piece of data.

Excel, shown in *Fig 5.6*, is used for this purpose. In many software applications searching takes place as an integral part of some other larger process, in fact many analysing processes include various simple and complex searches. For example, to create a pie chart requires that data be grouped according to various categories; a search is being performed to allocate each data item to its correct category.

What do we mean by searching and selecting and is there a difference? Both searching and selecting are processes that identify required data within a larger set of data. Commonly the term 'searching' is used to describe the process of actually retrieving the data; searching logically examines data items and compares them to some criteria. Any data that matches the search criteria forms part of the resulting information. Such results can be displayed one at a time as they are found or all the results of the search can be retrieved in preparation for further processing or prior to display. On the other hand, the term 'selecting' is generally used to describe the process of specifying the source of the data to be searched. The technique for selecting the source data depends on the information system and the nature of the search. It may mean selecting a particular file or files, it may mean selecting part of a file such as a paragraph in a text document, a particular field within a database, or even a particular range of pixels within an image. Searching is performed on the selected data using the specified criteria.

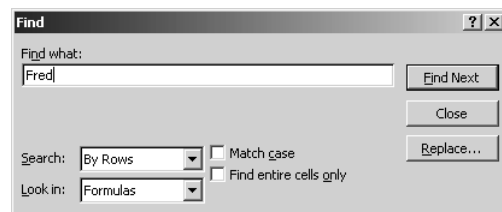


Fig 5.6

Find dialogue from Microsoft Excel.

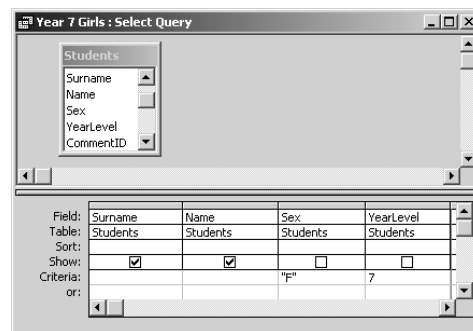


GROUP TASK Activity

Examine the 'Find' dialogue in various software applications. List and describe the various criteria that can be set prior to the search being initiated. In each case, how is the data source selected?

If the source data to be searched is not sorted into an appropriate order then searching requires each data item to be examined in turn. On the other hand, if the data is sorted appropriately then the search process can execute more efficiently. Consider manually searching the white pages telephone directory for a specific name, as the white pages is sorted by name and we wish to find a specific name, the search is a simple one. If the names were in a random order, or we were searching for a specific telephone number then this would be a most tedious task.

The required data is determined by applying criteria, where the criteria is commonly a rule or set of rules that must be correct for each found data item. For example, in *Fig 5.6* on the previous page, the criterion is the text 'Fred' therefore the find process searches for text that equals 'Fred'. The search process considers each data item and decides if the data item fulfils the criteria or rules, if the current data item fulfils the criteria then it becomes part of the results. The mechanics of the actual searching and selecting processes are commonly provided within most software applications; the user does not need to concern themselves with the detail of how the process is performed, rather they merely initiate the search after specifying the source of the data and the search criteria. For example, to retrieve the names of all the year 7 girls within a school's database requires first selecting the fields that contain the student's names within the correct database table. We then search the database for year 7 students who are also girls. *Fig 5.7* shows how this is specified as a query using Microsoft Access. The screen at the top of *Fig 5.7* shows a graphical representation of the structured query language (SQL) statement reproduced below the screen. In the HSC topic "Information Systems and Databases" we examine SQL in some detail.



```
SELECT Students.Surname, Students.Name
FROM Students
WHERE Students.Sex ="F" AND
Students.YearLevel=7;
```

Fig 5.7

Microsoft Access query to retrieve the names of all year 7 girls.



GROUP TASK Activity

Use an Internet search engine to perform searches that include the logical operators NOT, AND and OR. Describe the effect of each of these logical operators.



Consider the following:

- Blurring the edge of a line within a bitmap image.
- Reducing noise within a sampled audio file.
- Producing CMYK colour separations using a desktop publishing application.
- Kerning all AW character pairs within a desktop publisher document.
- A spreadsheet is used to determine the student with the highest mark in an exam.



GROUP TASK Discussion

Identify and describe how searching is used as an integral part of each of the above processes. How does the organisation of the data assist the searching process?

SORTING

Analysing information processes commonly involves sorting data, either sorting into alphabetical or numerical order or even sorting into different categories. When data is sorted, it becomes easier to understand – sorting transforms data into information. For example, an unsorted catalogue of all the different products stocked by a retailer is cumbersome and therefore of limited use, however when this same data is sorted into categories and then the products within each category are sorted alphabetically the catalogue becomes useable information. The catalogue is made easier to search; this is often the purpose of sorting data, to improve the efficiency of searches.



Sort

To arrange a collection of items in some specified order.

All digital data of all media types is represented as binary numbers therefore, sorting digital data is ultimately performed numerically. For alphabetical sorts it is primarily the numerical binary codes, commonly an extension of the ASCII system, which are used to determine the sort order. Let us consider how both numerical and alphabetical sorts are accomplished within software applications.

Numerical sorts consider the total value of the data item; hence an ascending numerical sort, as one would expect, arranges the data from smallest negative value to highest positive value. For example, -500, -5.6, -0.001, 2, 12 and 100 are in ascending numerical order; predictably, a descending numerical sort results in this list being reversed. Problems occur when data items contain characters that are not part of a valid number; in reality this is seldom an issue as the method of representation used for numbers does not permit invalid characters. The problem is encountered when attempting to perform a numerical sort on text data. Often to resolve the problem invalid data items are all placed at the start and then ignored, or the non-valid characters within each data item can be ignored and the remaining valid numbers sorted. Most software uses a combination of both these approaches; if the data commences with an invalid character then that data item is totally ignored, however if it commences with a valid number followed by invalid characters then the valid number forms the basis for sorting.

Alphabetical sorts compare corresponding characters from left to right; if two characters are found to be the same then the next corresponding characters are considered. For example, an ascending alphabetical sort places “Calf” before “Cat” as “l” comes before “t” in the alphabet. Problems commonly occur when numerical data is represented as text and is then sorted alphabetically, for example sorting -500, -5.6, -0.001, 2, 12 and 100 into ascending alphabetical order will, in most software applications, produce the result -0.001, 100, 12, 2, -5.6 and -500. So what is happening? Firstly, most software applications ignore all apostrophes ’ and hyphens when sorting alphabetically, hence the data actually sorted is really 500, 5.6, 0.001, 2, 12 and 100. Ignoring all hyphens and then sorting on the first character in each data item results in -0.001, 12, 100, 2, -500 and -5.6 as 0 comes before 1, which comes before 2, which comes before 5. Now we consider the second character when the first were the same; 0 comes before 2, so 100 appears before 12. What about -500 and -5.6? Most applications sort according to the following order: punctuation and other marks first, followed by the digits 0-9, and finally the characters A-Z; hence -5.6 comes before -500.



GROUP TASK Activity

Sort the numbers 23, 13, 2, 12, 33, 300, 1,45, 6 and 19 into ascending numerical order and then into ascending alphabetical order. Repeat the process using a spreadsheet and then using a word processor. Discuss any problems encountered.



Consider the following:

Text and numeric media types are commonly sorted as part of various analysis processes, however sorting processes are seldom performed on image, audio and video data. Sorting of image, audio and video media is generally restricted to sorting various attributes of the files used for storage.

In most operating systems it is possible to sort by various attributes of files stored on various secondary storage devices. *Fig 5.8* shows this facility within Explorer in Windows XP.

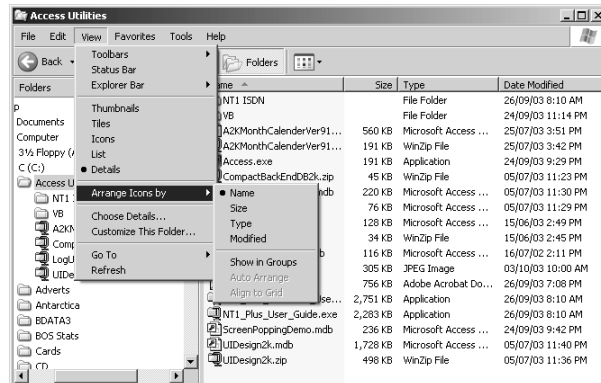


Fig 5.8

Screenshot from Explorer within Windows XP.



GROUP TASK Discussion

Why is sorting not commonly used for analysing image, audio and video data? Discuss with reference to the organisation of these media types.



GROUP TASK Discussion

In *Fig 5.8* it is possible to perform ascending or descending sorts on Name, Size, Type or Date Modified. Classify each of these different sorts as either numerical or alphabetical sorts. Justify your answers.



Consider the following:

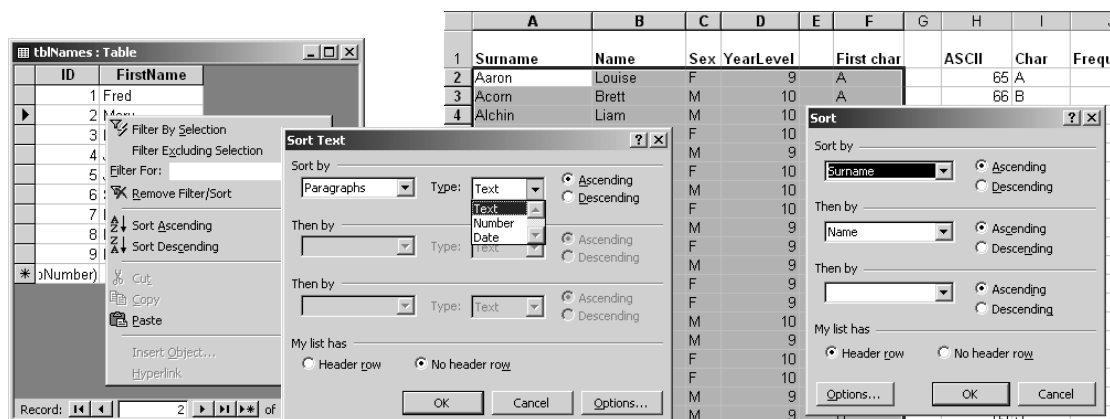


Fig 5.9

Sort functions in Microsoft Access, Word and Excel.



GROUP TASK Discussion

The sort functions used in databases, word processors and spreadsheets are implemented in different ways. Describe the differences and explain why these differences exist.

SET 5A

1. During analysis data moves from:
 - (A) RAM into secondary storage prior to analysis within the CPU.
 - (B) secondary storage directly to the CPU, once processed it is held in RAM.
 - (C) secondary storage into RAM and then to the CPU.
 - (D) the CPU into RAM and then onto secondary storage.
2. The analysing process:
 - (A) transforms data into information.
 - (B) does not alter the data.
 - (C) makes sense of data for humans.
 - (D) All of the above.
3. Fast chip based storage present of most hard disks is called:
 - (A) RAM
 - (B) cache
 - (C) a register.
 - (D) secondary storage.
4. 1, 12, 4, 500, 58, 9 has been sorted into:
 - (A) ascending numerical order.
 - (B) descending numerical order.
 - (C) ascending alphabetical order.
 - (D) descending alphabetical order.
5. Secondary storage can be considered the 'weakest link in the chain' because:
 - (A) it is significantly slower than RAM or the CPU.
 - (B) it is permanent storage.
 - (C) hard disks are more prone to failure than RAM or CPU chips.
 - (D) computers use secondary continuously whilst RAM and the CPU are used only when required.
6. If all other parameters are equal then a 32-bit CPU will:
 - (A) be half as fast as a 16 bit CPU.
 - (B) be double the speed of a 64 bit CPU.
 - (C) be half as fast as a 64 bit CPU.
 - (D) be four times as fast as a 16 bit CPU.
7. In regard to analysing, the most important property of RAM is:
 - (A) its total memory capacity.
 - (B) the speed at which it can deliver data.
 - (C) the design of the RAM module.
 - (D) its compatibility with the CPU.
8. When searching, each data item must be examined in sequence if:
 - (A) the data is sorted into an appropriate order.
 - (B) the data is not sorted into an appropriate order.
 - (C) the file being searched has been indexed.
 - (D) the search includes more than one field.
9. The time taken to locate a particular file on a hard disk can be measured using:
 - (A) areal density.
 - (B) seek and latency times.
 - (C) spindle speed.
 - (D) data transfer speed.
10. All sorting performed by computers:
 - (A) is ultimately performing a numerical sort.
 - (B) uses the ASCII code of each character.
 - (C) ignores apostrophe and hyphen characters.
 - (D) examines each corresponding character commencing on from the left.
11. Describe different measures used to compare the performance of hard disks.
12. Describe the relationship between RAM, secondary storage and the CPU during a typical analysing process.
13. Do you agree with the statement: "Each of the other information processes exists primarily to support the analysing process"? Justify your response.
14. According to the syllabus hardware requirements for analysing, include "large amounts of primary and secondary storage allowing for fast processing". Do you agree with this statement? Discuss both yes and no arguments.
15. Internet search engines rank or sort results based on some criteria. Examine a number of popular search engines and determine the criteria being used to rank the search results.

MODELLING/SIMULATIONS

A model is used to represent some real world system or thing; therefore modelling is the process or act of creating a model. For example, the plans for building a house are really a model of the final house, in this case a scale drawing, together with various specifications detailing the materials and method of construction. House plans are static models; they describe or represent a single house. A simulation alters various parameters of a model, often including time, to produce imitations of a system's operation. In essence, a simulation gives



Model

A representation of something. Computer models are mathematical representations of systems and objects.



Simulation

The process of imitating the behaviour of a system or object. A specific application of a model.

life to a model by altering and processing its inputs. For example, a flight simulator responds to various different inputs that change over time. The mathematical description of the rules and properties governing the behaviour and operation of the aircraft and its environment form the model. During simulation, inputs are collected from the pilot and generated for the environment to alter the results produced by the model's rules. Hence simulation is a process that imitates the behaviour of a system or object where the representation of the system or object is a model.

Computer modelling and simulation are analysing tasks; they utilise computer resources to represent something mathematically and to produce meaningful information. Computer modelling and simulation are commonly used when it is impractical, or even impossible to analyse the real system or object. For example, training pilots on real aircraft is costly and potentially dangerous. Computer simulators allow pilots to gain experience dealing with all types of potential problems in safety and at minimal cost. *Fig 5.10* shows a 747 flight simulator; in this case the cockpit itself is a realistic model of an actual 747 cockpit. There are also computer models and simulations that represent an imaginary view of the real world; many computer and video games are examples of such simulations.



Fig 5.10
Boeing 747 flight simulator cockpit.

Many models and simulations utilise images and video media to communicate information, however the data used to produce such information is numeric. For example, economists produce models and simulations of the stock market and other aspects of the economy; in this case, charts and graphs are produced from numeric information created during the simulation. Weather forecasting simulators collect and analyse vast quantities of atmospheric data from satellite and ground stations, the numeric results are subsequently used to construct images and video sequences such as those displayed each evening on the news. The initial numeric information may well be understood by meteorologists, however the general public better understands the final image and video data.



Consider the following article:

Forecasters rely on computer models

By Chad Palmer of USA TODAY Information Network, 30/6/2000

With the explosion of computer technology scientists have developed many computer models to help forecasters analyze and predict the weather.

Computer models depend on the fact that mathematical equations can describe the physical changes that govern the weather, just as equations describe movements of the solar system well enough for solar and lunar eclipses to be predicted years, even centuries, in advance.

Equations describing the solar system are complicated. Still, mathematicians and astronomers have been able to predict eclipses for centuries. The atmosphere's equations are much more complex. Solving them had to wait not only for more knowledge, but also for computers. And, even with the fastest computers, meteorologists can't forecast day-to-day weather for more than about a week ahead of time. The atmosphere is just too complex, among other reasons.

When researchers began developing the first computer models for the earth's atmosphere in the 1950s, they worked with computers that were extremely limited compared with today's. As a result, the first models were overly simplified, but still provided valuable insight into the atmosphere's future state. As computer technology advanced, the complexity of the forecast models increased and more of the dynamical and physical factors influencing the atmosphere were taken into account. This improved forecasts.

These models were never intended to replace human forecasters. Instead, the models were developed as aids. Human forecasters study the output from models over a long time period and compare the forecast output from the models to the actual verification for the forecasted time period. This is how model biases, model strengths, and model weaknesses are determined. Often, human forecasters will modify model output based on past experience in forecasting the weather and physical and dynamical reasons.



GROUP TASK Discussion

According to the above article the output from computer models are merely an aid for human forecasters. List and describe reasons why computer weather forecasting models are merely an aid for human forecasters.



Consider the following:

Computer modelling and simulation is used in many areas, including:

- Product design
- Space exploration
- Economic forecasting
- Nuclear power stations
- Training
- Urban planning
- Medical science
- Entertainment



GROUP TASK Research

Working in groups, choose one of the above dot points. Research specific examples that illustrate how modelling and simulation is used in your chosen area. Present your findings to the class.

WHAT-IF SCENARIOS

What-if scenarios allow you to consider more than one set of options; we do this often during a normal day. For example, each morning we make an assessment of the likely weather for the day. What if it rains? What if it's a really hot day? Based on our assessment we decide on what clothes to wear and perhaps whether we need a raincoat or an umbrella. Our brain performs the processing; it analyses the different possible inputs and produces an appropriate set of outputs. In our weather example it is likely that a number of different possibilities will be considered; we choose the most appropriate outputs to act upon. What-if scenarios created using computers perform similar processes; different sets of inputs are analysed to determine a corresponding set of resulting outputs. The "What-if" analysis process aims to produce the most likely outputs for each set of inputs. The aim is to predict the likely, or at least possible, consequences for each particular set of inputs; these predictions can then be used to make better informed decisions.

When performing 'What-if' analysis it is the inputs or data that is changed; the processing that transforms these inputs into information, in the form of predictions, remains the same. Therefore when designing a what-if scenario it is vital to understand the detailed nature of the analysis processes for all possible sets of inputs. In most cases these processes operate on numeric data using various mathematical and statistical calculations; for this reason spreadsheets are particularly suitable software tools for what-if analysis. Spreadsheets automatically recalculate each formula immediately after any input data is altered; therefore the information displayed always corresponds to the current data.



Consider the following:

The spreadsheet in *Fig 5.11* is used to calculate the regular payments, total payments and total interest for a loan. The results of various "What-if" scenarios are determined by altering the input data within cells C3 to C6.

The following formulas are used:

- $C8 = \text{PMT}(C4/C5, C5 * C6, C3) * -1$
- $C10 = C5 * C6 * C8$
- $C11 = C10 - C3$

	A	B	C	D	E
1					
2		Loan calculator			
3		Loan amount	\$ 450,000.00		
4		Annual interest rate	6.50%		
5		Payments per year	26		
6		Years to repay	25		
7					
8		Payment due per period	\$ 1,401.54		
9					
10		Total payments	\$ 911,000.84		
11		Total interest paid	\$ 461,000.84		
12					

Fig 5.11

Loan calculator implemented in Microsoft Excel.



GROUP TASK Activity

Reproduce the above spreadsheet. Using a variety of different sets of inputs compare the effect of making repayments monthly, fortnightly and weekly. Can you explain your results?



GROUP TASK Activity

Modify the spreadsheet to calculate the number of years to repay the loan when the payment due per period is part of the input data. Why would such an alteration be useful?

CHARTS AND GRAPHS

Charts and graphs are used to visually illustrate the relationships between two or more sets of data. For example, the rainfall each month for a particular town contains two sets of data, the months and the rainfall figures. Consider the example table and column graph in *Fig 5.12*. Within the table the precise value of each data item can be seen, however the graph more effectively shows the distribution of rainfall throughout the entire year. Different information is highlighted on the graph compared to the table.

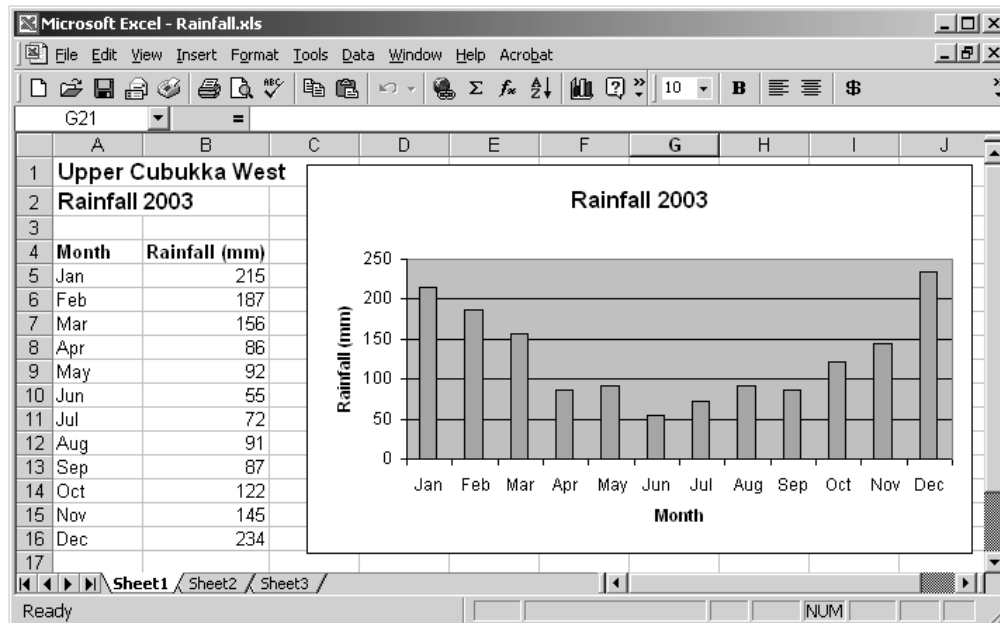


Fig 5.12

Rainfall data displayed in a table and as a column graph using Microsoft Excel.

Different types of graph or chart emphasise different types of information. Let us consider examples of the more common graph types together with their major purpose in terms of communicating information.

Column and bar graphs

Column graphs display data values vertically whereas bar graphs display data values horizontally. Both column and bar graphs are well suited to sets of data where the categories or entities are not numeric or have no inherent order; in this context the set of numeric values measure the same thing for various different entities. For example, in *Fig 5.13*, each state is a different entity; the order in which these entities appear is not important, whereas each numeric value is a measurement of the same quantity. A line graph would be inappropriate for graphing this data, as points on the lines between different states have no meaning.

The graphs in *Fig 5.13* are based on a single data series. Column and bar graphs can be created to graph multiple data series for each entity. Each data series can be shown as a separate column or bar, or they may be stacked together to show the total for each entity.

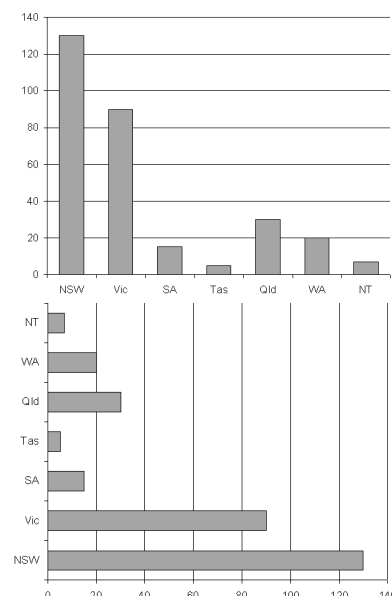


Fig 5.13

Column graphs and bar graphs display the relative differences between data values.

Line graphs

Line graphs are commonly used to display a series of numeric data items that change over time. They are used to communicate trends apparent in the data. Lines connecting consecutive data points highlight the changes occurring; when all such lines are plotted overall trends emerge.

When using line graphs the source data must be sorted by the data to be graphed along the horizontal or x-axis. For example, in *Fig 5.14* the horizontal axis contains the months of the year, if this data were not sorted correctly then the trends communicated by the lines connecting each data value would be incorrect.

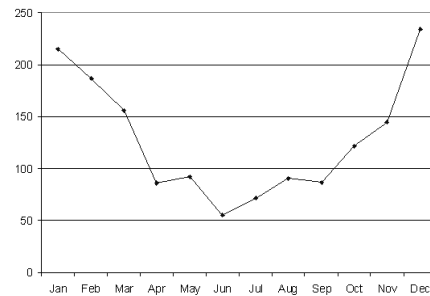


Fig 5.14

Line graphs highlight trends in a data series. Both axes should contain ordered data.

Pie charts

Pie charts show the contribution or percentage that each data item makes to the total of all the data items. For example, *Fig 5.15* clearly communicates that NSW contributes far more to the total than any of the other states and that Tas. and NT contribute the least.

The nature of pie charts means they are only able to plot a single data series. Pie charts do not provide information on the precise value of each data item rather they communicate the relative differences between each discrete category on the graph.

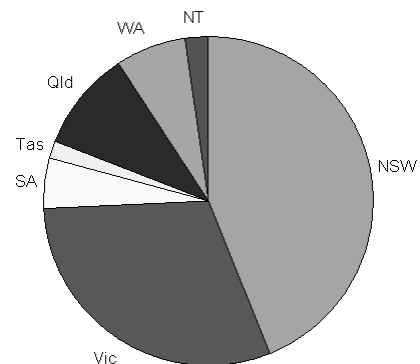


Fig 5.15

Pie charts highlight the contribution each data item makes to the total.

XY graphs

XY graphs are used to plot pairs of points. The source data being composed of a series of ordered pairs. Each ordered pair is composed of an X coordinate and a Y coordinate used to determine the position of a single point on the graph. When these points are connected using a series of smooth curves a continuous representation of the relationship between the X and Y coordinates is produced.

In most cases the source data for XY graphs is a series of samples taken at various intervals. For example, in *Fig 5.16* samples or ordered pairs have been produced by incrementing the X coordinate by 0.5 and then calculating the corresponding Y coordinate; the final graph is produced by connecting these points using a smooth curve.

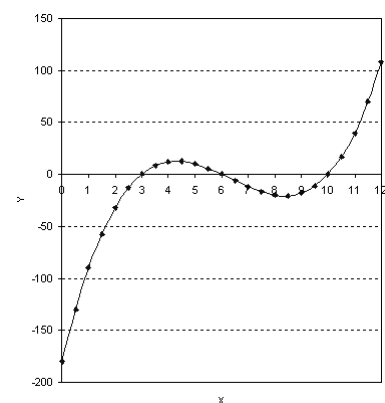


Fig 5.16

XY graphs are used to plot a series of ordered pairs.

In contrast to line graphs, it is not necessary for the X coordinates to be evenly spaced. It is quite common to obtain samples at random times which can then be connected to form a continuous curve. Furthermore, the curve can be extrapolated in an attempt to describe trends outside the range of the sample data.



Consider the following:

- Displaying the percentage of IPT students in each performance band for the past three HSCs.
- Determining the likely future percentage increase in real estate prices for a given region.
- Describing the relative popularity of different leisure activities.
- Finding a relationship between time spent studying and exam results.



GROUP TASK Discussion

For each of the above scenarios; identify appropriate source data and then describe and justify a suitable type of graph that could be used to analyse this data to best produce the required information.

FILE COMPARISON

There are various utilities available that compare either the properties or actual contents of files. Many of these utilities are designed to synchronise files stored on different computers and storage devices. For example, current versions of Microsoft Windows include a special type of folder called a 'Briefcase'. If this folder is copied to another computer, commonly a laptop, then it is a simple matter to synchronise the original files with those held on the laptop. Windows Briefcase (see Fig 5.17) does not examine the actual contents of files; rather it compares the modified dates and highlights any differences found.

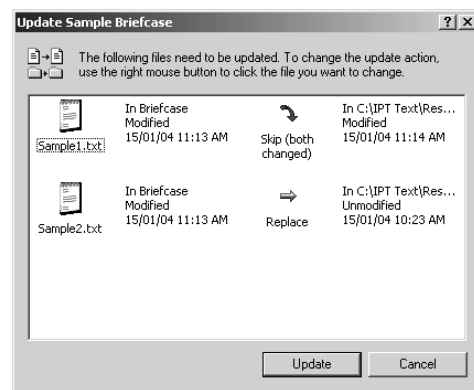


Fig 5.17

Window's Briefcase compares the modified dates of files.

File comparisons that examine the actual contents of files are available in various levels of complexity suited to a variety of needs. For example, WinDiff is an application that compares the contents of files line by line. This type of file comparison is particularly useful for comparing the contents of text files, including program source code. Any differences found are then highlighted; the screen in Fig 5.18 highlights differences found in lines 28 and 38. The darkly shaded lines are from the first file and the lightly shaded lines from the second file. Similar applications are available for comparing the contents of most common file types.

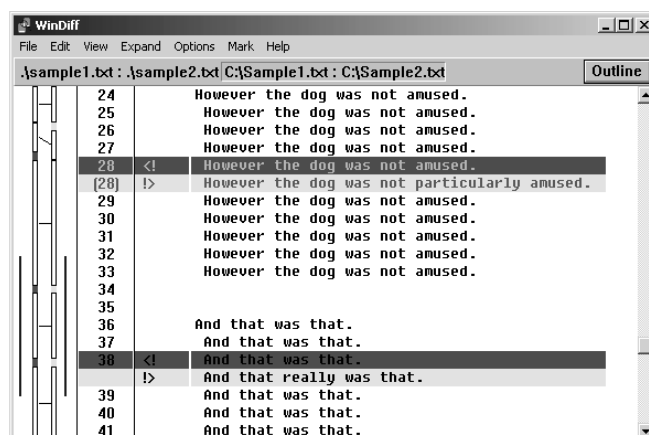


Fig 5.18

WinDiff is an analysis tool for comparing the contents of two files line by line.

Many database management systems include file comparison functionality. In Microsoft Access and SQL server such functionality is called replication. Replication is a system that allows the contents of records to be synchronised across multiple copies of the data. Replication is commonly used when a single database is shared over large distances; such distances make fast and reliable network connections difficult to maintain.

When synchronisation is initiated on a replicated database each modified record is compared to the same record in the original or master copy of the database. Generally, the most recent modification found is then copied to both databases, however it is possible to give priority to changes made in one database over those made in other copies of the database. The synchronisation process results in the original or master database having the most recent data; when further copies are later synchronised with the master the same process takes place, hence over time all copies of the database receive the changes made within any other copy of the database.



Consider the following:

Each of the following scenarios involved comparing the contents of different files:

- The Safe-T-Cam system (see page 100).
- The operation of an optical mouse (see page 87).
- Compressing video using block based coding (see page 63).



GROUP TASK Discussion

Review each of the above scenarios examined previously in the text. For each scenario discuss how comparing different files assists the information system to achieve its purpose.

NON-COMPUTER TOOLS FOR ANALYSING

Digital data that has been organised appropriately certainly improves the efficiency and accuracy of most analysing tasks, however there still remain numerous examples of information systems that either do not use computers at all or that include non-computer based analysis techniques. For example, the image in *Fig 5.19* shows a manual filing system used within a museum. Museums may well have their collection catalogued in detail on computer; such catalogues often include image and even video media. Despite the detailed nature of the digital data held much of the analysis of individual items still requires the actual physical artefact; in fact the contents of the digitised catalogue is largely the information derived from manual analysis of each artefact.



Fig 5.19
Manual filing system used within a museum.



GROUP TASK Discussion

Compile a list of non-computer data you encounter or use during a typical day. Discuss how each is analysed to obtain information.

In this section we discuss non-computer tools commonly used to analyse data and compare them to their computer-based equivalents, the aim being to illustrate the strengths and weaknesses of different approaches to analysis. We restrict our discussion to searching manual filing systems and non-computer models and simulations.

SEARCHING MANUAL FILING SYSTEMS

To search for particular data within a manual filing system is time consuming even when the files are sorted based on the field being searched. Furthermore, files can only be viewed by a single person at any one time and after use they must be returned to their correct position. Consider a manual filing system containing a file for each client sorted by their names. If the business has more than a hundred or so clients then locating an individual client's file, even when their name is known, takes a minute or so. If the files are not sorted on the field being searched then the time taken to perform the search increases enormously. Imagine trying to retrieve the address of all clients who live in a particular suburb; such a search involves laboriously examining every single client file. Such processes can be completed virtually instantaneously on similar computerised systems. Despite this, many businesses and organisations continue to maintain manual filing systems.

Some common reasons for maintaining manual filing systems include:

- **Cost** – hardware, software, data entry and training costs must be met upfront when implementing new computer based information systems. In comparison, manual systems require relatively minor upfront costs.
- **Volume of data** – most new businesses start up small, with correspondingly small amounts of data. As a consequence the time taken to manually analyse this data is not significant. For example, searching a collection of 50 files takes a matter of minutes and may well be more time efficient compared to starting up a software application, entering the search criteria and displaying the resulting information.
- **Training** – to use a computer-based filing system requires knowledge and understanding of the hardware and software. In contrast, manual filing systems require minor training. People easily grasp the organisation of data within a filing cabinet; each file contains all the data about a specific entity. In fact most operating systems use a manual filing system as a metaphor to assist users to understand the organisation of data within the computer. For example, the desktop is used to hold current or often used data items, hard disks contain folders which in turn contain files and an icon of a trashcan is used when deleting files. Such operating systems quite reasonably assume users understand manual filing systems.
- **Nature of the data** – most computerised information systems require data to be organised in a predetermined and highly structured manner, if additional fields are needed then the structure of the entire system must be altered. Manual filing systems do not enforce such rigidity; different files within a manual filing system can easily contain data of different types organised in different ways. Furthermore, the addition of extra data to an individual file has no effect on other existing files.



GROUP TASK Discussion

Most households maintain a manual filing system for bills, appliance manuals, mortgage and other documents. Discuss reasons why a manual filing system is more suited to the storage and analysis of household data.



Consider the following:

The motel in a particular country town has 12 rooms. Currently the motel uses a manual filing system. When a guest checks into the motel, they complete a registration card, which includes various demographic entries together with how they heard about the motel. Whilst this card is being completed the receptionist checks the room register and assigns the guest a room, which involves adding the guest's name to the room register. All details during the guest's stay, including the room number, are included on the guest's registration card. For example, food, beverage and movie hire details and charges. When a guest checks out of the motel all charges on the registration card are totalled, the guest is asked to verify the information on the card and finally the guest pays their account.



GROUP TASK Discussion

Identify and categorise each of the information processes occurring in the above scenario as either collecting, organising, analysing, storing and retrieving, processing or displaying.



GROUP TASK Discussion

The motel owner wishes to determine the effectiveness of an advertisement placed in the NRMA's accommodation guide. Their advert has been in the guide for the past 10 months. Discuss analysis techniques the motel owner could use to determine the ads effectiveness.

NON-COMPUTER MODELS AND SIMULATIONS

Most products today are designed and their operation is modelled and simulated using computers. However prior to production commencing a non-computer model or prototype is commonly built and tested; a prototype being a full size model of the final product. For example, motorcycle and car manufacturers create prototypes of each new vehicle prior to its production. The prototype is used for testing as well as for market research.



Fig 5.20

Motorcycle prototypes.

Why are such non-computer models and simulations built?

- To demonstrate feasibility – a working model or prototype clearly demonstrates that the product in fact operates and fulfils its objectives.
- To resolve design unknowns – not all variables can be included in a computer model. Non-computer models or prototypes allow the product to be tested in real world conditions.
- To resolve human factors issues – a non-computer model is needed to test the product is human friendly. For example, do people like the look of the product and can they understand how it will be used? A real world model or prototype is needed to answer such questions.
- To market an idea – an actual example or model of a proposed product is invaluable when trying to assess or create a market for a new product. Without a real world example most will view the product more as a theoretical idea rather than a product design ready for production.



Consider the following:

Mars Mice

In 2006 a group of mice-astronauts will orbit Earth inside a spinning spacecraft.

Their mission: to learn what its like to live on Mars.

Humans need gravity. Without it, as astronauts have vividly demonstrated, our bodies change strangely. Muscles lose mass, and bones lose density. Even the ability to balance deteriorates.

From long experience on the space shuttle and various space stations, we have some knowledge of how mammals, especially people, respond to 0-g. We have even more experience with 1-g on Earth. But we still don't know what happens in between.

What, for example, will happen to humans on Mars where the surface gravity is 0.38-g? Is that enough to keep human explorers functioning properly? And, importantly, how easily will they readapt to 1-g, once they return to Earth?

A team of scientists and students from the Massachusetts Institute of Technology (MIT), the University of Washington, and the University of Queensland, in Australia, plans to explore these questions. They're going to do it by launching mice into orbit.

"What we're doing," explains Paul Wooster, of MIT, and program manager of the Mars Gravity Biosatellite project "is developing a spacecraft that is going to spin to create artificial gravity." The satellite will spin at the rate of about 34 times each minute, which will generate 0.38-g, the same as gravity on Mars.

The team hopes to launch the Biosatellite in 2006. The mice will be exposed to Mars-gravity for about five weeks. Then, says Wooster, they'll return to Earth alive and well. The mice will descend by parachute and land near Woomera, Australia, inside a small capsule reminiscent of NASA's old Apollo capsules.

The research will focus on bone loss, changes in bone structure, on muscle atrophy, and on changes in the inner ear, which affects balance. "The main thing we're trying to do," says Wooster, "is to chart a data-point between zero-gravity and one-gravity."

As they orbit the earth, the mice, each in its own tiny habitat, will be painstakingly observed. Each habitat will have a camera, so that the researchers can monitor mouse activity. Each will have its own pump-driven water supply, so each mouse's water consumption can be tracked.

Each mouse's wastes will be collected in a compartment beneath its habitat; the compartment will contain a urinalysis system checking for biomarkers that indicate bone loss.

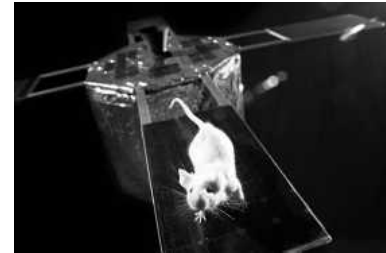
Each habitat will also be equipped with a body mass sensor, which will take frequent readings. This will also allow the researchers to track how the weight of the mice changes over the course of the five weeks.

Each mouse will also have toys to keep it busy. "We may give them a wooden block to chew on," says Wooster. That'll keep them happy, and will also prevent them from chewing on the habitat. They might have a small tube to run through.

No wheels, though, says Wooster, because NASA has learned that exercise can counteract some of the effects of low-gravity on astronauts. A mouse with a wheel in its cage can actually run several miles a day. "We don't want to give the mice a countermeasure in terms of exercise."

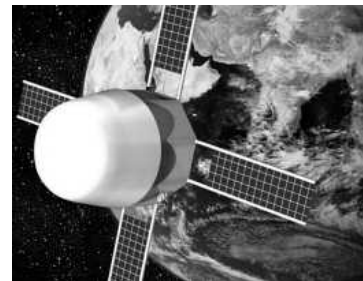
(Source: science.nasa.gov)

10/2009 Update: The satellite is yet to be launched. The project was discontinued on 24/6/2009, however there are indications that it may be recommenced with assistance from the MarsDrive organisation.



A mouse-astronaut candidate poses atop a model solar panel.

Credit: MarsGravity.org



An artist's rendering of the Mars Gravity Biosatellite in Earth orbit.

Credit: MarsGravity.org



GROUP TASK Discussion

Read the above article and discuss reasons why such a non-computer simulation would be necessary.



GROUP TASK Discussion

Make a list of the data to be collected and describe how this data could best be analysed to produce the required information.

SET 5B

1. A simulation:
 - (A) imitates the behaviour of a real or imagined system or object.
 - (B) is a specific application of a model.
 - (C) represents a real world system or thing.
 - (D) Both A and B.
2. Altering the inputs in a spreadsheet and observing the outputs is an example of:
 - (A) a simulation.
 - (B) data processing.
 - (C) what-if analysis.
 - (D) creating a model.
3. Pie charts are suitable for communicating:
 - (A) multiple data series.
 - (B) the different percentages each data item makes to the total of all data items.
 - (C) future trends likely to occur in the data.
 - (D) the relationship of a pair of data sets to each other.
4. An advantage of manual filing systems over similar computer-based systems is:
 - (A) sorting and searching is more efficient and accurate.
 - (B) each record can contain media of different types.
 - (C) all records must have an identical structure.
 - (D) more extensive training is needed to use manual filing systems.
5. Software that replaces files with more recent versions are likely to be:
 - (A) comparing the contents of files.
 - (B) comparing the dates files were created.
 - (C) comparing the dates files were modified.
 - (D) comparing each corresponding line of data within the files.
6. Non-computer models and simulations are built to:
 - (A) demonstrate feasibility.
 - (B) resolve human factors.
 - (C) market an idea.
 - (D) All of the above.
7. The data series plotted along the X axis must always be sorted when using a:
 - (A) line or XY graph.
 - (B) column or bar graph.
 - (C) pie chart.
 - (D) All of the above.
8. Which of the following is true?
 - (A) there can be many models created using a single simulation.
 - (B) there can be many simulations created using a single model.
 - (C) each simulation must use a different model.
 - (D) modelling is the process of creating a simulation.
9. The best chart type for displaying the number of cars, trucks, buses and motorbikes passing a given point would be a:
 - (A) line graph.
 - (B) column graph.
 - (C) pie chart.
 - (D) XY graph.
10. Spreadsheets are commonly used for creating 'What-if' scenarios because:
 - (A) they automatically recalculate all outputs each time an input is altered.
 - (B) most scenarios involve processing numeric data.
 - (C) commonly the processing utilises mathematical and statistical functions.
 - (D) All of the above.
11. Define each of the following terms:
 - (a) model
 - (b) simulation
 - (c) What-if scenario
12. The profit made by Eclectus Software Pty. Ltd. has risen each year. In 1997 profit was \$1.2 million, in 1998 it was \$1.5 million, 1999 \$2.1 million, 2000 \$2.4 million, 2001 \$2.8 million, 2002 \$3.5 million, and in 2003 profit was \$4.0 million.
 - (a) Graph the above data. Justify the type of graph you use.
 - (b) Use your graph to estimate likely profits for 2004, 2005 and 2006.
 - (c) Do you think your profit estimates in (b) will prove to be accurate? Discuss.
13. Comparing the modified date of files is commonly used as the basis for synchronising files on a desktop computer with those on a laptop. Why is this a suitable technique when there is a single user of both machines but not when there are multiple users?
14. Why would a new business choose to use a manual filing system rather than a computer-based filing system? Discuss.
15. Imagine the RTA proposed to replace all practical driving tests with realistic simulator tests. Discuss advantages and disadvantages of such a proposal.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH ANALYSING

Social and ethical issues arising from the analysis of data commonly involve concerns in regard to privacy, security and accuracy. In Chapter 1, we considered each of these areas, however we focused on the privacy, security and accuracy of data during and after collection. In this section we assume the data has been collected responsibly and accurately. We are concerned with social and ethical issues occurring as this data is transformed into information.

Areas for consideration include:

- Unauthorised analysis of data. Is analysis of the data authorised?
- Data incorrectly analysed. Do the results of the analysis correctly reflect the data?
- Linking databases for analysis. Have privacy concerns been addressed?

UNAUTHORISED ANALYSIS OF DATA

According to National Privacy Principle 2 (NPP2) of the *Privacy Act 1988*, organisations that hold data on individuals are required to disclose how such data will be used. Essentially NPP2 specifies restrictions and requirements in regard to how an organisation can analyse its data. For example, if personal data is collected by a business to enable it to complete a sales transaction then the business is not authorised to use such data for unrelated purposes. If they wish to analyse their sales data to better target future marketing then they must first disclose this purpose to each individual whose data they wish to use. Such disclosure must contain a clear indication of how an individual can choose to be excluded from the process. Commonly such disclosure is included in the fine print at the time the data is collected, together with a simple question indicating agreement. There is a difference between legally collecting personal data and legally analysing such data; disclosure of the purpose and detail of both these processes require the consent of the individual.



Consider the following:

- Fred buys a new BBQ at a large department store. The store offers 12 months interest free terms; which Fred agrees to. Some six months later the BBQ develops a major fault. Fred attempts to have the fault rectified by the store with little success. As a consequence Fred stops making repayments. Eventually the problem is resolved and Fred completes all his repayments. However, some 2 years later Fred is refused credit for a car loan; apparently on the basis that he has defaulted on repayments in the past.
- Wilma has worked for the same employer for 20 years. Unfortunately she has a significant conflict with a new supervisor and subsequently seeks other employment. Wilma experiences trouble obtaining her entitlements from her previous employer. It seems they have examined various files on her personal computer and have reached the opinion that she had been performing extensive amounts of private business using the companies resources; hence their objection to paying all her entitlements.



GROUP TASK Discussion

Identify and discuss issues apparent in each of the above scenarios that relate to the unauthorised analysis of data.

DATA INCORRECTLY ANALYSED

Although the data used for analysis may well be correct the actual analysis process can itself be flawed, the result being incorrect or at least misleading information. Such inaccuracies routinely occur when attempting to identify future trends based on past data. For example, weather forecasts are never completely accurate; rather they describe the most likely weather patterns.

There may be undetected errors within the analysis process. For example, a query may exclude records that should actually have formed part of the data set. The resulting information will be incorrect or at least incomplete. Perhaps certain variables have not been included within the analysis. This is particularly a problem when computer models and simulations are built; it is often impossible to include all variables within such models. This is the primary reason why full working prototypes are created for most products prior to production; such prototypes can be tested in the real world.

Techniques used to communicate information after analysis can result in misleading information being communicated. This is particularly so in regard to graphs or charts. The scale of the axes and the type of graph influences the relative importance of differences and trends communicated. Also, the categories used to group data can be selected in such a way that summary information becomes misleading.



Consider the following:

Consider the data and three graphs shown in *Fig 5.21*. All three graphs display the results from throwing a standard die one thousand times, yet the information communicated is quite different.

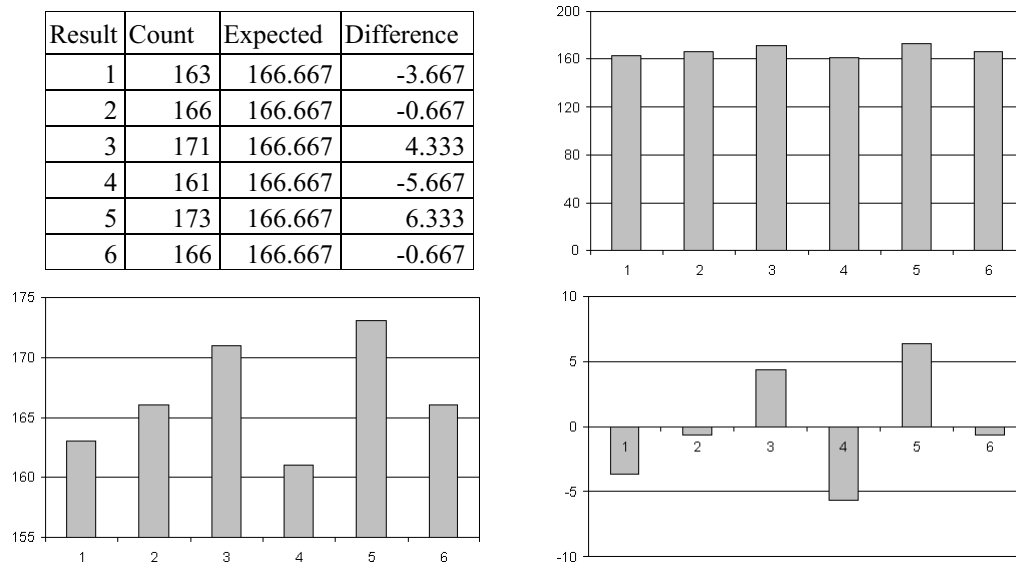


Fig 5.21

Results and graphs from throwing a die experiment.



GROUP TASK Activity

The table was actually produced using a spreadsheet simulation rather than by throwing a real die. Create such a model using a spreadsheet.



GROUP TASK Discussion

Describe the information communicated by each graph. Are these graphs reasonable analysis tools for analysing the simulation? Discuss.

LINKING DATABASES FOR ANALYSIS

Linking together multiple databases can be used to formulate new information that was not present in any of the individual source databases. This type of analysis is known as ‘data mining’, an analysis process that discovers new unintended relationships among the data. Data mining is usually applied to large databases in an attempt to find new patterns that result in new meanings being derived from the data. For example, retail stores use data mining to determine which products to stock and even how to place them within each store. They can use such techniques to target the marketing of particular products to customers more likely to purchase. Such information increases the profitability of the business, but is it socially and ethically acceptable?



Consider the following:

- A telephone company wishes to target people who are most likely to change their phone carrier. They use their own customer data combined with general census data and data within the existing telephone book. A relationship is found that indicates older people with low incomes are more likely to change carrier based on expected levels of personal service and young people are more price sensitive. As a consequence the telephone company creates two new sales packages to reflect the findings.
- A medical research company determines that a small but significant proportion of people who have suffered from a particular fatal disease also have an abnormality in a specific gene within their DNA. The results are published widely. As a consequence a large life insurance company makes it mandatory for all new customers to submit DNA samples to test for the abnormality prior to insurance being granted.
- A pattern matching algorithm is developed to recognise an individual’s handwriting; in particular their signatures. Software that implements this algorithm is used by the federal police to link signatures on various documents held by government departments; the aim being to build up more precise profiles of potential criminals.
- Web harvesting software is available that scours the web for email addresses and websites of potential customers. When ‘The Lord of the Rings’ movies were first released the software product WebQL was used to search the web for email addresses and websites of people with an interest in the Tolkien series. Unsolicited emails were then sent to these people.



GROUP TASK Discussion

For each of the above scenarios, identify the different data sources and describe the nature of the new information discovered.



GROUP TASK Discussion

Each of the above scenarios includes invasion of privacy concerns. Identify privacy concerns within each scenario and then discuss these concerns in terms of your knowledge of the Australian Privacy Act 1988.



HSC style question:

Each day farmer Fred records the minimum temperature, the maximum temperature and any rainfall. He has been recording these details on paper for more than 25 years.

Fred has purchased a laptop and has been teaching himself to use a word processor, spreadsheet and flat file database. His aim is to graph all his historical weather data to see if there is any evidence of global warming on his property. Fred intends to share the results from his weather analysis system with the local community to illustrate the effects of global warming.

- Define the term *data* and identify the data for Fred's weather analysis system.
- Define the term *purpose* and identify the purpose of Fred's weather analysis system.
- Propose how Fred could organise and analyse his weather data to achieve the system's purpose.
- Outline areas where bias could potentially affect the validity of Fred's results.

Suggested Solution

- Data is the raw facts input into a system which are analysed to provide information that can be used by humans. In this system the data collected includes daily measurements of rain, minimum and maximum temperature, and the date when each set of data was collected.
- The purpose of a system is a statement identifying who the information system is for and what it needs to achieve.

This particular system is for Fred initially and then for wider publication, to determine if there has been a significant amount of change in the weather over time indicating trends caused by global warming. Interestingly, even though Fred's stated purpose for the system is to illustrate the effects of global warming, this particular system can only indicate the existence or otherwise of changes in weather patterns on Fred's property.

- Fred could use a spreadsheet as the main application; however a database could be used to simplify the initial data entry of his vast historical paper records.

In terms of the organisation of the data on the spreadsheet:

- One worksheet is used for all the data. Fred could enter all his historical data directly or if he used a database to enter the data then he would import the data into the spreadsheet from this database.
- Each row on the worksheet contains the data for each day. Four columns are required; one for the date, one for minimum temperature, one for maximum temperature and one for rainfall.

In terms of analysing the data:

- Calculations (formulas) to summarise Fred's data would be entered on a second worksheet. As more than 25 years of data is present it would be appropriate to find the average minimum and maximum temp for each month of each year. Also the total rainfall for each month of each year would be calculated.

- Two charts could then be created that include the summarised information. One chart graphs the monthly information and the other the yearly.
 - Each chart includes a line graph for average minimum temperature and a line graph for average maximum temperature. A column graph (on the same chart) would be used for total rainfall. The x-axis contains the dates (or years) in ascending order.
 - These charts should show any trends. In terms of evidence of global warming, Fred would likely expect to see an upward temperature trend and perhaps a downward rainfall trend – the yearly chart would likely highlight any trends.
 - This analysis would also need to consider natural fluctuations that are unrelated to global warming.
 - Fred could also extrapolate his results in an attempt to make predictions about future temp and rainfall in his area.
- (d) Bias is the effect of someone's opinion on the information output from the system. This can occur with the original collected data so that the data collected does not reflect the true original value. For instance, it could be that Fred was sure that he heard it rain in the night, and yet when he went out to check his rain gauge in the morning it registered zero. Bias would occur if he then registered a positive reading for rain for that morning to support his strong feeling that it had rained. Bias could also be present within the collection itself. For example, perhaps a tree that once shaded the measuring equipment has been dying and Fred has ignored this so that the temperature readings have increased. The data then incorrectly support or exaggerate the effects of global warming.

If Fred strongly believes in global warming then bias could affect Fred's analysis. If his initial analysis indicates no gradual warming and drying of the climate, but in fact points to a cooling in his temperature readings, he could be tempted to skew or even delete some readings that indicated cooler temperatures and high rainfall. For instance, he may remove some particularly cold days or days when there was high rainfall. He may even attempt to justify such changes by claiming faults in the thermometers or rain gauge.

Fred could also skew his analysis of the results by altering the scale or the detail included on his charts. This may make any temperature increases (or rainfall decreases) appear more significant than they actually are. For instance, it is likely there will have been some significantly warmer years in the past and cooler years more recently. Fred could choose to only chart each second year and he could select years such that years that do not support global warming are excluded. Also, if the charts indicate a small rise in temperature he could alter the temperature scale of the axis so that this small increase is exaggerated.



GROUP TASK Research

The Bureau of Meteorology (BOM) analyses temperature, rainfall and other climate data across Australia. Research and determine the data collected by the BOM for your location.



GROUP TASK Activity

It is possible to download historical weather statistics from the Bureau of Meteorology. Download and analyse the available data for your location to determine any evidence of global warming.

CHAPTER 5 REVIEW

1. Computers are used for analysing data because:
 - (A) after analysis the resulting information is always correct.
 - (B) they are able to accurately model and simulate real world systems.
 - (C) they can examine vast quantities of data with incredible speed and accuracy.
 - (D) they can make better and more intuitive decisions when the data is incomplete.
2. Which of the following statements is incorrect?
 - (A) RAM is volatile.
 - (B) Secondary storage is non-volatile.
 - (C) Increasing the amount of RAM is likely to improve performance.
 - (D) Increasing the amount of secondary storage is likely to improve performance.
3. Searching is simplified when:
 - (A) the data is numeric.
 - (B) the data is not sorted.
 - (C) the data is sorted.
 - (D) the data is text.
4. The logical statement “A=8 AND A>4” is:
 - (A) Always true.
 - (B) Never true.
 - (C) True when A>4
 - (D) True when A=8
5. Sorting 15, 12, 21, 2, 9 and 3 into ascending alphabetical order would result in:
 - (A) 2, 3, 9, 12, 15, 21
 - (B) 21, 15, 12, 9, 3, 2
 - (C) 12, 15, 21, 2, 3, 9
 - (D) 12, 15, 2, 21, 3, 9
6. If the individual pixel values within a full colour bitmap image were sorted into ascending numerical order then the resulting image:
 - (A) would not be affected.
 - (B) would be a random mix of colours.
 - (C) would be similar in appearance to a rainbow.
 - (D) could not be viewed as it has been corrupted.
7. Sorting is rarely performed on individual image, audio and video files because:
 - (A) each file is really a distinct data item.
 - (B) the order in which each element appears within the file is significant.
 - (C) sorting would alter the data.
 - (D) All of the above.
8. The process of determining the outputs from various sets of inputs is called:
 - (A) modelling.
 - (B) simulation.
 - (C) searching.
 - (D) what-if analysis.
9. Incorrect information derived from correct data is likely to be the result of:
 - (A) data incorrectly analysed.
 - (B) unauthorised analysis of data.
 - (C) linking databases incorrectly.
 - (D) Any of the above.
10. A business, without the knowledge of its employees, analyses the websites visited by each of its employees. This is an example of:
 - (A) data incorrectly analysed.
 - (B) unauthorised analysis of data.
 - (C) linking databases incorrectly.
 - (D) All of the above.
11. Define each of the following terms and describe their role in the analysing information process.
 - (a) Secondary storage
 - (b) primary storage
 - (c) CPU
12. List and describe four different types of graph.
13. List and describe reasons why most records are maintained using computer-based databases rather than manual filing systems.
14. An inventor has created a computer-based model of a new ergonomic keyboard design. The inventor is now working on the creation of a real prototype. Discuss reasons why the inventor would build a real prototype.
15. Explain how searching is likely to be used within each of the following processes:
 - (a) Reducing the total number of colours within a bitmap image.
 - (b) Removing unwanted background noise from a sampled audio file.
 - (c) The Board of Studies calculating the total number of students enrolled in Year 11 IPT.

In this chapter you will learn to:

- document the storage and retrieval process in an information system
- describe the characteristics and operation of hardware devices used for storage and retrieval
- use a range of hardware devices and associated software to store and retrieve information and data
- store and retrieve data using a network
- compare different file formats for storing the same data, explaining the features and benefits of each
- use software features to secure stored data and information
- retrieve and use data in an ethical way

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Storing and retrieving – the two-step process by which data or information can be saved and reloaded to allow for:

- other processing to take place
- a temporary halt in the system
- backup and recovery
- the transfer of data or information

Hardware for storing and retrieving

- hardware secondary storage devices, including:
 - magnetic disks
 - optical disks
 - network storages
 - flash memory
 - magnetic tapes
- the characteristics of hardware, including:
 - random or sequential access
 - volatile or non-volatile
 - permanent or non-permanent
- the trend to faster and greater storage capacity over time

Software in storing and retrieving

- hardware interface software
- file management software
- database management systems
- file formats for different data types
- Internet browser
 - used to access a machine independent data store
 - using search engines to access data
- encryption/password protection
- security of stored data whether stored centrally or distributed

Non-computer tools, including:

- paper based storage systems
- microfiche
- libraries

Social and ethical issues, including:

- the security of stored data
- unauthorised retrieval of data
- advances in storage and retrieval technologies and new uses such as data matching

TOOLS FOR INFORMATION PROCESSES: STORING AND RETRIEVING

Storing and retrieving is a two-part process; storing saves data or information and retrieving reloads data or information. Storing and retrieving supports all other information processes; it provides a mechanism for maintaining data and information prior to and after other information processes. The actual data or information is unchanged by storing and retrieving processes, rather the physical method of representing the data changes. For example, when saving data on a hard disk the storing information process physically represents the data using magnetic fields; when this data is later reloaded the retrieval process converts these magnetic fields into varying electrical signals that can be used by other hardware devices; in particular the CPU.

The CPU can only process a limited amount of data at any one time; consequently it is necessary to maintain data and information both before and after processing. The CPU stores and retrieves data directly from primary storage; primarily RAM. However primary storage is volatile and non-permanent; to permanently store data requires secondary storage. As was discussed early in Chapter 5, data is retrieved from secondary storage into primary storage in preparation for processing by the CPU. Once the data has been processed it is returned to primary storage, and finally is stored on secondary storage.

Secondary storage is non-volatile; it does not require power, and is used to maintain a more permanent copy of the data or information. In this chapter we concentrate on the storage and retrieval of data to and from secondary storage; in particular we consider:

- the role of storing and retrieving within information systems,
- characteristics of storage hardware,
- the operation of common examples of secondary storage devices,
- software used for storing and retrieving,
- non-computer storage systems and
- social and ethical issues associated with storing and retrieving.

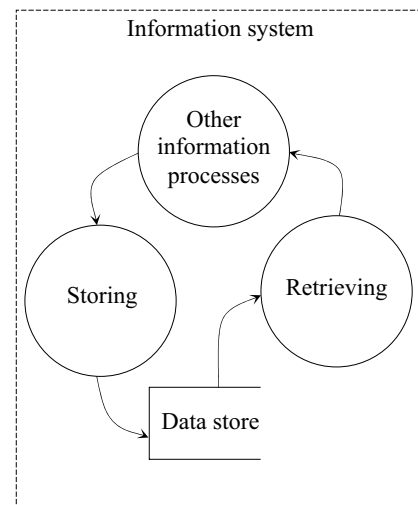


Fig 6.1

Storing saves data/information and retrieving reloads data/information.



GROUP TASK Research

There are many different types of RAM available. Use the Internet to research and classify various types of RAM according to their data access speed, storage capacity and cost.

THE ROLE OF STORING AND RETRIEVING

Storing and retrieving is about preserving data, it allows data to be reused at a later time. What are the reasons for wishing to preserve data? Let us consider some common answers to this question.

TO ALLOW OTHER PROCESSING TO TAKE PLACE

Secondary storage has a much greater capacity than primary storage; furthermore it is not uncommon for files to be larger than the total capacity of primary storage. Consequently data, and even program instructions, must be retrieved from secondary storage as required by the process occurring at that particular time. Similarly once primary storage is full then data that has been processed is saved to secondary storage. In essence secondary storage is being used as an extension of primary storage; when such a system is formally implemented within an operating system the portion of secondary storage used as RAM is called virtual RAM. Fig 6.2 shows a dialogue from Windows XP where the amount of virtual memory or virtual RAM is specified.

The situation becomes even more critical when various different processes are occurring at the same time; each process having its own data needs. For example, as I write these words my computer is running Microsoft Word, Internet Explorer, Microsoft Outlook, the Windows XP operating system together with various other software utilities for networking, virus detection, scanning, multimedia and faxing. Each of these processes uses data that is being swapped between secondary storage and primary storage as the need arises. It is the operating system's job to ensure each process is delivered the appropriate data and instructions at the required time.

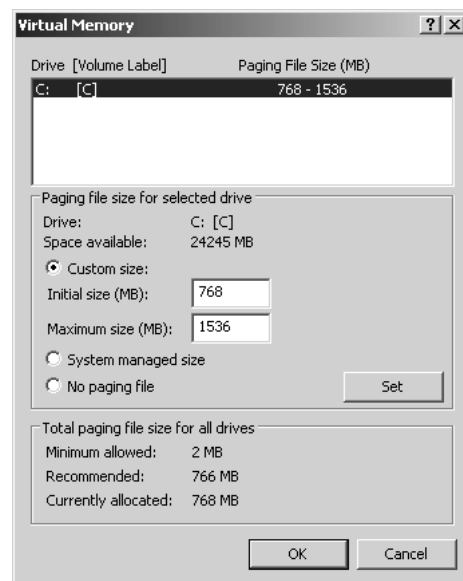


Fig 6.2
The amount of virtual memory or virtual RAM can be specified in Windows XP.



Consider the following:

The primary purpose of a file server is to store and retrieve data for a number of computers within a network. For this to occur all data must pass through the file server's primary storage (RAM) on its way to the network and then again on its way back to the file server's secondary storage.



GROUP TASK Modelling and discussion

Construct a diagram to describe the flow of data in the above discussion. Why is it necessary for the data to move through the file server's RAM? Discuss.



GROUP TASK Discussion

File servers commonly receive simultaneous requests from different computers to either store or retrieve data. These requests appear to be processed simultaneously. How is this possible? Discuss.

TO ALLOW FOR A TEMPORARY HALT IN THE SYSTEM

It is uncommon for all information processes present in an information system to be completed in a single session. As a consequence provision must be made to halt the operation of the system for a period of time; obviously this requires all data to be permanently stored. A simple example would be a student completing an assignment. The assignment is unlikely to be completed in a single session; therefore the student saves the assignment, halts the system and then at some latter time reloads the saved data to complete the assignment. In this example, the collecting information process is interrupted for a period of time; this is commonly the case for most collecting processes. For example, an ordering function within an information system is activated each time a new order arrives, entering the order being a collecting information process. Between entering orders the computer is used for various other processes, hence orders must be stored to allow for a temporary halt in the ordering system. Furthermore, the collected data must be stored if it is to be used at a latter time by various analysing, processing and displaying information processes.



Consider the following:

The systems flowchart in *Fig 6.3* at right describes the logic and flow of data for an information system used to process the results of an assessment task.

It is not necessary to understand the meaning of each symbol on this diagram; systems flowcharts are not specified within the current IPT syllabus. However an understanding of the processing taking place is necessary. Firstly the teacher marks the assessment task. These marks are entered by hand into the teacher's mark book. At a later time the marks are entered into the computer where they are stored in the school database. At the same time the student names are being retrieved from the school database. Once all the marks have been entered they are scaled and stored in the school database. Finally a printout of the results is generated and students are given their results.

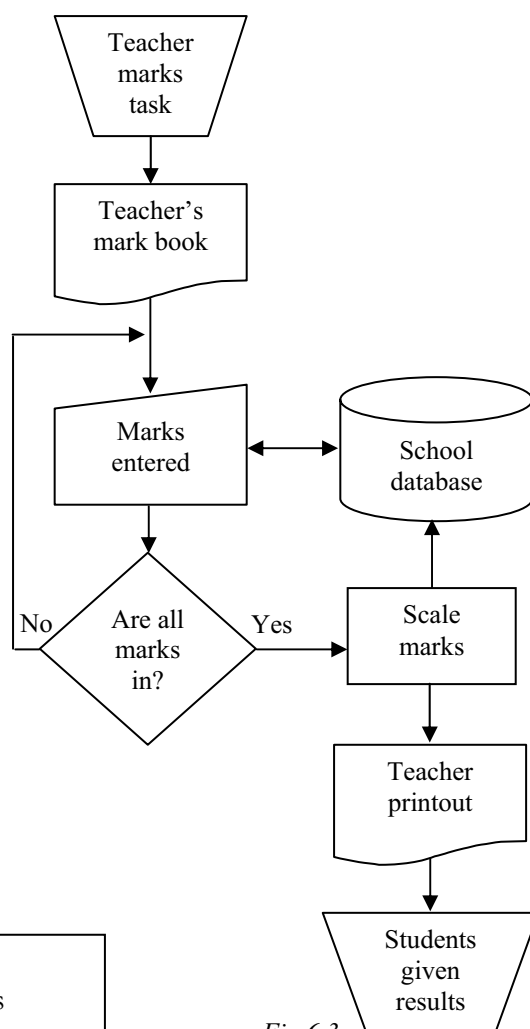


Fig 6.3
Systems flowchart for results
from an assessment task.



GROUP TASK Classify

Classify each information process occurring in the above scenario as one of the 7 syllabus information processes.



GROUP TASK Identify and Discuss

Identify times within the above scenario where a halt in the system is possible. How does the ability to halt the system at these times assist the operation of the information system? Discuss.

BACKUP AND RECOVERY

Making a backup of data is the process of storing or copying the data to another permanent storage device, commonly recordable CD, magnetic tape or a second hard disk. Recovery of data is the opposite process where the data is retrieved or restored from the backup copy and placed back into the system.

The aim of creating backups is to prevent data loss in the unfortunate event that the original data is damaged or lost. Such damage most often results from hard disk failures; in fact it is inevitable that all hard disks will eventually fail. Some other reasons for data loss includes software faults, theft, fire, viruses, intentional malicious damage and even intentional changes that are later found to be incorrect. For backup copies to most effectively guard against such occurrences requires backups to be made regularly and that these backup copies be kept in a fireproof safe or at a separate physical location.



Backup

To copy files to a separate secondary storage device as a precaution in case the first device fails or is lost.

Even the most reliable computer will eventually break down and the consequences can be devastating if no backups have been made. Consider a small business with 100 clients; a total loss of data means loss of all client records, orders and invoices, together with any correspondence and marketing materials. Even if much of this information is maintained in paper-based storage the cost of recovering from such a loss is enormous compared to the minor costs involved to maintain regular backups.

There are two main types of backup that are commonly used; full backups and partial backups. A full backup includes all files whereas a partial backup includes only those files that have been created or altered since the last backup was made. Most operating systems include an archive bit stored with each file to simplify partial backups; each time a file is created or altered the archive bit is set to true. Backup and recovery utilities examine this bit to determine files to be included in each partial backup. Incremental partial backups set each archive bit to false once each file has been copied, whilst differential partial backups copy the files but leave the archive bit set to true. A common backup strategy involves completing a full backup (which sets all archive bits to false), followed by a series of partial backups. If a failure occurs then the full backup is restored first. If incremental backups were made then each must be restored in the order they were made. If differential backups were made then, once the full backup has been restored, only the most recent differential backup needs to be restored as it contains all changes since the last full backup.

The frequency at which backups are made depends on how critical the data is to the organisation. Commonly a full backup is made once a week with an incremental backup made daily. A further safeguard against data loss is to rotate the media used for backups; commonly three complete sets being used. This means that should one set of backups become corrupted then the previous set can be used for data recovery. In addition, maintaining different sets of backups means the system can be restored back to many different historical points in time. This is useful for restoring data that was inadvertently changed and for returning to a point prior to corruption occurring, such as before a virus attack.



GROUP TASK Research

Research and document the backup strategy used at your work or your school. How long do these backups take to produce and is it necessary to halt the system to perform each backup?



GROUP TASK Discussion

RAID is a system that uses multiple hard disks to store data. Should one disk fail then the other disks include sufficient data to not only rebuild the lost disk but to continue system operation. Do you think such a system removes the need to make regular backups? Discuss.

TO ASSIST THE TRANSFER OF DATA/INFORMATION

When we view a web page, receive an email or access data across a network we are, amongst other things, retrieving files from a storage device on a remote computer. The data or information on the remote computer must be stored before it can be retrieved and transferred to other computers. Furthermore, the data, once received by the local computer, must be stored locally prior to further processing and display. Hence the storing and retrieving information process is integral to the transmitting and receiving process.

There are software applications, in particular database applications, operating across networks where transferred data is stored within RAM on the receiving computer, however in general, most data received is stored locally as a file within secondary storage. For example, web browsers store a copy of every file retrieved from a website locally within a temporary Internet files folder on the hard disk; the browser retrieves these files from this folder prior to display. *Fig 6.4* shows the Parramatta Education Centre IPT page behind the contents of the temporary Internet files folder. In this screen shot the temporary Internet files folder was first cleared, hence each of the files shown is required to correctly view the web page shown. Similar storing and retrieving processes occur as an integral part of the transferring of most data across networks.

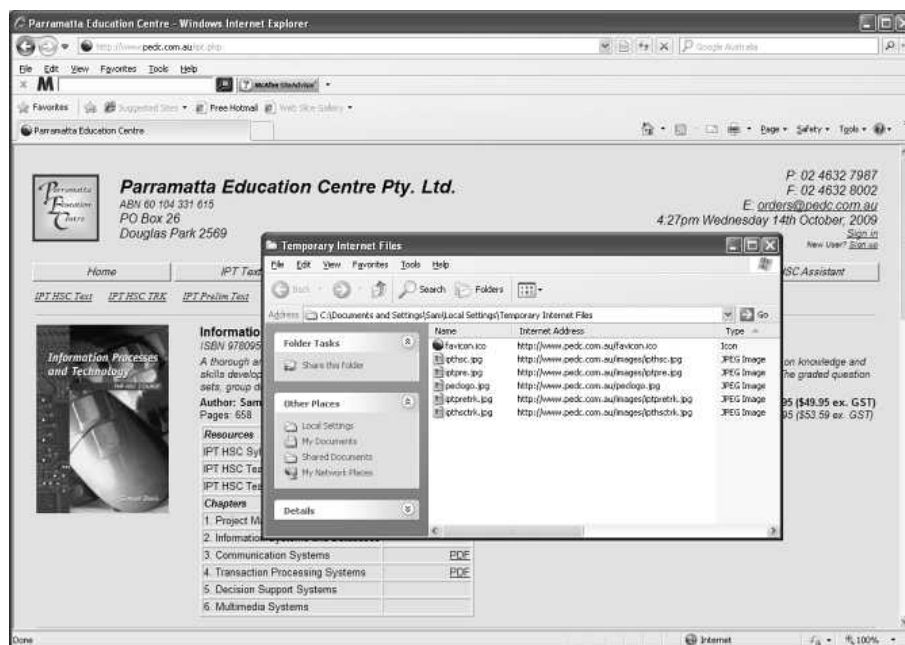


Fig 6.4

All files accessed using a web browser are transferred and stored locally.



GROUP TASK Discussion

When viewing a web page for the first time it often takes some time for all the images to appear, however on subsequent visits these same images appear virtually instantly. How can this be explained? Discuss.

HARDWARE IN STORING AND RETRIEVING

In this section we consider the characteristics and operation of a variety of commonly used secondary storage devices. Although our discussion is restricted to the operation of secondary storage devices it is important to remember that primary storage, such as RAM and ROM, is also hardware and that primary storage plays a vital role in the storing and retrieving information process.



GROUP TASK Discussion

RAM is certainly storage hardware, however it is integral to all seven of the information processes. Briefly discuss how RAM is used in each of the seven information processes.

Before we commence our examination of particular hardware storage devices let us discuss the meaning of some terminology commonly used to describe characteristics of such devices.

RANDOM OR SEQUENTIAL ACCESS

Random access refers to the ability to go to any data item in any order. Once the location of the required data is known then that data can be read or written directly without accessing or affecting any other data. The word random is used because the data can be accessed in any order, however in reality accessing any data item at random is unheard of; an equivalent, and perhaps more accurate term is direct access.

Sequential access means the data must be stored or retrieved in a linear sequence. For example, in *Fig 6.5* the sixth data item is needed so the preceding five data items must first be accessed. In terms of hardware devices, tape drives are the only widely used sequential storage devices. The time taken to locate data makes sequential storage unsuitable for most applications apart from backup.

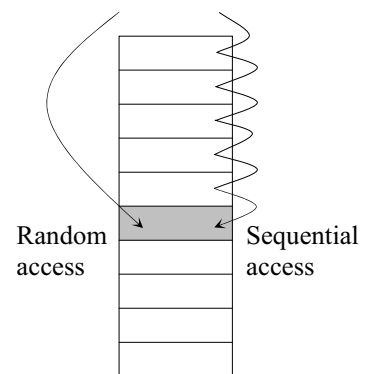


*Fig 6.5
Random access versus
sequential access.*

Do not confuse random access files and sequential files with random access and sequential access storage devices. Random and sequential when used in relation to files describe the way software applications access files, when used in regard to storage devices these same terms relate to the physical storage of the data. Both types of file can be stored on either device; however a random access file stored on a sequential access device must physically be read and written sequentially, similarly a sequential file stored on a random access device despite being able to be physically accessed randomly will be read and written sequentially.



Consider the following:

Commodore released the Personal Electronic Transactor (PET) in 1977. The original PET used standard audio cassette tapes as its sole secondary storage medium, and had a massive 4K of RAM!



GROUP TASK Research

Use the Internet to research how data was stored on early personal computers that used audio cassette tapes.



*Fig 6.6
The Commodore PET 2001
released in 1977.*

VOLATILE OR NON-VOLATILE

Volatile computer storage requires a continuous electrical current to maintain data; if no electrical current is present then the data will be lost. On almost all computers RAM is volatile, if you do not save your data to secondary storage it is lost from RAM should a power failure occur. Dynamic RAM (DRAM) chips are particularly volatile; each storage area on these chips must be refreshed regularly to maintain their data whereas static RAM (SRAM) chips merely require electrical current to be present.

To reduce the effects of the volatility of RAM computers performing critical tasks are connected to uninterruptible power supplies (UPSs). The purpose of a UPS is to provide sufficient power to allow the contents of RAM to be written to secondary storage and then for the computer to be shutdown gracefully. At the time of writing non-volatile RAM chips had just been developed; currently such chips are used in specialised applications, however it is likely that eventually they will become part of all computer systems.

Predictably non-volatile storage does not require power to maintain stored data. Virtually all types of storage, apart from RAM, can be classified as non-volatile. Examples include ROM, magnetic disks and tapes, all types of optical storage and even flash memory.

PERMANENT OR NON-PERMANENT

No storage device is totally permanent; in reality there are only degrees of permanence. The meaning of the terms permanent and non-permanent largely depends upon the context in which they are used. Let us consider common uses of the terms permanent and non-permanent as they apply in different contexts.

Volatile memory such as RAM is certainly less permanent than any of the non-volatile forms of storage. Hence when comparing RAM with secondary storage it is common and appropriate to classify RAM as non-permanent and secondary storage as permanent.

When comparing different secondary storage devices permanence can be used to imply the inability to alter or erase data. Consider data stored on a hard disk, it can easily be altered or even erased, hence hard disks can be described as non-permanent. On the other hand the data on a non-recordable DVD or CD-ROM can be described as permanent; it cannot be altered or erased.

Another common use of the term permanent is in regard to archived copies of data. Commonly businesses make a complete copy of their financial records at the end of each financial year. This copy is placed into permanent storage, perhaps in a safe or even in a safety deposit box within their bank's safe. In this context it is not the medium on which the data is stored that determines permanence rather the term permanent describes the purpose of maintaining the secure copy.

A further common use is applied to backup copies of data, particularly in regard to networks. Files or complete storage devices that are included within regular backups are said to be permanent whereas data not included in such backups is said to be semi-permanent or even non-permanent.



GROUP TASK Discussion

It is often said that volatile storage devices are non-permanent and non-volatile storage devices are permanent. Do you agree? Discuss.

SET 6A

1. The CPU stores and retrieves data directly to and from:
 - (A) secondary storage.
 - (B) primary storage.
 - (C) non-volatile storage.
 - (D) permanent storage.
2. Virtual memory is used:
 - (A) when there is insufficient secondary storage.
 - (B) to remove the need to retrieve data from secondary storage.
 - (C) when the amount of RAM is insufficient.
 - (D) to speed up the processing of data.
3. What is the purpose of secondary storage?
 - (A) To allow other processes to take place.
 - (B) To allow for the system to be halted.
 - (C) To assist the transfer of information.
 - (D) All of the above.
4. Incremental backups are performed to:
 - (A) ensure a complete copy of the data is maintained should a problem occur.
 - (B) reduce the time used to perform backups.
 - (C) ensure multiple backups are maintained.
 - (D) secure data against unauthorised access.
5. Storage that requires power to maintain its contents is best described as:
 - (A) volatile storage.
 - (B) non-permanent storage.
 - (C) non-volatile storage.
 - (D) permanent storage.
6. The aim of creating backups is to:
 - (A) prevent unauthorised access.
 - (B) detect incorrect data.
 - (C) protect against data loss.
 - (D) remove the need for users to save their work.
7. When viewing a web page:
 - (A) no data is stored locally, all the data remains on the remote web server.
 - (B) all files used by the page are stored in RAM on the local computer.
 - (C) each file needed to view the page is first stored locally in secondary storage.
 - (D) the files within the page are sent directly to the display hardware.
8. Which of the following is NOT a characteristic of sequential storage?
 - (A) To retrieve a data item requires retrieval of each preceding data item.
 - (B) Individual data items can be retrieved from any physical part of the media without accessing any other data.
 - (C) Tape is the only widely used sequential storage media.
 - (D) For most applications sequential retrieval of data is slower than direct or random retrieval.
9. If the archive bit for a file is set to true then:
 - (A) the file will be included in an incremental backup but not in a full backup.
 - (B) the file will be included in a full backup but not in an incremental backup.
 - (C) the file will be included in both an incremental backup and a full backup.
 - (D) the file will not be included in any backups.
10. Which of the following is NOT true of the storing and retrieving information process?
 - (A) It supports all other information processes.
 - (B) It alters the actual data or information.
 - (C) It allows data to be reused
 - (D) It maintains data prior to and after other processes.

11. Storing and retrieving assists each of the other information processes.
Explain how storing and retrieving assists the
 - collecting,
 - organising, and
 - analysing information processes.Include examples within each of your explanations.
12. List and describe the differences between primary and secondary storage.
13. For each of the following, compare and contrast the meaning of the terms:
 - (a) Volatile and non-volatile
 - (b) Permanent and non-permanent
 - (c) Random and sequential
14. Commonly commercial software is installed from CD-ROM. The installation involves various information processes.
 - List the sequence of information processes that would typically occur.
 - For each step in your sequence identify the hardware devices being used.
15. A small business receives on average 15 orders per day. These orders are processed as they are received using a commercial software package. The computer used to process the orders is also used for email, web access, word processing and various other administrative tasks.
Recommend and justify an appropriate backup strategy.

OPERATION OF SECONDARY STORAGE HARDWARE

In this section we consider the operation and characteristics of magnetic storage, both disks and tape; optical storage such as various CD and DVD based technologies; network storage and finally flash memory technologies. Each of these secondary storage technologies is used to store and retrieve digital data in a non-volatile form.

MAGNETIC STORAGE

To understand the underlying operation of magnetic storage devices requires a basic knowledge of certain magnetic principles:

1. Magnets exert forces on each other known as magnetic fields. Such forces move from the north to the south pole of the magnet.
2. Magnetic fields are greatest at the poles.
3. Electrical currents produce magnetic fields.
4. There are only a few elements, primarily iron, cobalt and nickel, which can be magnetised. Materials that include these elements and that can be magnetised are known as ferromagnetic materials.
5. Different ferromagnetic materials behave differently when placed in a magnetic field.
 - A. Some materials are easily magnetised by weak magnetic fields but when the field is turned off they quickly demagnetise; these materials are known as soft magnetic materials and are used during the process of storing or writing data.
 - B. Some soft magnetic materials conduct electricity well when in the presence of a magnetic field but are poor electrical conductors when not. This phenomenon is called the magneto-resistance (MR) effect. MR materials are used during the process of retrieving or reading data.
 - C. Some materials require a strong magnetic field to become magnetised however they retain their magnetisation when the magnetic field is turned off. These materials are known as hard magnetic materials and are used to produce permanent magnets. Such materials are the basis of magnetic storage media.

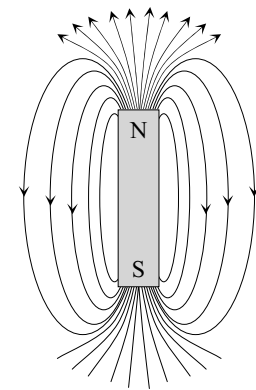


Fig 6.7

Magnetic forces move from north to south poles and are greatest at the poles.

To further assist our discussion let us first examine a microscopic detail of a typical piece of magnetic storage medium that already contains stored data (see Fig 6.8 at right). This detail could be a section of a floppy disk, a hard disk platter or even a piece of magnetic tape; in each case hard magnetic material is used and the storage principles are the same.

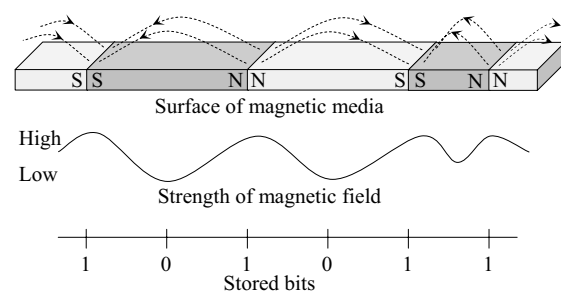


Fig 6.8

Microscopic detail of magnetic storage medium.

Digital data is composed of a sequence of binary digits, zeros and ones. These zeros and ones are equally spaced along the surface of the magnetic medium. High magnetic forces are present where the direction of the magnetic field changes; these points are really magnetic poles. It is the strength of the magnetic force that determines a one or a zero, not the direction of the magnetic force. Low magnetic forces occur between two poles and represent zeros. High magnetic forces are present at the poles and represent ones.



Consider the following:

At the time of writing (2009) the number of bits stored per inch (BPI) on the surface of a hard disk ranges up to around 1,000,000 BPI at the centre of each disk platter; this measure is commonly called linear density. This means a track on a hard disk can store some 40000 bits per millimetre. If *Fig 6.8* is the surface of a hard disk platter then the real width of the medium depicted would be approximately 1.5 ten thousandths of a millimetre; rather too small to print! Currently magnetic tape is available with a linear density of around 100,000 BPI resulting in some 4000 bits per millimetre.



GROUP TASK Research

Investigate the linear density of various hard disks and magnetic tapes. During your research determine the relationship between linear density and areal density.

Storing or writing magnetic data

Magnetic data is written on to hard magnetic material using tiny electromagnets. These electromagnets form the write heads for all types of magnetic storage devices. Essentially an electromagnet is comprised of a copper coil of wire wrapped around soft magnetic material (see *Fig 6.9*). The soft magnetic material is in the shape of a loop that is not quite joined; this tiny gap in the loop is where the magnetic field is produced and the writing takes place.

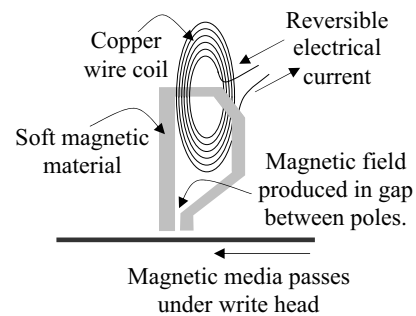


Fig 6.9

Detail of magnetic write head.

When an electrical current is present in the coil the enclosed soft magnetic material becomes magnetised, one end of the material becoming a north pole and the other a south pole. Hence a magnetic field is produced flowing from the north to the south. If the direction of the current through the coil is reversed then the direction of the magnetic field produced is also reversed. The magnetic field is strong enough for the hard magnetic material on the medium to be magnetised. A binary one is represented each time the direction of the magnetic field changes as a consequence of reversing the current into the coil. Zeros are represented when the direction of the current flow is constant and hence the direction of the magnetic field remains constant.

Retrieving or reading magnetic data

MR materials are the basis of most modern read heads; commonly this material contains around 80 percent nickel and 20 percent iron. Such materials are particularly sensitive to small changes in magnetic forces when a constant current is flowing through the material; that is they alter their resistance more noticeably. When stronger magnetic forces are detected, representing a 1, the current flow through the MR material increases and hence the voltage increases; similarly when the force is weaker the current and voltage decreases. These voltage fluctuations reflect the original binary data and are suitable for further processing by the computer.

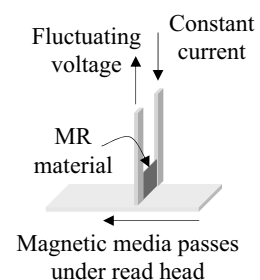


Fig 6.10

Detail of an MR read head.

MAGNETIC HARD DISKS

Hard disks store data magnetically on precision aluminium or glass platters. The platters have a layer of hard magnetic material (primarily composed of iron oxide) into which the magnetic data is stored. On top of this material is a layer of carbon and then a fine coating of lubricant. The carbon and lubricant layers improve the durability of the disk and slow down corrosion of the magnetic layer. Each platter is double sided, so two read/write heads are required for each platter contained within the drive's casing. At the time of writing most drives contain two to five double-sided platters requiring four to ten read/write heads. The casing is sealed to protect the platters and heads from dust and humidity.

Data is arranged on each platter into tracks and sectors. The tracks are laid down as a series of concentric circles. At the time of writing a typical platter contains some one hundred thousand tracks with each track split into hundreds of sectors. The diagram in *Fig 6.11* implies an equal number of sectors per track; on old hard disks this was true however on newer hard disks this is not the case, rather the number of sectors increases as the radius of the tracks increase. Each sector stores the same amount of data, in most cases 512 bytes. The read/write heads store and retrieve data from complete sectors.

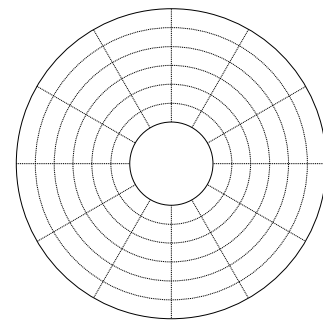


Fig 6.11
Each disk platter is arranged into tracks and sectors.

There are two motors within each hard drive; a spindle motor to spin the platters and an actuator assembly to move the read/write heads into position. The spindle motor operates at a constant speed; commonly from around 5,000 to 15,000 revolutions per minute. Whilst this is occurring the read/write head is moved in and out by the actuator assembly to locate the heads precisely over the required sectors on the disk platters.

Each read/write head is attached to a head arm with all the head arms attached to a single pivot point, consequently all the read/write heads move together. This means just a single read/write head on a single platter is actually operational at any instant. Each read/write head is extremely small, so small it is difficult to see with the naked eye. What is usually seen is the slider that houses the head. The air pressure created by the spinning platters causes the sliders to float a few nanometers (billionths of a metre) above the surface of the disk.

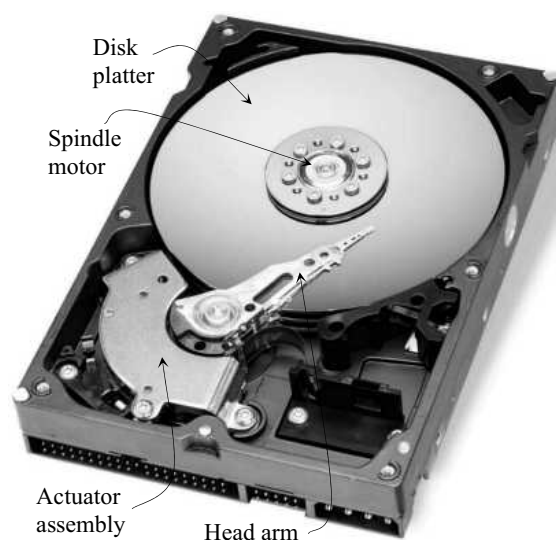


Fig 6.12
Internal view of a hard disk drive.

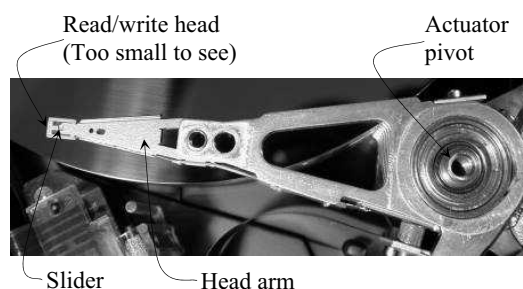


Fig 6.13
Expanded view of a head arm assembly.

Sophisticated circuits are required to control the accurate performance of the drive; in fact the processing power contained within a modern hard disk drive far exceeds the power of computers produced during the 1980s, furthermore they contain similar amounts of RAM in the form of cache. Hard drive circuits control the operation of the motors, communication with the CPU as well as checking on the accuracy of each read or write operation. Most hard disks contain their own built-in cache to significantly speed up access times. Data on sectors near the requested data is read into cache; commonly such data is subsequently required, consequently it can be accessed much faster from cache.

Because the operation of a hard drive involves mechanical operations they will never reach the speeds possible with chip based storage technologies. Hard disks provide an economical means of permanently storing vast quantities of data. At the time of writing 250GB hard drives were common and drives exceeding 1TB were readily available. Currently, with the assistance of cache, hard drives are able to store and retrieve data at speeds exceeding 100MB per second.

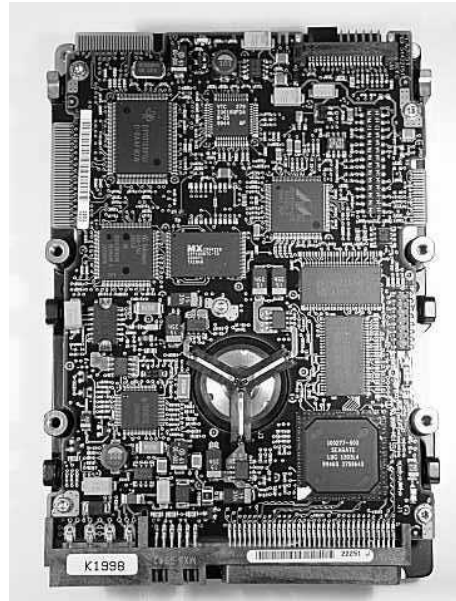


Fig 6.14

Underside of a hard disk drive showing the circuit board containing processing and cache chips.



Consider the following:

Older hard disk drives used the track (or cylinder) number, head number and sector number to determine the address of each sector (or block) of data. These addresses, known as CHS addresses, were translated via the computers BIOS (Basic Input Output System). Unfortunately such a system limited the size of hard disks to 1024 cylinders, 255 heads and 63 sectors per track equating to a capacity of 8.4GB.

As newer higher capacity hard drives became available and variable sectors were present on each track a new addressing system known as LBA (Logical Block Addressing) was introduced; this system essentially bypasses the computers BIOS altogether. LBA assigns each block (or sector) of data a unique sequential number; for example a drive with a total of 490,350,672 sectors would use LBA addresses from 0 to 490,350,671. The circuits within the hard drive translate the LBA address into the required physical address on the disks.



GROUP TASK Activity

Explain how 1024 cylinders, 255 heads and 63 sectors per track equates to a storage capacity of 8.4Gb?



GROUP TASK Research

Research specifications with regard to currently available hard disk drives. Determine the storage capacity, claimed data transfer rate, number of platters, total number of sectors and the storage size of each sector.

MAGNETIC TAPE

Magnetic tape has been used consistently for data storage since the early 1950s; the first such device being released commercially in 1952 by IBM (see Fig 6.15). At this early stage magnetic tape was the principal secondary storage technology; hard disk technologies first appeared in the late 1950s. The IBM 726 pictured featured six data tracks running parallel to the length of the tape, a seventh track was used for error checking. The linear density was around 100 bits per inch with a read/write speed of approximately 12,500 bits per second; current high performance magnetic tapes have linear densities exceeding 100,000 bits per inch and read/write speeds of more than 100 megabytes per second. These early devices were based on audio tape technologies; this has remained a common trend, many of today's tape drives borrow many of their components from audio or video tape drives.



Fig 6.15
The IBM 726 magnetic tape drive released in 1952.

Today magnetic tape is contained within cassettes or cartridges. Such cartridges range in size from roughly the size of matchbox to the size of a standard VHS tape. Tape is currently the most convenient and cost effective media for backup of large quantities of data. A single inexpensive magnetic tape can store the complete contents of virtually any hard disk; currently magnetic tapes (and tape drives) are available that can store up to 1TB of compressed data at only a few cents per gigabyte. The ability to backup the entire contents of a hard disk using just one tape far out way the disadvantages of sequential access; both backup and restore procedures are essentially sequential processes.



Fig 6.16
Various types of magnetic tape cartridges.

There are two different technologies currently used to store data on magnetic tape, helical and linear. Helical tape drives use technology originally developed for video and audio tapes; in fact the majority of the components, often including the actual tape cartridges, are borrowed directly from camcorders. Linear tape technologies were designed specifically for archiving data; hence in terms of data storage most linear systems perform their task more efficiently than helical systems.

Helical Technology

Helical systems contain a relatively large drum containing two pairs of read and write heads, each pair operates in isolation to the other. The tracks written by each pair of read/write heads cross each other at an angle of 40 degrees. Where two tracks intersect the magnetic forces combine. As both original forces are of equal strength, the direction of the combined magnetic force always remains closest to the direction of both the original forces and further away from either of their opposite forces. In Fig 6.17 the original forces are shown with open arrow heads and the combined force with a closed arrow head; the length of the arrows (or vectors) is an indication of the strength of each force. The result of these criss-cross tracks is a doubling of the linear density of the tape.

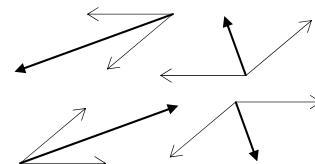


Fig 6.17
Original and combined force possibilities at the intersection of two tracks.

The drum, containing the read/write heads, is tilted slightly and rotates at high speed (commonly around 2,000 rpm). The tape is wrapped halfway around the drum and moves slowly (about 2cm per second) in the opposite direction to the rotation of the drum. As the drum is tilted, the tape contains relatively short diagonal data tracks; each track storing some 128 kilobytes of data. *Fig 6.18* shows the data tracks for one set of read and write heads. During a typical writing or storing process the write head stores the data on a complete track, as this track arrives at the corresponding read head the data is retrieved and verified to ensure correctness.

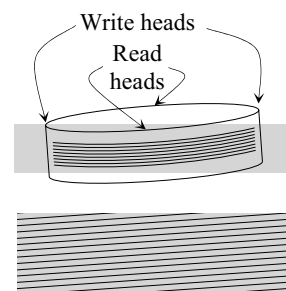


Fig 6.18
Detail of read/write drum and diagonal data tracks.

Helical tape systems work well for domestic video and audio applications where the data is rarely rewritten and even data retrieval is relatively uncommon; you can only watch Grandpa's home video so many times! Because of the enormous digital camcorder market such systems are economical to produce, however they are not designed for the multiple and intensive demands of large corporate organisations performing massive and regular data backup processes. Tape wear is the major problem. Slow moving tape is in contact with a rapidly spinning read/write drum causing friction and subsequent wear. Also the tape must negotiate a maze of posts and rollers in order for it to wrap around the rotating drum; wear occurs as the tape contacts each roller and also as it flexes to negotiate the maze. *Fig 6.19* below shows the tape path for a helical tape compared to two different types of linear tape system.

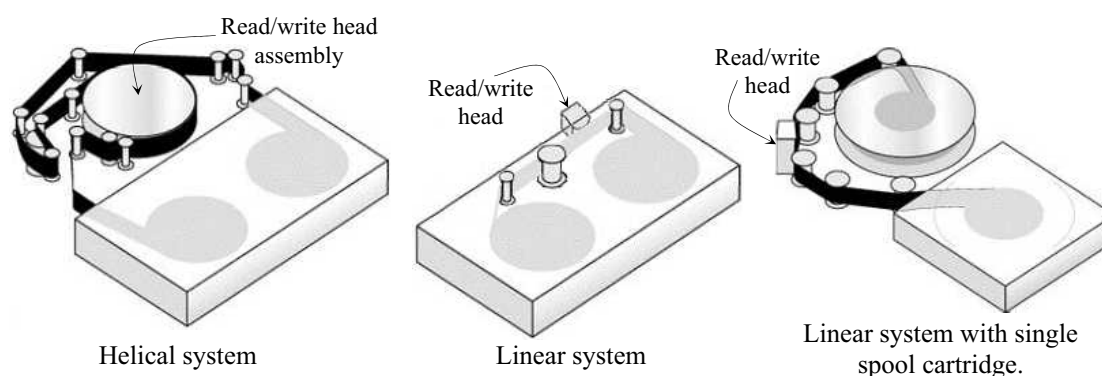


Fig 6.19
Detail of tape paths for helical and linear tape systems.

Consider the complexity of the helical system depicted above. When a tape is loaded this system must first extract the tape from the cartridge and wind it around a maze of posts and rollers using a complex system of motors and pulleys. In contrast, no such components are needed for the linear system, as the tape never leaves the cartridge. These motors and pulleys add an enormous number of mechanical parts and mechanical parts are most prone to failure.

The single spool linear system depicted at right in *Fig 6.19* aims to maximise the amount of tape within the cartridge by using a permanent spool within the tape drive. Notice that there are no acute angles in the tape path and furthermore the data side of the tape never contacts any rollers. Contrast this with the helical system, the data side of the tape contacts many rollers during its path through the drive.



GROUP TASK Investigation

Determine the type of tape backup system used by either your school or work. What is the capacity of a single tape used by this system?

Linear Technology

Linear systems read and write continuous tracks parallel to the length of the tape. Each set of read/write heads contains two read heads with a write head in between; this allows data to be written and verified in either direction (see Fig 6.20). A complete track is written by each set of read/write heads, when the end of the tape is encountered the tape reverses and the whole read/write assembly moves up or down slightly in order to write the next track in the opposite direction.

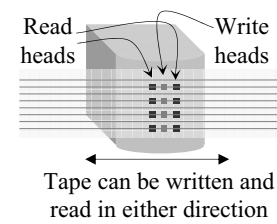


Fig 6.20
Detail of read/write assembly
and parallel data tracks.

There are a large variety of different types of linear tape drive; some contain just a single set of read/write heads whilst others contain as many as eight sets of heads. The tape cartridges and actual tape are similarly diverse. Some systems write just 8 parallel tracks whilst others write many hundred of parallel tracks. The cartridges are broadly of two types; Fig 6.19 on the previous page depicts these two types. Traditional double spool cartridges are used in lower capacity systems and single spool cartridges for higher capacity systems.



Consider the following:

An example of the currently popular Quantum LT0-4 HH is shown in Fig 6.21 at right; this linear tape drive uses lasers and optical marks on the reverse side of the tape to accurately align the read/write heads. The model shown is able to store 800GB (1.2TB compressed) on each tape cartridge.



Fig 6.21
The Quantum LTO-4 HH linear tape drive.

Fig 6.22 shows a Super AIT (SAIT) drive and cartridge produced by Sony. The system uses helical scan technology and single spool cartridges. Each cartridge contains its own memory chip where details of the contents of the tape are stored. A single cartridge can store up to 500GB (1.3TB compressed).



Fig 6.22
Sony's SAIT helical scan tape drive.



GROUP TASK Discussion

Discuss how lasers combined with optical marks and memory chips on tape cartridges help to increase the performance of these drives.



GROUP TASK Research

Using the Internet, or otherwise, find out the capacity and read/write speed of currently available tape drives. Can you determine whether each drive is based on helical or linear technology?

OPTICAL STORAGE

Optical storage processes are based on reflection of light; either the light reflects well or it reflects poorly. It is the transition from good reflection to poor reflection or vice versa, that is used to represent a binary one (1); when reflection is constant a zero (0) is represented. This is similar to magnetic retrieval where a change in direction of the magnetic force represents a binary one and no change represents a zero. To illustrate optical storage imagine shining a torch across a busy highway at night, you would see the light reflected back as each vehicle passed through the beam of light; ones being represented each time a vehicle enters the beam and again as it leaves the beam (see Fig 6.23). If this data were recorded at precise intervals, say every hundredth of a second, the result would be a sequence of binary digits.

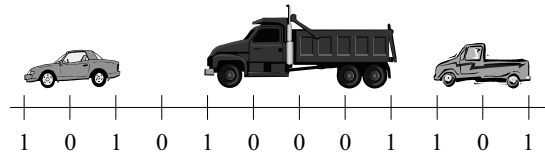


Fig 6.23
The transition between good and poor reflection is read as a binary one (1).

As the data is so tightly packed on both compact disks (CDs) and digital versatile disks (DVDs) it is essential that the light used for optical storage processes be as consistent as is possible; lasers provide such light. The word laser is really an acronym for “light amplification by stimulated emission of radiation”. Different types of atoms, when excited, give off radiation in the form of different types of light; under normal conditions the light is emitted in all directions, for example neon advertising signs. A laser controls this process by using particular atoms within a precisely controlled environment. Essentially a laser produces an intense parallel beam of light composed of electromagnetic waves that are all identical; accurately focussing this light produces just what is needed for optical storage and retrieval processes. Relatively weak lasers are used during the retrieval of data and much higher-powered lasers when storing data. Higher-powered lasers produce the heat necessary to alter the material used during the CD or DVD burning process; in fact similar lasers are used during the initial stages when manufacturing commercial CDs and DVDs.

Before we consider the detail of the optical storing and retrieving processes let us consider the nature of both CD and DVD media. CDs contain a single spiral track that commences at the inner portion of the disk and spirals outward toward the edge of the disk (see Fig 6.24). This single track is able to store up to 680 megabytes of data. DVDs contain similar but much more densely packed tracks, each track can store up to 4.7 gigabytes of data. Furthermore, DVDs may be double sided and they may also be dual layered. Therefore a double sided, dual layer DVD would contain a total of four spiral tracks; in total up to 17 gigabytes of data can be stored.

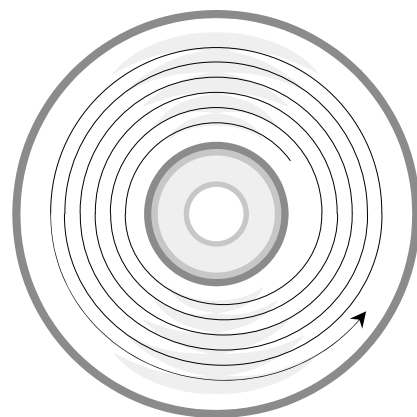


Fig 6.24
CDs and DVDs contain spiral tracks.



GROUP TASK Activity

An audio CD is able to store up to 74 minutes of stereo sound using 16bits per sample and 44,100 samples per second. Compare the capacity of an audio CD with the 680MB data capacity quoted above. Suggest possible reasons for any differences.

Each spiral track, whether on a CD or a DVD, is composed of a sequence of pits and lands. On commercially produced disks the pits really are physical indentations within the upper side of the disk. *Fig 6.25* depicts the underside of a disk, this is the side read by the laser, and hence the pits appear as raised bumps above the surrounding surface. On writeable media the pits are in fact not pits at all; rather they are areas that reflect light poorly; more on this when we discuss optical storing. The essential point is that pits reflect light poorly and lands reflect light well regardless of their physical structure.

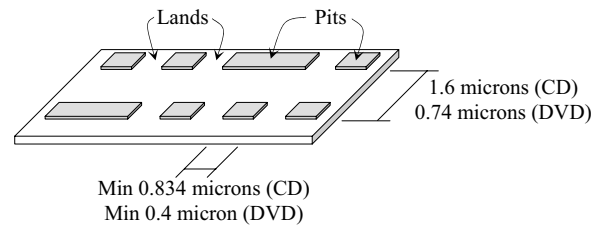


Fig 6.25

Magnified view of the underside of an optical disk.

The dimensions shown in *Fig 6.25* indicate an approximate 50 percent reduction in both track pitch and pit length for DVDs compared to CDs; these physical size differences account for about a four-fold increase in the storage capacity of DVDs compared to CDs. In reality, an almost seven-fold increase in capacity has occurred; the remaining increase is largely due to improvements in error correction techniques.

In *Fig 6.25* the measurements are expressed in microns, one micron is one millionth of a metre or one thousandth of a millimetre. As a consequence of these incredibly small distances the length of pits that would be needed when ones appear together or close together is so small that it is likely to cause read errors. Also tracking problems can occur when the pits or lands are too long, this would occur when a large number of zeros are in sequence. The solution is to avoid such bit patterns occurring in the first place. The eight to fourteen modulation (EFM) coding system is used; EFM converts each eight-bit byte into fourteen bits such that all the bit patterns include at least two but less than ten consecutive zeros. This avoids such problems occurring within a byte of data, but what about between bytes? For example, the two bytes 10001010 and 11011000 convert using the EFM coding system to 1001001000001 and 01001000010001. When placed together the transition between the two coded bytes is ...0101...; our rule of having at least two zeros is broken. To correct this problem two merge bits are placed between each coded byte; the value of these merge bits is chosen to maintain our at least two zeros but less than ten rule. Obviously once the data has been read the merge bits are ignored.

Both CDs and DVDs are approximately 1.2mm thick and are primarily clear polycarbonate plastic. On commercially produced disks the pits are stamped into the top surface of the plastic, which is then covered by a fine layer of reflective metal (commonly aluminium), followed

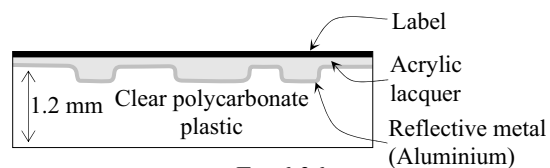


Fig 6.26

Cross section of a typical commercially produced CD or single-sided single layer DVD.

by a protective acrylic lacquer and finally some sort of printed label. On recordable and rewriteable media a further layer is added between the polycarbonate and the reflective layer; this is the layer whose reflective properties can be altered. It is actually quite difficult to damage a disk by scratching its underside, in contrast the label side of a disk is easily damaged; try scratching both sides of an old CD-R with a pen, you'll see what I mean. Double-sided DVDs are essentially two single-sided disks back to back. Double layer DVDs contain two data layers where the outside layer is semi reflective; this allows light to pass through to the lower layer. The laser is accurately focussed onto the layer currently being read.

Retrieving or reading optical data

Retrieving data from an optical disk can be split into two processes; spinning the disk as the read head assembly is moved in or out to the required data and actually reading the reflected light and translating it into an electrical signal representing the original sequence of bits. To structure our discussion we consider each of these processes separately, although in reality both occur at the same time.

- **Spinning the disk and moving the read head assembly**

To read data off an optical disk requires two motors, a spindle motor to spin the disk and another to move the laser in or out so that the required data passes above the laser. The spindle assembly contains the spindle motor together with a clamping system that ensures the disk rotates with minimal wobble. The read head assembly is mounted on a carriage, which moves in and out on a pair of rails. In modern optical drives the motor that moves the carriage responds to tracking information returned by the read head. This feedback allows the carriage to move relative to the actual location of the data track.

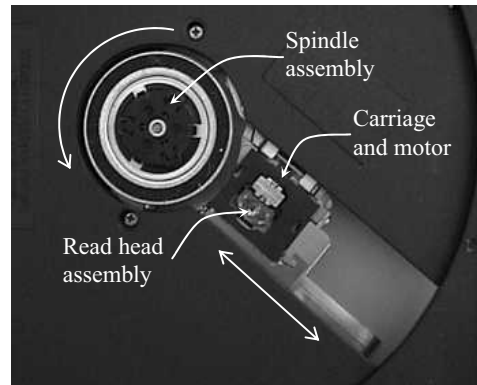


*Fig 6.27
Detail of a CD/DVD drive from a
laptop computer.*

At a constant number of revolutions per minute (rpm) the outside of a disk rotates much faster than the inside. Older CD drives, and in particular audio CD drives, reduce the speed of the spindle motor as the read head moves outwards and increase speed as the read head moves inwards. For example, a quad speed drive spins at 2120 rpm when reading the inner part of the track and at only 800 rpm when reading the outer part. The aim being to ensure approximately the same amount of data passes under the read head every second; drives based on this technology are known as CLV (constant linear velocity) drives.

Most CD and DVD drives manufactured since 1998 use a constant angular velocity (CAV) system, which simply means the spindle motor rotates at a steady speed. CLV technology is still used within most audio drives, which makes sense, as there really is no point retrieving such data at faster speeds. However for computer applications, such as installing software applications, faster retrieval is definitely an advantage. As a consequence of CAV, such drives have variable rates of data transfer. For example, a 24-speed CAV CD drive can retrieve some 1.8 megabytes per second at the centre and 3.6 megabytes per second at the outside. Quoted retrieval speeds for CAV drives are often misleading; for example a CAV drive designated as 48-speed can only retrieve data from the outside of a disk at 48 times that required for normal CD audio. These maximum speeds are rarely achieved as very few CDs have data stored on their outer edges.

Current CAV drives have spindle speeds in excess of 12000 rpm; faster than most hard disk drives. Such high speeds produce air turbulence resulting in vibration. When most drives are operating the noise produced by this turbulence can be clearly heard. Furthermore, the vibration is worst at the outside of the disk, just where the data passes under the read head at the fastest speed, hence read errors do occur. Such problems must be resolved if the ever increasing speed of optical data retrieval is to continue.



Consider the following:

How does CLV work? Essentially the speed of the spindle motor is controlled by the amount of data within the drive's temporary storage or buffer. When the amount of data in the buffer exceeds a certain threshold the motor is slowed and hence the buffer begins to empty. Similarly if the data in the buffer is less than a certain threshold then the motor speeds up. Unfortunately it takes time to speed up and slow down the spindle; this time becomes significant once rates of data transfer approach 16 times that required to read an audio CD (about 16 times 150 kilobytes per second or roughly 2 megabytes per second). This is the primary reason for the development and production of CAV drives.



GROUP TASK Discussion

Buffers are used primarily to assist the movement of data between hardware devices operating at different speeds. CAV optical drives also contain a buffer; discuss how such a buffer would operate during data retrieval from a complete track.

• Reading and translating reflected light into electrical signals

There are various different techniques used to create, focus and then collect and convert the reflected light into electrical signals. Our discussion concentrates on the most commonly used techniques.

Let us follow the path taken by the light as it leaves the laser, reflects off the pits and lands, and finally arrives at the opto-electrical cell (refer to *Fig 6.28*). Firstly, remember that lasers generate a single parallel beam. This beam passes through a diffraction grating whose purpose is to create two extra side beams; these side or tracking beams are used to ensure the main beam tracks accurately over the pits and lands. Unfortunately the diffraction grating causes dispersion of the beams. To correct this dispersion the three beams pass through a collimator lens; whose job is to make the beams parallel to each other. A final lens is used to precisely focus the beams on the reflective surface of the disk.

As the disk spins both tracking beams should return a constant amount of light, as they are reflecting off the smooth surface between tracks (see *Fig 6.29*). If this is not the case then the carriage containing the read assembly is moved ever so slightly until constant reflection is achieved. In essence, the tracking beams are used to generate the feedback controlling the operation of the motor that moves the read head in and out.

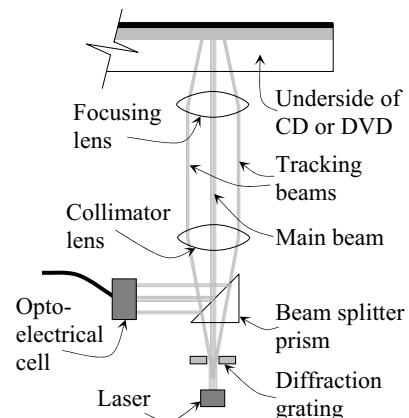


Fig 6.28
Detail of a typical optical storage read head.

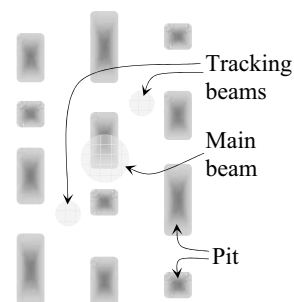


Fig 6.29
Magnified view of main and tracking laser beams.

The reflected light returns back through the focussing and collimator lenses and then is reflected by a prism onto an opto-electrical cell. The prism is able to split the light beam based on its direction; light from the laser passes through, whereas light returning from the disk is reflected. The term 'Opto-electrical' describes the function of the cell; it converts optical data into electrical signals. Changes in the level of light hitting the cell cause a corresponding increase in the output current. Constant light causes a constant current. Hence the fluctuations in the electrical signal correspond to the stored sequence of bits.

The electrical signal is then passed through a digital signal processor (DSP). The DSP removes the merge bits, converts the EFM codes back into their original bytes and checks the data for errors. Finally the data is placed into the drive's buffer where it is retrieved via an interface to the computer's RAM.



Consider the following:

If you shine a torch directly at a wall a circle pattern is seen, however if the torch is angled then the pattern becomes elliptical. Modern optical read heads are able to detect the difference between such patterns returned by the two tracking beams.



GROUP TASK Discussion

How could such information be used by optical drives to improve the performance of the retrieval process? Discuss.

Storing or writing optical data

There are two different technologies used to store data on optical disks; recordable which actually means data can only be written once but not erased, and rewriteable meaning the data can be erased and rewritten many times. Examples of both technologies are available for writing both CDs and DVDs. CD-R is the acronym used for recordable compact disks and DVD-R for similar DVDs. CD-RW stands for rewriteable compact disk. The standard for rewriteable DVD is currently a bit of a mess. Three competing standards exist; DVD-RAM, DVD-RW and DVD+RW, presumably just one of these standards will eventually prevail; I guess it's likely you already know the winner! Fortunately the basic principles and operation of all rewriteable optical disks is similar.

• Recordable or write once technology

The essential difference between recordable media and commercially produced stamped disks is the addition of a layer of dye between the clear plastic and the reflective metal. There are various different dyes used by different manufacturers, however initially they are all relatively clear and when exposed to heat turn opaque or cloudy. Drives capable of burning data onto recordable disks contain lasers that can operate at two power levels, low power for retrieving and higher power for storing or burning data.

In order to protect the dye layer from corrosion the reflective metal layer is commonly a mix of silver and gold; increasing the percentage of gold in the mix substantially increases the life expectancy of the data. Disks manufactured with 100 percent gold reflective layers are estimated to last for more than 200 years

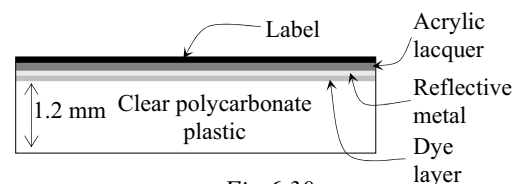


Fig 6.30

Cross-section of a recordable optical disk.

Storing data on optical disks first involves coding the data; this is essentially the reverse of the processes performed by the DSP during retrieval. The coded data is sent at a constant rate to the drive's processor. The processor responds to ones in the sequence of binary data; zeros merely cause a slight delay. If the laser is off and a one is encountered then it is turned on at high power, conversely if the laser is on and a one is encountered then it is turned off. Whenever the laser is on it produces heat and hence the dye layer turns opaque. As this is occurring the disk spins and the carriage moves slowly outwards. The result being a spiral track where the burnt or opaque areas on the track are the equivalent of the physical pits found on commercially produced disks, hence the recorded disks can be read on conventional optical drives.



Consider the following:

Writing a precisely placed spiral track on an otherwise flat surface is a difficult task, furthermore ensuring each pit (really areas of opaque dye) is of the correct length and is spaced accurately makes the task almost impossible. To solve these problems all blank recordable, and also rewriteable, disks are stamped during manufacture with a groove containing a wobble pattern along the path of the spiral track. The groove is followed during the writing process and the wobble pattern is used to ensure correct timing; the aim being to ensure the correct track pitch and linear distances between bits are maintained.



GROUP TASK Discussion

Based on your knowledge of tracking beams, explain how the spiral groove and wobble pattern could be used to ensure the correct track pitch and linear distances between bits are maintained?

• Rewriteable technology

Rewriteable media contains a recording layer composed of a crystalline compound sandwiched between two insulating layers. The crystalline compound currently used is a mixture of silver, iridium, antimony and tellurium. This unusual mix of elements normally reflects light well, however it has some interesting characteristics. If it is heated to between 500 and 700°C its crystal structure breaks down and so does its reflective properties. If, once cooled, the compound is then reheated to around 200°C it returns to its original reflective crystalline state. These characteristics form the basis of rewriteable storage. The high temperatures mentioned above must be localised within a microscopic area and these areas must be cooled quickly; this is the purpose of the surrounding insulating layers.

The laser used for storing data on rewriteable media has three different power levels. The highest level is able to heat the recording layer to between 500 and 700°C and is used for writing, the middle level heats to around 200°C and is used for erasing, and the lowest is used for reading data.



*Fig 6.31
A variety of different
rewriteable media. All are
the same physical size.*

The process of storing data on new rewriteable media is essentially the same as that used for recordable media. The only significant difference being the much hotter temperatures needed to break down the crystalline compound.

Rewriting data is slightly different; there are two techniques commonly used. One involves first erasing all the data, that is the laser is set at a constant erase power level whilst the entire data track is rotated above the laser. The disk can then be written as if it were new. A second technique allows new data to be directly written over existing data. This technique involves alternating the power of the laser between write power and erase power each time a one is encountered within the data.



Consider the following:

Currently CD-RW disks cost approximately four times that of a CD-R, however CD-RW disks can be reused more than 1000 times. Unfortunately the reflective properties of CD-RW disks are such that they cannot be read by many older CD-ROM drives, including most CD audio drives.



GROUP TASK Identify and justify

CD-R and CD-RW are suited to different applications. Identify applications where CD-R is more suitable and applications where CD-RW is more suitable. Justify your answers.

FLASH MEMORY

Flash memory is commonly seen in the form of memory cards; these cards provide removable storage for various electrical devices, for example digital cameras, MP3 players, PDAs, video game consoles, laptop computers and even mobile phones. Fig 6.32 shows a variety of different types of flash memory cards. Flash memory is not just used for removable storage; it is now becoming available as an alternative to magnetic hard disk drives (HDDs) in the form of flash solid state drives (SSDs). Flash memory is also included as an integral part of many devices. For example, BIOS chips, mobile phones, cable modems, DVD players, network routers, motor vehicles and even kitchen appliances. So what is flash memory?

Flash memory is electronic, solid-state and non-volatile; now what does that mean? Electronic devices use electricity; that is they manipulate electrons. Flash memory is a type of electronic storage that represents data by trapping or storing electrons. The essential difference between flash memory and other types of electronic storage, such as RAM, is the ability to trap electrons even when no power is present. This makes flash memory non-volatile. Solid state means there are no moving parts. Mechanical parts take time to do their job, generate noise and are prone to wear and failure. In contrast, flash memory is fast, silent and reliable. Furthermore, flash memory operates reliably within a much wider temperature range than magnetic or optical storage devices. For example, flash memory developed for motor vehicles is certified to operate from -40°C to $+125^{\circ}\text{C}$.



*Fig 6.32
A variety of removable flash
memory devices.*

Flash memory cards can be used in a variety of different devices. For example, *Fig 6.33* shows a variety of devices that include a Sony memory stick slot. Using a single memory stick you can take photos, edit them on your computer, view them on your TV and then send them to Grandma using your PDA!

If flash memory is so wonderful then why hasn't it replaced magnetic and optical storage? The answer is cost; compared to magnetic and optical storage flash memory is expensive. For example, presently (2009) a 250GB flash solid state drive (SSD) retails for about \$1,000, yet a 1TB (1000GB) hard disk drive retails for around \$100. Currently it's not economically feasible to include large capacity flash SSDs in most computers, however this is likely to change over the coming years. All the large microchip manufacturers are continually investigating a variety of high capacity solid-state non-volatile secondary storage technologies, flash technology being just one of the technologies under consideration.



Fig 6.33
Just some of the devices utilising
Sony's memory stick technology.



Consider the following:

What's a Solid State Disk (SSD)

A solid state disk/drive (SSD) - is electrically, mechanically and software compatible with a conventional (magnetic) hard disk.

The difference is that the storage medium is not magnetic (like a hard disk) or optical (like a CD) but solid state semiconductor such as battery backed RAM, EPROM or other electrically erasable RAM-like chip such as flash.

This provides faster access time than a hard disk, because the SSD data can be randomly accessed in the same time whatever the storage location. The SSD access time does not depend on a read/write interface head synchronising with a data sector on a rotating disk. The SSD also provides greater physical resilience to physical vibration, shock and extreme temperature fluctuations. SSDs are also immune to strong magnetic fields which could sanitise a hard drive.

The only downside to SSDs is a higher cost per megabyte of storage - although in some applications the higher reliability of SSDs makes them cheaper to own than replacing multiple failing hard disks. When the storage capacity needed by the application is small (as in some embedded systems) the SSD can actually be cheaper to buy because hard disk oems no longer make low capacity drives. Also in enterprise server acceleration applications - the benefit of the SSD is that it reduces the number of servers needed compared to using hard disk based RAID on its own.

Historically RAM based SSDs were faster than flash based products - but in recent years the performance of the fastest flash SSDs has been more than fast enough to replace RAM based systems in many server acceleration applications.

(Extract of an article on www.storagesearch.com)



GROUP TASK Research

Using the Internet, or otherwise, research current developments in non-volatile solid-state storage solutions. Are any of these new developments seen as a real alternative to current secondary storage devices?

NETWORK STORAGE

Have you ever wondered how banks, government departments, web-based email systems, in fact any large computer network manages to store and retrieve data for many thousands of employees and customers? What's more, they manage to do this fast and securely. For example, consider EFTPOS, it is rare to have to wait more than a few seconds for a transaction to be approved, similarly logging into your hotmail account takes mere seconds. Furthermore, the large proportion of this time is attributed to the transmitting and receiving of the data rather than its storage and retrieval. Our aim in this section is to introduce some of the hardware used to perform high speed and secure network storage and retrieval processes. This includes not only providing data access to users of the system, but also creating backup copies of such vast quantities of data.

We shall consider the two most commonly used technologies: RAID (Redundant Array of Independent Disks) and also tape libraries. RAID provides fast data access combined with inbuilt fault tolerance. Tape libraries, as the name suggests, provide access to multiple magnetic tapes. Such libraries are primarily used for automated backup processes, however they also provide relatively fast retrieval of archived data.

RAID (Redundant Array of Independent Disks)

RAID utilises multiple hard disk drives together with a RAID controller. The RAID controller manages the data flowing between the hard disks and the attached computer; the attached computer just sees the RAID device as a normal single hard disk. The RAID controller can be a dedicated hardware device or it can be software running on a computer. In most cases the computer attached to the RAID device is a server on a network. This means a RAID device can be added to an existing network with minimal changes to existing server software and no changes to any other machines on the network.

Simple RAID systems contain just two hard disks whilst large systems may contain many hundreds of disks. The RAID controller's job is to manage all these drives to improve data access speeds and fault tolerance. RAID is based on two basic processes, striping and mirroring. Striping improves read/write access times and mirroring improves fault tolerance and read times. Let us consider the operation of each of these processes.

Striping splits the data into chunks and stores chunks equally across a number of hard disks. During a typical storing or retrieving process a number of different hard drives are writing/reading different chunks of data simultaneously (see Fig 6.36). As the relatively slow physical processes within each drive occur in parallel, a significant improvement in data access times is achieved.

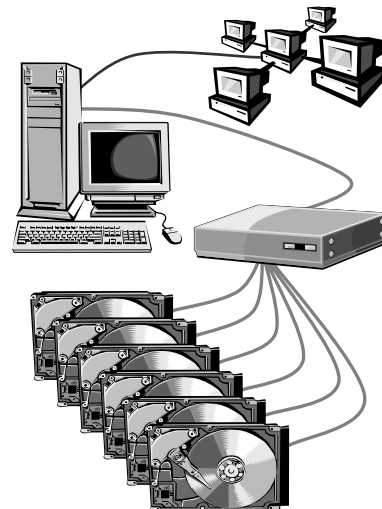


Fig 6.34

Main components of a RAID mass storage system attached to a network.



Fig 6.35

A variety of different RAID devices.

Mirroring involves writing the same data to more than one hard disk. *Fig 6.36* shows the simplest example of mirroring using just two hard disks where both disks contain identical data. When identical copies of data are present on different hard disks the system is said to have 100% data redundancy. Should one disk fail then no data is lost, furthermore the system can continue to operate without rebuilding data after the complete failure of a disk. Hence mirroring makes it possible to swap complete hard disks without halting the system; this is known as 'hot swapping'. Many larger RAID systems also include various other redundant components, such as power supplies; these components can also be 'hot swapped'. Data redundancy and the ability to 'hot swap' components improve the system's fault tolerance.

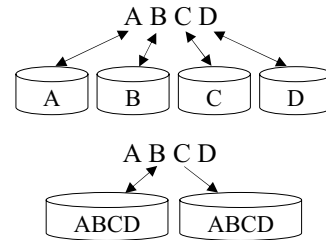


Fig 6.36

Striping (top) and mirroring (bottom) processes are the basis of RAID systems.

So mirroring improves the fault tolerance of the system, but what about read and write access times? Write access times are not reduced; in fact they may actually increase slightly due to the extra processing performed by the RAID controller. When mirroring, all data is written simultaneously to multiple hard disks; hence the time taken is similar to writing all the data to a single drive. On the other hand, retrieving data is quicker as any of the drives containing the data can be used; the RAID controller can make a choice, if the first drive is busy with another process the data can be retrieved from a different drive.



Consider the following:



GROUP TASK Research

Use the Internet, or otherwise, to determine the most commonly used RAID levels. Describe how each of these RAID levels implements striping and mirroring.

Tape libraries

Have you ever made a complete backup copy of a hard disk? It involves manually swapping media and a good deal of time; these are major disincentives. Now imagine performing the same process for all the data held by a large organisation; hundreds or even thousands of tapes need to be swapped taking days or even weeks to complete. Clearly the backup process needs to be automated, this is the purpose of tape libraries.

Tape libraries, such as the one shown in *Fig 6.37*, include multiple tapes and multiple tape drives. A robotic system moves tapes between the storage racks and the tape drives. The tape drives are just normal single drives whose operation has been automated.

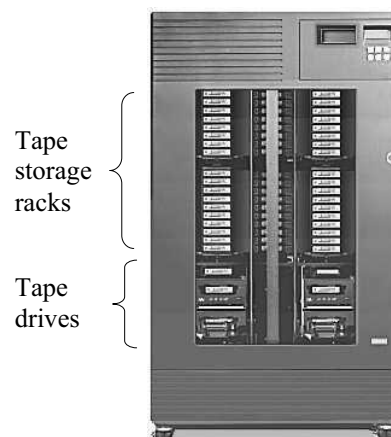


Fig 6.37

Qualstar's TLS-58132 tape library stores up to 340 terabytes of data.

Various different size tape library devices are available to suit the demands of different information systems. The smallest, such as Sony's TSL-SA400C in *Fig 6.38*, hold just four tapes and use a single drive; these devices provide capacities suited to most small businesses. Larger devices hold hundreds or even thousands of tapes and contain many drives. Large government departments and organisations link multiple tape library devices together; such systems hold hundreds of thousands of tapes and many thousands of tape drives. Backup processes on such large systems continue 24 hours a day, seven days a week.



Fig 6.38
Sony TSL-400C tape library.



Consider the following:

StorageTek's StreamLine™ SL8500 shown in *Fig 6.39*, has a minimum configuration of 1448 tapes which uses 64 tape drives. This modular system can be increased by combining up to 7 units to hold up to 70,000 tapes using 448 tape drives.

The system is capable of using tapes and tape drives of various types. Current tape capacities range from 20GB to 500GB per tape. Data storage speeds from 8 terabytes per hour up to 250 terabytes per hour are achievable depending on the type of tapes used, together with the configuration of the system

The average time taken for the robotics to place a tape in a drive is around 6.25 seconds. To identify individual tapes the robotic arms contain barcode readers, each tape being individually bar-coded. Furthermore redundant robotics, power supplies and electronics can be optionally installed to increase the fault tolerance of the system.

Clearly such systems are aimed at large corporate and government organisations that maintain extensive large-scale computer systems. Such systems are held in secure air-conditioned environments.



GROUP TASK Activity

Determine the minimum and maximum storage capacity of the tape library system described above.



GROUP TASK Discussion

RAID and tape libraries help to secure data, but do they make it 100% secure? Discuss.



Fig 6.39
Exterior and interior of
StorageTek's StreamLine™
SL8500 tape library.

SET 6B

1. Which of the following best describes how binary data is represented on magnetic media?
 - (A) A one is represented by a north pole and a zero by a south pole.
 - (B) The direction of the magnetic field is used. One direction for ones and the other for zeros.
 - (C) High magnetic forces represent ones and occur where the direction of the magnetic force changes. Low forces represent zeros.
 - (D) A magnetic force exists where a one is represented and does not exist where zeros are represented.
2. Which of the following terms does NOT describe MR materials?
 - (A) soft magnetic
 - (B) conduct electricity better when close to a magnetic field.
 - (C) used during the storing process.
 - (D) used during the retrieving process.
3. The primary advantage of magnetic tape over other types of secondary storage is:
 - (A) the speed of data access.
 - (B) the ability to retrieve data sequentially.
 - (C) that tape is much cheaper.
 - (D) that tapes can be removed and stored off-site.
4. In a RAID device the process of striping is best described as:
 - (A) Storing the same data on multiple drives.
 - (B) Splitting up data, and storing each chunk simultaneously on different drives.
 - (C) A technique for improving read times.
 - (D) A method for improving fault tolerance.
5. The EFM coding system, together with merge bits, are used:
 - (A) for error checking during the retrieval of data.
 - (B) to restrict the length of both pits and lands so read errors do not occur.
 - (C) to ensure both pits and lands are of sufficient length to be read accurately.
 - (D) Both (B) and (C)
6. Helical tape systems:
 - (A) use many components from audio and video tape drives.
 - (B) write tracks at an angle to the length of the tape.
 - (C) tend to wear out tape more rapidly than linear systems.
 - (D) All of the above.
7. Which of the following best describes how binary data is represented on optical media?
 - (A) Lands represent zeros and pits represent ones.
 - (B) A change from land to pit or pit to land represents a zero whilst no change represents a one.
 - (C) A change from land to pit or pit to land represents a one whilst no change represents a zero.
 - (D) Changes in reflection are read as ones, whilst constant reflection is read as a zero.
8. Which of the following is true of sectors on hard disks?
 - (A) Each track is always split into the same number of sectors.
 - (B) All sectors on a particular hard disk store the same quantity of data.
 - (C) The physical area of each sector is always the same.
 - (D) Commonly the number of sectors per track increases as the radius of the track decreases.
9. The significant difference between CAV and CLV drives is:
 - (A) Data passes under the read head of a CLV drive at a relatively constant speed; this is not the case with CAV drives.
 - (B) The spindle motor operates at varying speed on a CLV drive but at a constant speed on a CAV drive.
 - (C) CAV drives vibrate more as they spin at much greater speed than CLV drives.
 - (D) The time taken to vary the speed of rotation in a CLV drive limits data transfer rates, hence CAV drives have higher data transfer rates.
10. Commonly the read/write head of an optical drive generates three laser beams. Why are three laser beams needed?
 - (A) So that three data tracks can be read or written simultaneously.
 - (B) One beam is used to read or write the data, whilst the others ensure the head remains centred on the data track.
 - (C) One beam is used for the actual data and the other two are used for correcting errors within the data.
 - (D) The use of three beams means that the laser does not need to be precisely focussed on the data track.

11. Describe the components and operation of the read/write head within a hard disk during:
 - (a) a storing process.
 - (b) a retrieval process.
12. Describe the components and operation of the read/write head within a CD-R drive during:
 - (a) a storing process.
 - (b) a retrieval process.
13. Describe how data is organised on the following storage media:
 - (a) hard disks.
 - (b) magnetic tape.
 - (c) optical disks.
 - (d) RAID devices.
14. Compare and contrast:
 - (a) hard disk storage with magnetic tape storage.
 - (b) recordable CDs and rewriteable CDs.
 - (c) RAID devices and tape libraries.
 - (d) Mirroring and striping used by RAID devices.
 - (d) Flash memory with RAM.
15. Research both the storage capacity and data transfer rates for a variety of different models of RAID devices and tape libraries. Make up a table to summarise your results.

SOFTWARE IN STORING AND RETRIEVING

Software controls and directs the operation of all hardware, including all the various types of storage devices. Software causes hardware to perform processes that ultimately assist in achieving the system's purpose. So what software is used to perform storing and retrieving processes, and what does it do? To answer this question we first consider the various types of software operating behind the scenes to interface with storage hardware. We then consider the format of data files and how these formats affect storing and retrieving processes.

Virtually all application software utilises storing and retrieving processes, however there are particular types of software whose central purpose is managing the storing and retrieving of data. We consider examples of such software, namely file management software, database management systems and Internet or web browsers. Finally we discuss techniques for securing stored data, namely passwords and encryption of data.

THE HARDWARE TO SOFTWARE INTERFACE

It would be inefficient for every software application to direct and control all aspects of the storing and retrieving process. Rather such processes are split into various sub-processes performed by different programs, each piece of software being dedicated to a particular part of the storing and retrieving process. To identify and describe the software components involved in storing and retrieving let us consider a typical storing process; namely saving a file from within a software application to the hard disk. These steps could easily be adjusted to describe any storing process or reversed to describe a retrieving process:

1. Typically the user interacts with the application to initiate the save; this involves selecting a location for the file, and specifying a file name and storage format. This is a collecting information sub-process.
2. The application informs the operating system and passes it the location and name of the file. The operating system is now in control of the storing process.
3. The operating system directs the device driver associated with the appropriate storage device to proceed with the storing process. We discussed device drivers in Chapter 3 (see page 103); essentially the device driver provides a software interface between the operating system and the actual storage device.
4. Once the storage device is ready to commence the storing process it's device driver informs the operating system. The operating system then instructs the application to commence sending data directly to the device driver.
5. It is the job of the application to organise the data into the appropriate file format prior to sending it to the device driver. The device driver in turn passes the data from the application to the actual hardware storage device.
6. Within the storage device is further software, often called firmware, together with a buffer. The data stream progressively arrives and is held in the buffer as it waits its turn to be processed by the firmware. Firmware is permanently stored software within the hardware device; essentially the brain of the device.
7. The firmware controls the mechanics of the storage device to physically move components and store the data. It also reorganises the data as it leaves the buffer to suit the requirements of the actual storage device. For example, on a hard disk the data is split into appropriately sized chunks corresponding to the size of individual sectors on the disk. The firmware is not concerned with the file's format; it just sees the data as a stream of binary digits.

Throughout the whole process the operating system maintains ultimate control. Control messages are being relayed back and forth between the hardware and up through the various different software programs. These messages control the data transfer as well as ensuring the accuracy of the data.

Fig 6.40 at right depicts the software and hardware components, together with arrows indicating the exchange of both data and control messages. This diagram applies to both storing and retrieving processes, hence the data arrows point in either direction. Essentially retrieving processes are the reverse of storing processes.

The dotted line around the software components in Fig 6.40 also indicates software that executes on the main CPU; the firmware within the storage device being executed on a dedicated processor within the storage device. The gap between the two dotted rectangles represents the physical wires connecting the computer to the storage device. The arrangement of these wires, together with the connectors and rules for transferring data are part of an interface standard; the most common of these standards being ATA (Advanced Technology Attachment) and SCSI (Small Computer Standard Interface). The older parallel IDE (Integrated Drive Electronics) standards are often referred to as simply ATA, however the acronym PATA (Parallel ATA) is also used. The more recent Serial ATA (SATA) interface has largely replaced the Parallel ATA interface. Software with the ability to operate these interfaces is contained within the computer's BIOS (Basic Input Output System) and is loaded as part of the initial startup process.

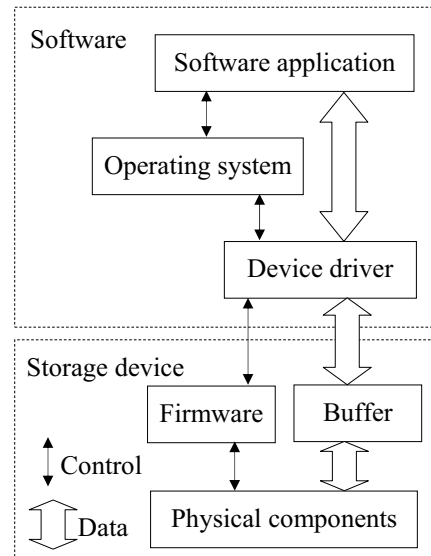


Fig 6.40

The interface between storage devices and software applications.



Consider the following:

The operating system and the device drivers are stored on secondary storage. Our discussion above requires the operating system and the device driver stored on the hard disk to be loaded prior to retrieving data. It's a Catch-22 situation; you can't access the hard drive without the operating system and the device driver, yet to load this software requires the hard drive to be operational!

Fortunately, the hard disk contains firmware instructions and its own processor. Furthermore, the computer's BIOS is also firmware held on a dedicated chip. Both these firmware components are crucial to the successful start-up of all computers.



GROUP TASK Discussion

Obviously the operating system and device drivers do somehow get loaded from secondary storage. Discuss how this occurs.



GROUP TASK Investigation and discussion

Some firmware can be updated or edited whilst other firmware is completely permanent. Can you update the firmware within your home or school computer? Is firmware hardware or is it software? Discuss.

FILE FORMATS FOR DIFFERENT DATA TYPES

Organising data into a particular file format in preparation for storage is clearly an organising information process, however the file format chosen has implications in regard to the efficiency of storing and retrieving processes. The file format influences the size of the file and also the way in which the file may be retrieved.



Consider bitmap image files

An image saved as a Windows bitmap (.BMP) uses significantly more storage than the same image saved as a JPEG file; the JPEG format includes the ability to significantly compress the data. Clearly storing and retrieving a compressed JPEG file takes less time than the larger BMP file. Furthermore, most files, including bitmap image files, are retrieved sequentially. As most bitmap files are arranged into rows of pixels commencing with the top (or bottom) row and ending with the bottom (or top) row the complete image cannot be displayed until the retrieval process is complete. Some bitmap formats, including JPEG, include the ability to arrange rows of pixels in non-sequential order such that a low-resolution version of the image is first displayed. As further rows are retrieved the resolution increases until eventually the complete image is displayed; JPEG files arranged in this manner are called progressive JPEGs.



GROUP TASK Calculate

A 505 by 391 pixel bitmap image has a colour depth of 24 bits. Saving this image as a Windows BMP requires 578KB of storage, however when saved as a JPEG the size is just 29KB. Explain how the BMP size can be calculated and express the compression of this BMP to JPEG as a ratio.



GROUP TASK Discussion

Compressed image file formats are used extensively on the web and also within digital cameras. Why is this? List and discuss reasons.



Consider the following:

Video data files are commonly organised in such a way that they can be progressively displayed as the file is being retrieved; this process is known as streaming. On many websites it is possible to jump to a later scene without the need to download all preceding scenes. On other websites the user must wait for all intermediate scenes to be retrieved before the later scene is displayed.



GROUP TASK Research

How can the differences described above be explained? Discuss.



GROUP TASK Discussion

Database Management Systems are able to retrieve specific records within a single file. How is this ability similar to the ability to jump directly to scenes within a video file? Discuss.

FILE MANAGEMENT SOFTWARE

File management software is used to logically organise files on secondary storage devices. Most operating systems include file management software. For example, Windows Explorer is an integral part of the Windows family of operating systems. Such software is not concerned with the data within files but rather with the manipulation of complete files. The aim is to present a logical arrangement of the files to the user and to provide processes for manipulating files within this arrangement.

What do we mean by the term logical when referring to the arrangement of files? Different storage devices physically store data differently, however file management systems are able to translate this physical arrangement of data into a consistent logical arrangement. When using file management software the directory structure appears similar regardless of the type of storage device. For example, the screen from Windows Explorer shown in *Fig 6.41* includes hard disks, removable storage and network drives, yet all are presented to the user in a similar consistent manner. Furthermore, from the user's perspective, opening any of these devices and manipulating files is performed using identical actions. In effect file management software hides the physical details of where and how files are stored and manipulated.

So what is this logical arrangement? Files are arranged into a hierarchical structure of directories or folders. Each storage device has a root directory, which may contain both files and other directories; each of these directories also contains files and/or further directories and so on. For example, in *Fig 6.42* the file *abc.gif* is within the directory called *Images*, which is within the directory called *IPT Text*. Actually a directory is merely another file, it contains the name, location and various other details of each of its files. Recall our discussion in regard to archive bits on page 192; there is a similar bit set for files that are directories. The file management system, in consultation with the operating system, reads this bit to determine files that are directories.

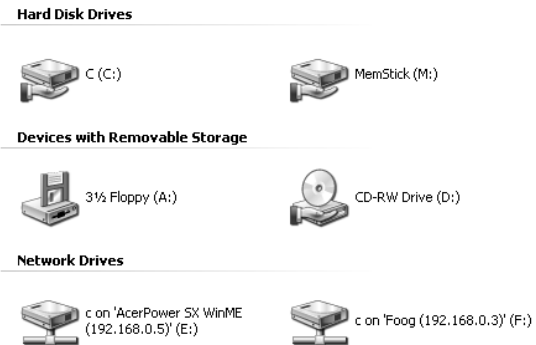


Fig 6.41
Various types of storage device are accessed using the same user interface.

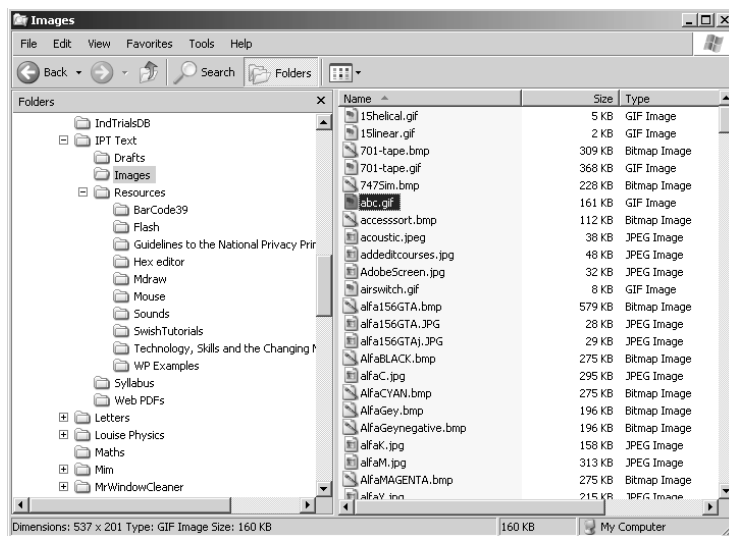


Fig 6.42
Screen shot from Windows Explorer included with Microsoft's Windows XP operating system.



GROUP TASK Activity

Most file management software includes new, cut, copy, paste and rename functions. Use each of these functions to manipulate files on either your home or school computer. List and describe other functions available.



Consider the following:

Data on most current hard disks is physically stored in individual 512-byte sectors. Many operating systems utilise a storage system known as FAT, which combines multiple sectors into clusters (typically from 4 to 64 sectors per cluster); each file resides within a particular number of complete clusters. A file allocation table (FAT) on the disk contains an entry for every cluster. These entries indicate whether a cluster is free, damaged or being used to store part of a file. If it is being used then this entry either points to the next cluster holding data for the file or it contains a flag indicating it is the last cluster for the file.

The directory file contains entries for each file within the directory. Each of these entries includes the address of the first cluster on the disk containing the file. When the operating system wishes to access a file it retrieves the address of the first cluster from the directory file; subsequent cluster addresses being obtained from the FAT. These addresses are submitted to the hard disk, which responds by retrieving the data within the sectors corresponding to the specified cluster addresses.



GROUP TASK Discussion

Deleting a file does not actually remove the data. Based on the above information, discuss what is likely to be occurring during a typical delete operation.



GROUP TASK Discussion

A file containing exactly 30000 bytes is being stored. Assuming each cluster contains four 512-byte sectors, calculate the number of clusters used and describe the changes made within the file allocation table.

DATABASE MANAGEMENT SYSTEMS (DBMS)

Database management systems are software applications used to store and retrieve data within databases. Most databases contain various types of data arranged into multiple tables where each table is composed of records. We discussed the organisation of such data back in Chapter 4 (p154-155); it may be worthwhile reviewing these pages. Now imagine the size of a database maintained by even a small organisation or business; it is likely to contain many tables and many thousands of records and furthermore many users have simultaneous access to this data. For example, a user can be entering an order for a customer whilst another user is analysing sales trends; they are both accessing the same data. In large organisations the number of users and the number of records becomes massive, perhaps many thousands of users and many millions of records. It is the job of the DBMS to manage the storing and retrieval of this data so that all users have access in a logical and efficient manner. Clearly DBMS software does not simply retrieve complete files.



GROUP TASK Calculate

Most schools maintain their timetable in a database. Details of each student being held in one table, details of each class in another, and a further table linking each student to their classes. Calculate the approximate number of records held in each of these tables within your school's timetable database.

In regard to storing and retrieving, DBMS software must manage:

Retrieval of just the set of records requested by a user.

Most retrieval processes read an entire file, for database applications this is unworkable. Even a single database table will commonly be much larger than the amount of RAM within even large servers, furthermore transmitting such massive amounts of data to a workstation would be most inefficient, hence DBMS software reads individual records, rather than individual files. For example, when searching for a particular customer's record a DBMS retrieves each customer record in turn until it finds the required record; during this process, records that don't match are simply discarded from RAM. This record based retrieval explains why database files must be highly organised; every record composed of fields, where each field contains data of the same type and length. Such structured organisation allows the DBMS to identify the precise location of individual records within the stored database.



Consider the following table's data dictionary

Products			
Field Name	Data type	Size (bytes)	Description
ProductID	Long Integer	4	Primary key
ProdName	Text	50	Name of product
ProdDesc	Text	100	Description of product
CategoryID	Long Integer	4	Link to ProductCategory table
WholePrice	Currency	8	Wholesale price



GROUP TASK Calculate

Imagine a company sells a total of 5000 different products whose details are held in the above Products table. Calculate the total amount of storage required.

Multiple users accessing and possibly editing the same data.

As DBMS software retrieves records rather than files then editing can also be controlled based on records rather than complete files. Imagine two users have retrieved the same record, if both users subsequently make changes to this record then which version of the record should be stored? The DBMS must implement a strategy whereby records can be locked; commonly DBMSs provide two different strategies; pessimistic locking and optimistic locking.

Pessimistic locking, as the name suggests, is somewhat negative. The first user to start editing the record effectively locks the record and hence subsequent users must wait for the updated record to be stored before they can commence editing; often a visual aid (see Fig 6.43) is used to inform the user. Such a strategy requires the DBMS to be informed and lock the record whenever a user commences editing a record. Such a strategy adds considerably to the amount of processing required of the DBMS.

Fig 6.43

Microsoft Access displays a  symbol when pessimistic locking is active and another user is editing a record.

Optimistic locking is a much more positive strategy; based on the assumption that conflicts will rarely occur. Such a strategy does not require the DBMS to be informed as editing commences, rather the DBMS checks for record changes prior to storing each record. If another user has made a change then there are two possible options, either the record can be overwritten or the current changes can be discarded. Commonly the user is given the task of making this decision via a warning message. Fig 6.44 shows the default message generated by Microsoft Access. In either case all but one user is destined to lose their changes.

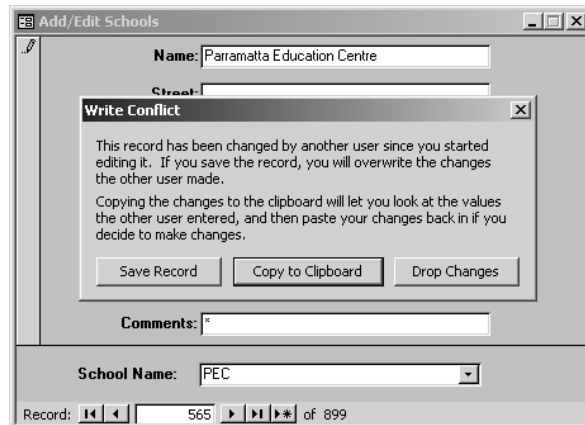


Fig 6.44

Microsoft Access provides 3 options in response to write conflicts when optimistic locking is enabled.



Consider the following:

Similar problems to those described above can occur within software applications that read and write complete files. For example, what happens when a word processor file is opened and edited on more than one computer?



GROUP TASK Investigation

Open and attempt to edit the same word processor file simultaneously on a number of computers. Try using different software applications to simultaneously open the file. Describe your findings.



GROUP TASK Discussion

Do your findings from the above investigation have similarities to the two different locking strategies discussed above? Discuss.

Securing data by restricting user's access to records.

Restricting access to complete files can be accomplished using various operating system and network functions, however access to databases commonly involves restricting access to particular tables or even particular records within a table. These tasks require a detailed knowledge of the database structure; hence DBMS software is in the best position to accomplish such tasks.

Each user is assigned a set of permissions. For example, an order entry clerk may be able to read customer details but not change them, yet they may be able to both add and edit invoices. Commonly, users are required to enter a user name and password each time they use a particular database, other DBMS systems utilise the network user name to verify the identity of the current user. In either case the identity of the user is determined and their data access rights assigned accordingly.



GROUP TASK Activity

Many discussion forums on the Internet permit read-only access for guests; a username and password being required to add or edit data. These forums are really databases and users interact via a web page linked to a DBMS. Use such a forum and describe the security used.

INTERNET OR WEB BROWSERS

Web browsers are software tools that essentially collect and display data and information retrieved from web servers. In one sentence we've mentioned three of the seven syllabus information processes: collecting, displaying, and retrieving. In reality most software, including browsers, utilise all the seven information processes. So far in this text we have considered browsers as applications for collecting data (p109), we have examined the organisation of data used by browsers (p156) and then earlier in this chapter (p201) we discussed how browsers utilise storage and retrieval processes to assist the transfer of data. These were all relatively technical discussions. Let us discuss browsers a little differently; from the point of view of the user.

Browsers are such common software applications; virtually every computer with an Internet connection has a browser. Browsers provide the human interface between users and the vast store of information out there in cyberspace. Browsers allow users to navigate and explore the web with virtually complete ignorance in regard to the underlying processes occurring. From the user's perspective browsers provide access to a vast store of information. Furthermore, they assist users to locate specific information via search engines.

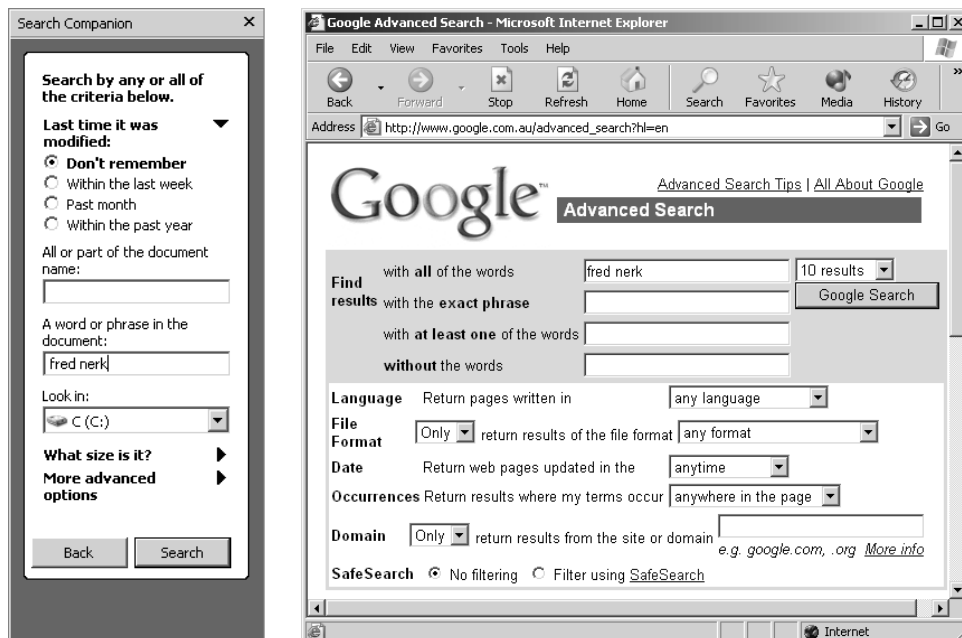


Fig 6.45

Search screens from Windows XP and the search engine Google.

To display a web page the user enters a URL or clicks on a hyperlink, the browser responds by opening that file; to the user the procedure is essentially the same as opening a file on a local hard disk. In essence, the location of the file is specified and then the data is retrieved. What if you don't know the location or name of the file you wish to retrieve? Well you search for it, of course! To search the web you use a search engine, to search your hard disk you use the search function provided with the operating system; compare the screens in Fig 6.45 above, there are clear similarities.



GROUP TASK Discussion

"It is simply unnecessary for users to understand the underlying processes occurring within software applications, or even within complete information systems." Do you agree? Discuss.



Consider the following:

Web pages stored on web servers all over the world can be read and displayed on virtually any computer connected to the Internet. The only requirement being that the computer has a web browser installed. Hence all the web pages contained within the entire World Wide Web (WWW) form an enormous data store that is independent of the hardware on which the pages are viewed. This machine independence of web pages is perhaps the primary reason behind the incredible success of the WWW.



GROUP TASK Discussion

Why do you think the machine independence of web pages is such a significant factor contributing to the success of the WWW? Discuss.

SECURING STORED DATA

Information such as credit card details, medical records, financial records and various other types of personal or sensitive data must be secured to protect against unauthorised access; in many cases this is a legal requirement. Security measures are also implemented to assist in the management and operation of information systems. For example, it would be inappropriate for a managing director to change network configuration settings or for a data entry operator to alter programming code.

There are two different strategies commonly used by software to secure stored data against unauthorised access. The first, and most common line of defence is to use passwords to restrict or prevent access. The second line of defence is to scramble or encrypt the stored data.

Password protection

Passwords can be used to secure individual files, directories or even entire storage devices. A combination of user names and passwords are used by operating systems, network software and various other multi-user applications to confirm the identity of users. Read/write access to data, together with various other permissions, are assigned based on the user name.

Passwords for individual files are set by the file's related software application. For example in Microsoft Word a password can be set via the tools menu within the 'Save As...' dialogue. In many software applications setting a password also causes the file to be encrypted; this is not always the case. In many applications the file can simply be opened using another application and the raw data can be read.

Data secured by passwords is only secure whilst the passwords remain secret. There are numerous techniques and also software applications available for working out passwords. Furthermore, remembering many different passwords is difficult, hence people tend to either use the same password for multiple systems or they write down their passwords. There have been cases where the user names and passwords for entire systems have been typed into totally unsecured text files, which are easily accessible to hackers.



GROUP TASK Activity

Develop a list of recommendations that users should follow to protect the security of their passwords.

Encryption and decryption

The science of developing and analysing encryption and decryption technologies is called cryptography. The military have used cryptography to secure messages for hundreds of years. In fact many of the techniques and strategies now widely used evolved from these military applications. Cryptography has now become a major industry due to the widespread need to secure sensitive digital data.

Encryption alters raw data in such a way that the resulting data is virtually impossible to read. Therefore should unauthorised access occur the infiltrator just sees a meaningless jumble of nonsense. Of course, this would be a pointless exercise if authorised persons cannot reverse the process and decrypt the data. To enable decryption, secret information, called keys, are used. The key contains sufficient information to encrypt and/or decrypt data to the required level of security. Some systems use a single key for both encryption and decryption whilst others use a different key for each process.

Single key encryption is commonly called symmetrical or secret key encryption. The same key being used to decrypt the data as was used for encryption. Such systems are commonly used to encrypt data held on secondary storage devices. The device itself, or at least the attached computer, does all the encrypting and decrypting. As a consequence it is not necessary for the secret key to be shared, although it must be securely protected. If the user or computer decrypting the data is different from the one who encrypted the data then the secret key must be shared with both parties. A secure encryption technique is needed to communicate the secret key. Solving issues such as this is the job of cryptographers; one solution is the use of systems that use two keys.

Two key systems utilise a public key for encryption and a private key for decryption; they are known as asymmetrical or public key systems. Each user of the system has a public key and a private key. The public key can be distributed freely to anybody or any computer, however the private key must never be divulged. Let us consider a typical transfer of data, say from Fred to Jane (see Fig 6.46). Jane has her own personal public and private key, as does Fred. Fred first sends a plain message to Jane requesting her public key. Jane responds by sending Fred a copy of her public key; Fred uses this key to encrypt the message. He then sends the encrypted message to Jane. Jane receives the message and decrypts it using her private key. The message is secure during the transfer as only Jane's private key is able to decrypt the message, and Jane is the only one who has this key. It doesn't matter if Jane's public key is intercepted during the transfer as it can only be used for encrypting messages, not decrypting them. Our example used two people, in reality the transfer may well be between two computers.

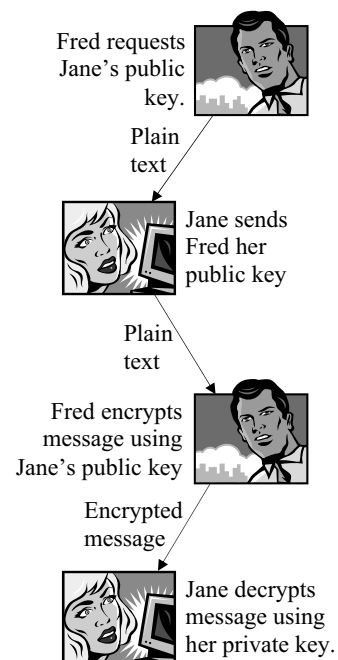


Fig 6.46
Typical transfer using a public or two key system.



GROUP TASK Research

Phil Zimmerman developed a public key software encryption system known as PGP (Pretty Good Privacy). There are significant restrictions on the export of this software. Use the Internet to research the nature and reasons for these export restrictions.



Consider the following:

It is common for systems that store highly sensitive data to use a combination of encryption techniques. In many organisations users carry flash memory-based smart cards containing their private keys. These cards must be inserted into a reader before any data can be decrypted and viewed. On file servers' data is encrypted using a different technique, often involving further levels of encryption.

The data stored on many file servers is encrypted, and the key for decrypting this data is itself held on a removable flash device attached to the file server. During retrieval the file server uses the key on its flash device to decrypt the data, then prior to transmission the data is encrypted using the public key of the current user. Once the user receives the data it is decrypted using the private key on their smart card.

However what if a user's smart card is stolen? Surely the thief then has complete access. To counteract this possibility a password can be used to confirm the user's identity corresponds with the owner of the smart card. But passwords can be guessed, or users can divulge their password. Such problems can be overcome using biometric data, such as fingerprints, to replace passwords; the biometric data being used to confirm the identity of the user. *Fig 6.47* shows one such device, a keyboard incorporating a fingerprint scanner and a smart card reader.



Fig 6.47
Precise Biometric's Precise 100SC integrates a fingerprint scanner, smart card reader and keyboard.

Even more elaborate schemes can be used. Some storage systems use a different key to encrypt every file. They then encrypt each of these individual keys using the key on the server's flash card. Such systems allow the key on the flash card to be changed at any time without the need to decrypt and then encrypt all the data on the entire storage device. Similarly the use of smart cards for users means their public and private keys can easily be altered at any time.



GROUP TASK Discussion

Why do you think smart cards are used within the systems described above? Why not just use fingerprints or passwords? Discuss.



GROUP TASK Discussion

"Ultimately all security depends on the honesty and integrity of the system's users." Do you agree with this statement? Discuss in terms of the above discussion.



GROUP TASK Discussion

Physical products are protected using physical means, such as locks and security guards. What is so different about digital data that means it requires all this extra security? Discuss.

SET 6C

1. When an application retrieves data from a storage device:
 - (A) all data must pass through the operating system.
 - (B) the application communicates directly with the device driver.
 - (C) the data passes from the device driver directly to the application.
 - (D) the firmware within the storage device passes the data to the operating system and then onto the application.
2. An uncompressed bitmap image has a resolution of 300 by 200 pixels and each pixel is restricted to a palette of just 8 colours. The size of this file will be closest to:
 - (A) 480Kb
 - (B) 180Kb
 - (C) 60Kb
 - (D) 22.5Kb
3. Firmware within storage devices is used to:
 - (A) control mechanical operation.
 - (B) reorganise the data.
 - (C) communicate with the device driver.
 - (D) All of the above.
4. The organisation of data displayed by most file management software:
 - (A) shows the physical arrangement of the data on the storage device.
 - (B) shows the logical arrangement of files on the storage device.
 - (C) is different depending on the nature of the storage device.
 - (D) includes the structure of tables and records.
5. In terms of the physical storage of data, a directory is:
 - (A) a container for files.
 - (B) used to locate files on the device.
 - (C) just like any other file.
 - (D) always within the root directory.
6. When deleting a file using most file management software:
 - (A) the data is physically erased.
 - (B) the data is not physically erased.
 - (C) a single directory entry is removed.
 - (D) Both (B) and (C)
7. In terms of the storing and retrieving process, the most significant difference between file management software (FMS) and DBMS software is:
 - (A) FMS processes records, whilst a DBMS processes complete files.
 - (B) FMS processes complete files, whilst a DBMS processes records.
 - (C) unlike DBMS software, FMS is included with most operating systems.
 - (D) unlike FMS, DBMS software allows multiple users access to the same data.
8. A file has been password protected by the user of an application, however a second user who does not know the password is able to open and view the information. Which of the following is the most likely?
 - (A) The file has been encrypted using the public key of the second user.
 - (B) The password used is equivalent to the second user's network user name.
 - (C) The original file was copied and the second user opened the copy.
 - (D) The file was not encrypted and the second user opened it within a different application.
9. Public key encryption systems require:
 - (A) the same key for both encrypting and decrypting.
 - (B) no sharing of public keys.
 - (C) that private keys be shared.
 - (D) the sharing of public keys.
10. A single character is stored in a text file. As expected, the file is reported to contain a single byte of data, however on the hard disk it physically occupies 32,768 bytes. Why is this?
 - (A) The file also contains formatting, font, error checking and other data.
 - (B) Files must occupy complete clusters, and this disk uses 64 sectors per cluster.
 - (C) Files are stored using complete sectors. Each sector is 32,768 bytes on this hard disk.
 - (D) The character is being stored as a bitmap image rather than as ASCII text.

11. Describe, as a series of steps, the software used and processes occurring as a file is opened from within a software application.
12. Previously in this section we described web pages as being “machine independent”, meaning they can be accessed and displayed by virtually any computer running web browser software.
Explain how each of the following examples of “machine independence” is achieved.
 - (a) A particular model of hard disk drive can be installed and used on virtually any type of computer system.
 - (b) An audio CD can be played in almost any optical drive. They can be played in cars, on a home stereo, in a DVD drive or even in a computer’s CD or DVD drive.
 - (c) JPEG image files are used by many devices. Most scanners, digital cameras and even many mobile phones produce JPEG files. Furthermore they can be displayed on most computers; in fact it would be rare to find a computer that does not contain software capable of displaying JPEG files.
13. Calculate the approximate storage capacity of each of the following files:
 - (a) A bitmap image file with a resolution of 640 by 480 pixels at a colour depth of 24 bits.
 - (b) A stereo sound file of 20 seconds duration. The file contains 16 bit samples and 10,000 samples are used for each second of both left and right channels.
 - (c) A database table containing 5000 records where each record contains 4 fields. One field holds integers within the range -32768 to 32767, another holds True/False data represented as 1 or 0, and the last two are text fields with length 4 and 10 characters respectively.
14. Explain the difference between each of the following:
 - (a) File management software and database management systems.
 - (b) Optimistic and pessimistic locking strategies.
 - (c) Passwords and encryption.
 - (d) The two keys used in a public key encryption system.
15. Jack performs most of his banking over the Internet. For this to occur data must be sent securely from Jack to the bank and vice versa, hence a public key encryption system is used. At the start of each session Jack enters his user name and password into the system, the bank’s system responds by retrieving and sending back the current balance and details for each of his accounts.
Explain the encryption and decryption processes occurring during the start of Jack’s banking session.

NON-COMPUTER TOOLS FOR STORING AND RETRIEVING

In previous chapters we discussed the nature of non-computer storage in terms of its organisation and its ability to be analysed. In this section we concentrate on its disadvantages, and advantages, in terms of storing and retrieving processes. On the whole, data stored within non-computer systems is manually retrieved and manually stored. Obviously this takes significantly more time than an equivalent computerised system. Furthermore, the physical space required for computerised data is negligible compared to that required to store the data using non-computer methods.



Consider the following:

- **Paper-based storage**

Most offices still maintain paper-based filing systems; why is this? Surely any data that can be stored on paper can be more efficiently stored on computer-based systems. In terms of physical space and data access times this is true. For example, a single DVD has a physical volume similar to that of just a few sheets of paper, yet a DVD can store all the data held within a large bank of filing cabinets. Furthermore, hundreds of megabytes can be retrieved from a DVD in seconds; even retrieving a single megabyte of text data from a filing cabinet involves removing thousands of pages and takes days or even weeks. Prior to computers large quantities of data were stored in paper systems, today this seems amazing!



GROUP TASK Discussion

We discussed reasons for maintaining manual filing systems in Chapter 5 (p186). Discuss the bulleted points on page 186 in terms of the storing and retrieving information process.

- **Microfiche**

Microfiche was once commonly used to store the contents of newspapers, magazines and other text and image data. A microfiche card is a small sheet of clear photographic film onto which a miniature image of each page of a publication has been exposed; therefore creating a microfiche card is a similar process to taking a traditional photograph using film. To read microfiche cards requires a microfiche reader; essentially the reader is a magnifying device together with a backlight.

The widespread use of computers has resulted in microfiche being rarely used for storage of new data and publications. However it was once the primary technology for archiving records. For example, births, deaths and marriages records are archived on microfiche, as are images of all the various parts within most old motor vehicles. These microfiche records prove invaluable when tracing family history or attempting to restore an old vehicle. In fact it is common for automotive parts outlets to retain microfiche cards and readers to assist in locating part numbers.



*Fig 6.48
A microfiche reader with
microfiche card inset.*



GROUP TASK Discussion

Identify and describe reasons why microfiche was once such a commonly used medium for archiving data.

- **Libraries**

The purpose of libraries is to store collections of information and to provide efficient processes for its retrieval. Currently much of the information held by libraries is not in digital form, however the catalogues used during retrieval are virtually always computerised. These computerised catalogues allow users to readily search the libraries collection to identify relevant information.

Most libraries now provide Internet access together with a collection of CD-ROM and other computer-based data, the large majority still maintain extensive collections of non-computer data, primarily in the form of books. There are some libraries that are attempting to digitise their entire collection, is this the way of the future or will printed media continue to be used?

Let us examine some of the reasons why non-computer library collections, primarily printed media, are likely to exist for the foreseeable future:

1. Printed media is transportable; it is a self-contained information store that does not require any special technology for retrieval. Hence books, magazines and other printed media can be used on planes, trains, buses, at the beach and even on the lounge room couch.
2. Books and other printed media are ergonomically sound. A single book can be held in the hand and its information read directly without any specialised training. The individual reader decides on their posture and any furniture used whereas the technology associated with digital data is imposed on the reader as a consequence of the required hardware. For printed media our hands, eyes and brain are the technologies used for data retrieval; surely this is a more natural and hence ergonomically sound process.
3. The structure of printed media is intuitive. For example, flicking back and forth through a newspaper to locate items of interest. Commonly we read a small sample of numerous articles, before settling on one of interest. Such random browsing is difficult with digital media where articles are categorised, and one must navigate a logical but complex series of menus and links before the actual article is displayed.
4. Printed media is readily accessible to all; it requires no expensive equipment, no Internet connection and no power. In most countries libraries are government institutions, where books can be borrowed and read by all.
5. Many books do not merely contain information; rather they are works of art. These works of art were created as books and therefore changing the media on which they are delivered also alters the artistic value of the content. An obvious example is a coffee table book; the texture of the paper and binding together with the quality of the photographs is more significant than the raw data within.


GROUP TASK Discussion

“Printed media is machine independent, where the machine is the human body.” Do you agree? Explain your response.


GROUP TASK Discussion

Libraries containing collections of primarily print-based media have existed for thousands of years. Do you think computers and digital data will eventually cause their extinction? Discuss.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH STORING AND RETRIEVING

Social and ethical issues in regard to the storing and retrieving of data are largely concerned with ensuring stored data can only be accessed and used by authorised persons for authorised purposes. Previously in this chapter we examined the use of passwords and encryption techniques, such tools are effective in terms of preventing unauthorised access to data, however they do not protect the data against unauthorised use by authorised users. In Chapter 1 (p19-20) we discussed the security of data and information, and in particular some of the strategies used to address security concerns. In this section we consider examples of particular social and ethical issues arising as a consequence of the storage of data; to assist your discussion of these issues it would be worthwhile reviewing page 19 and in particular page 20.



Consider the following:

Many taxation office employees have access to individual's taxation records; such access is necessary for the completion of their duties. As a consequence it is possible, and perhaps likely, that some of these employees will access and read their friend's tax returns. Such events are difficult to prevent; privacy laws are a deterrent but in this case somewhat difficult to enforce. The breach must first be detected; to do this requires costly and constant monitoring of user's access to individual records. Furthermore, such detailed monitoring of employees raises further ethical issues in regard to privacy.



GROUP TASK Discussion

Imagine you are employed by the tax office. Would you be tempted to read your friends' tax returns? How would you feel about being constantly monitored? Discuss.



Consider the following:

A shop owner would not leave their cash register full of cash and the front door open at the close of business, however many businesses effectively do this with their sensitive data and information. They simply do not recognise the risks and possible effects of unauthorised access to such data.

Furthermore, on the whole even large businesses are unable to detect unauthorised access has even occurred let alone be able to identify the perpetrator. History tells us that all security measures are eventually circumvented; hence regardless of the security systems in place no data is ever truly secure.



GROUP TASK Discussion

Describe possible effects of unauthorised access to sensitive business data.



GROUP TASK Research

Do you think secure public key encryption systems will ever be broken? Use the Internet to gather various opinions to assist your response.



Consider the following:

In 1990 the federal government approved legislation authorising the use of a system known as “The Parallel Data Matching Program”. This legislation was required to override various provisions existing within the *Privacy Act 1988*, in particular to legalise the use and linking of personal data as part of the data matching process. The Data Matching Agency (DMA) was subsequently created to implement the system under the control of Centrelink. The DMA uses data sourced from various government departments and agencies including the Departments of:

- Social Security (DSS)
- Veterans' Affairs (DVA)
- Employment, Education, Training and Youth Affairs (DEETYA)
- Health and Family Services
- the Australian Taxation Office (ATO)
- Centrelink

The data matching process links the individual personal records held by each of these departments in an attempt to identify various fraudulent and illegal activities. In many cases, tax file numbers are used, however it is also common for names, addresses and other private data to be the basis of the data matching process.

The purpose of the DMA is to detect:

- instances of tax evasion
- fictitious or assumed identities
- incorrect payments from support agencies
- inaccurate income disclosures

The DMA has access to and processes data on virtually every single individual resident of Australia; hence initially everyone is a suspect. Should inconsistencies be determined then the presumption is that the person is guilty. This is the opposite of other investigative procedures whereby evidence points to particular individuals who are then investigated in an attempt to gather further evidence.



GROUP TASK Discussion

Do you think the fraudulent activities detected by the DMA justify its extensive use of personal information? Discuss



GROUP TASK Discussion

The Privacy Commissioner has described the above data matching process as “the information society’s equivalent of drift-net fishing”. What do you think the commissioner meant by this statement? Discuss.



GROUP TASK Discussion

Data matching does not merely look for perfect matches; it also links records that have a certain level of similarity. Obviously such matches are often incorrect. List and describe possible consequences of such incorrect matches being assumed to be accurate.



HSC style question:

- (a) Describe the movement of data and the restrictions on the speed of data access between secondary storage, primary storage and the CPU during a process that analyses large amounts of data.
- (b) Explain how binary digits are represented on magnetic tape.
- (c) Outline reasons why most organisations still maintain paper-based filing systems in addition to their computer-based storage systems.
- (d) Describe TWO techniques that aim to secure digital data so it cannot be read by unauthorised users.

Suggested Solutions

- (a) During analysis data is retrieved by the CPU from RAM (primary storage). If the data is not present in RAM then it must be retrieved from secondary storage (usually a hard disk) into RAM. The CPU operates much faster than RAM and RAM operates much faster than secondary storage. Because large amounts of data are being analysed then RAM cannot be filled quickly enough from the slower secondary storage, so RAM cannot keep up with the demands of the CPU. The analysis will only operate at the speed of the slowest device – secondary storage in this case.
- (b) The bits are equally spaced along a track on the surface of the magnetic tape. When the direction of the magnetic field changes the magnetic force is greatest – such points represent binary ones. Binary zeros are represented where the magnetic field does not change direction and hence the force is lower.
- (c) Possible reasons organisations maintain paper-based filing systems include:
 - The existing computer system does include the functionality required to store all the data used by the organisation and it is not cost effective to update to a computer system that can perform these functions.
 - The paper records are not required by other information processes, therefore there is no need for them to be digitised.
 - The original of many documents must be kept for legal reasons. For example, original signatures and seals placed by courts cannot be reproduced digitally.
 - The organisation does not own and cannot justify purchasing the hardware to digitise their paper records.
 - Some data is not suited to computer-based storage. For example, hand written notes, instruction manuals, cash register receipts, etc.
- (d) Passwords can be used so that the system can identify that a user is who they say they are. Permission to read data being based on the user name.
 Encryption involves using an algorithm to scramble the data using a key. The key must be known during the decryption process. Therefore people who do not have the key see scrambled data.

CHAPTER 6 REVIEW

1. Optical media includes:
 - (A) Hard disks and tape cartridges.
 - (B) Hard disk drives and tape drives.
 - (C) CD drives and DVD drives.
 - (D) CDs and DVDs.
2. Flash memory is solid state and non-volatile, this means:
 - (A) it is portable and is difficult to destroy.
 - (B) power is required to maintain the data, but no mechanical parts are used.
 - (C) it contains no moving parts, and requires no power to maintain its contents..
 - (D) it is contained on a microchip and does not require power for data storage.
3. Microfiche stores data:
 - (A) magnetically
 - (B) photographically
 - (C) optically
 - (D) electrically
4. Data is stored on a single continuous track on all:
 - (A) CDs.
 - (B) DVDs.
 - (C) magnetic tapes.
 - (D) hard disks.
5. Electromagnets produce magnetic forces when power is applied, they are used during:
 - (A) optical storing processes.
 - (B) optical retrieving processes.
 - (C) magnetic storing processes.
 - (D) magnetic retrieval processes.
6. The process of linking records from multiple data sources is known as:
 - (A) data retrieval.
 - (B) data matching.
 - (C) record linking.
 - (D) drift-net fishing.
7. Drives capable of storing data on rewriteable optical media:
 - (A) have a laser capable of operating at two different intensities.
 - (B) contain MR material within their read/write head.
 - (C) contain lasers capable of operating at three levels of intensity.
 - (D) produce significant levels of vibration that commonly cause read and write errors.
8. Software that assists the user to copy, delete and paste complete files is known as:
 - (A) a database management system.
 - (B) a tape library.
 - (C) an operating system.
 - (D) file management software.
9. If a collection of data will only ever be encrypted and decrypted by a single machine or user then:
 - (A) a password is sufficient security.
 - (B) single key encryption is suitable.
 - (C) public key encryption should be used.
 - (D) All of the above should be used.
10. The read/write heads in a linear tape drive commonly have each write head positioned between a pair of read heads. Why is this?
 - (A) So the tape can be maintained in the correct vertical position.
 - (B) To enable data to be read, written and then reread without the need to rewind.
 - (C) So data can be written, then verified in either direction.
 - (D) It is cheaper to produce, as such components are part of domestic camcorders.
11. List and describe the main components of each of the following devices:
 - (a) hard disk drive.
 - (b) DVD drive.
 - (c) RAID device.
 - (d) tape library.

12. Discuss each of the following:
- How can flash memory cards be used to help secure data?
 - What are the differences between helical tape and linear tape systems?
 - RAID devices help to protect data against various problems. What are these problems?
 - Tape libraries use multiple small magnetic tapes, in some cases many thousands of them. Doesn't it make more sense to just use much larger tapes?
- 13 For each of the following scenarios:
- Identify and describe any social and ethical issues arising.
 - Suggest a method for securing the data to prevent such issues arising in the future.
- The hard disk on a file server fails. This results in many employees not being able to work for a total of five days whilst a new disk is installed, all the software loaded and configured and finally the data is restored from backups. Most employees are not particularly concerned however management subsequently fires the entire IT department.
 - A mail order business commences trading over the Internet. Unfortunately they begin receiving complaints from customers that their credit card details are being used to purchase goods from all over the world.
 - Various employees, via casual chitchat, form the opinion that their private and sensitive business emails are being read by at least one of the company's network administrators. These suspicions are shared with management, who respond by developing a code of conduct that includes a statement discouraging such activity. However, a sub-clause is included permitting senior management to read any emails as they see fit.
- 14 During the storing and retrieving process the actual data is unchanged, however its physical representation changes and so too does the method of binary representation. In essence the raw data is being reorganised various times as part of storing and retrieving processes.
- Identify and describe each reorganisation of data that occurs during the process of:
- saving a file to a hard disk..
 - retrieving a file from a CD-ROM.
 - saving a file to a RAID device.
- 15 Storage devices are composed of various sub-systems that are ultimately composed of individual hardware and software components. Each component possesses characteristics that make it suitable for its particular task.
- For each of the following components, describe:
- how the component is used by the device during storing and/or retrieval of data.
 - characteristics of the component that makes it suitable for the task it performs.
- Electromagnets
 - Lasers
 - Spindle motors
 - Opto-electrical cells
 - The dye layer within a CD-R.
 - The crystalline layer within a CD-RW.

In this chapter you will learn to:

- select appropriate hardware configurations for a specified type of processing
- edit text data using word processors, desktop publishing, hypertext and database management systems
- edit numeric data using spreadsheets and database management systems
- edit image data using paint, draw and animation packages
- edit video data using animation packages
- edit audio data using mixing software
- diagrammatically represent data processing
- identify examples of potential human bias in data processing
- recognise that processes can overlap, be concurrent or independent or not significant in a specific system

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Processing – a method by which data can be manipulated in different ways to produce a new value or result (eg calculating a total, filtering an email, changing the contrast of an image, changing the volume of a wave file)

Hardware in processing

- hardware with fast processors, a lot of RAM and large storage capacity for image, video and audio processing
- increased processing speed, by:
 - increased clock speeds
 - increased bus capacity
- historical and current trends in CPU development

Software for processing text, numeric, image, video and audio data

Non-computer tools and processing

- documenting procedures to be followed when processing

Social and ethical issues associated with processing

- ownership of processed data
- bias in the way participants in the system process data

Integration of processes

- the interrelationships between the processes in a given system
- one tool (such as software to develop a multimedia presentation) may involve several processes

TOOLS FOR INFORMATION PROCESSES: PROCESSING

The processing information process manipulates data by editing and updating it therefore data is changed and a new value or result is produced. For example, correcting spelling, altering a photo, editing sound effects within a video or changing the price of a product. Each of these processes results in new data different to the original; furthermore this is the only information process that makes such changes.

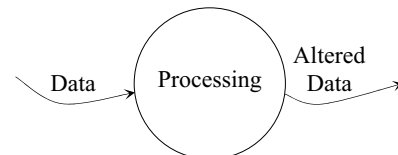


Fig 7.1
Processing alters data.

Processing is central to the operation of all information systems as all information processes depend on processing to achieve their purpose. Hence we commence this chapter with a discussion of the relationship between processing and its integration within other information processes. We then consider the hardware and software used for processing, non-computer techniques for documenting the processes within a system and finally social and ethical issues associated with processing.

THE INTEGRATION OF PROCESSING AND OTHER INFORMATION PROCESSES

Other information processes use data but do not produce or alter it. The organising information process rearranges and represents the data in a different way, but the actual data being organised is not altered. Analysing makes sense of the data transforming it into information; again the actual data is unchanged. Storing and retrieving maintains the data, it changes the data's physical representation, but it does not alter its actual contents. In all of these examples the central processing unit (CPU), or some other processor, is executing instructions to manage and control the operation of the information process.

Surely some data must be produced or altered in order to perform any of the other information processes? This is true, however these changes occur to data used to assist, direct or control the process; no change is made to the data being processed. For example, when using a keyboard the state of the caps lock light is updated each time the caps lock key is pressed; this is a processing task occurring during a collecting process. During a save or storing process the current date and time is saved with each file. Clearly the date and time is new data produced each time the file is saved, consequently a processing information process occurs within the storing process. In essence, every information process encompasses various processing information processes.

If we break down any information process into sufficient detail a multitude of processing information processes are eventually uncovered. This is why they are collectively called information processes; ultimately they are composed of processing processes! In fact every time a processor executes an instruction a processing information process occurs. Data must be produced as a result of each and every instruction executed; if no new data is produced then the execution of the instruction

would be pointless. This is what a central processing unit (CPU) does; it produces new results using instructions and based on input in the form of other data. Sometimes the result changes the input data, sometimes control data is created, but data is always produced.

The above discussion appears to have uncovered a contradiction; other information processes do not produce or alter data yet they are composed of processing information processes which by definition do produce new data. The contradiction is resolved by understanding the nature of the data being created. Other information processes are composed of processing processes that do not alter the actual data within the information system, rather they alter data used to control and direct the operation. Furthermore, the level of detail at which the system is examined determines what is considered actual data. For example, consider the CPU as a complete system. At this level the actual data includes the contents of various registers that store the result of each instruction; in all cases data in one or more of these registers is altered, however this may or may not result in changes to the actual data within the larger information system. For instance, an analysing information process that sorts data causes the CPU to make numerous comparisons, eg. Is *Cow* greater than *Elephant*? Each comparison alters the contents of a register to either *True* or *False*, however the actual data being compared is not altered, rather its position relative to other data items is changed.



Consider the following:

Searching is an analysing information process, therefore we might expect no new value or result to be produced. Consider searching for the number of times the text *at* appears within the sentence *The cat sat on the mat*. At the highest level the input data is the sentence together with the search text *at*; the resulting output from the process is an integer representing the number of times the search text appears; in our example *at* would be found 3 times. Fig 7.2 describes this process using a dataflow diagram. Notice that the search process does not alter the input data, but clearly processing is taking place to produce the new result, namely Number of occurrences. Clearly this analysing process must include at least one processing information process.

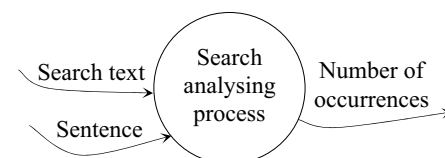


Fig 7.2

Search analysing process

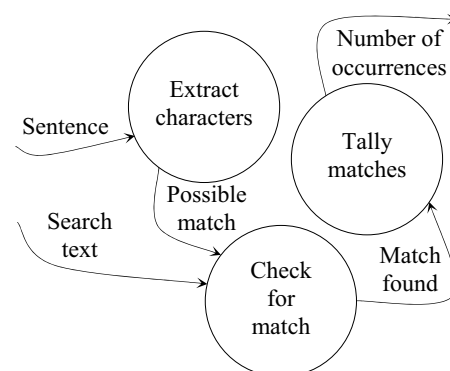


Fig 7.3

Detailed search dataflow diagram.

Now let us break down this search further into a more detailed dataflow diagram (see Fig 7.3). Notice that the various sub-processes use a variety of different data, however the initial input data and the final data output is the same as was indicated on the higher-level dataflow diagram.

Consider each sub-process in Fig 7.3. The Extract characters sub-process generates possible matches. In our *The cat sat on the mat* example it would first output *Th*, then *he*, then *e<space>*, etc... In this example we are searching for a two character string, therefore we successively extract two characters. The data output each time this process executes is different.

Similarly the Check for match process outputs *True* if the Possible match data equals the Search text and *False* if it does not, hence the data output changes. Finally the Tally matches process updates the total matches found each time its input data is *True*. Each of the three sub-processes produces new values or results, however the original input data to the system remains unchanged.

The above example illustrates how information processes are interrelated. Each of the seven syllabus information process includes and/or utilises a mix of other information processes. This is particularly true with regard to processing, as all information processes are ultimately composed of processing information processes.



GROUP TASK Activity

Work through the detailed dataflow diagram above using the sentence *The cat sat on the mat* and the search text *at*.



GROUP TASK Discussion

“All information processes can be broken down into a series of sub-processes. These sub-processes are not necessarily processing processes; they could be any combination of any of the seven information processes. However, ultimately when these sub-processes are completely broken down a series of processing information processes will indeed be the result.” Do you agree? Discuss.



Consider the following:

A database server performs all security, data retrieval and storage tasks for a business. Individual client computers connect to the server via a local area network when they require data; all other processing is performed locally. The local or client personal computers send the database server requests (most often as SQL statements). These requests may involve adding new records, updating existing records or returning sets of records. The central database server performs all the processing necessary to execute the request and respond appropriately. For example, the server's response may include transmission of a set of records to the client or it may simply be a confirmation that the request has been processed.

Examples of typical tasks performed by various users of this system include:

- Creating and printing new invoices.
- Adding and editing customer records.
- Generating monthly and yearly sales graphs.
- Updating the wholesale and retail price of products.
- Posting marketing material to specific groups of customers.



GROUP TASK Activity

For each of the above dot points, identify the information processes likely to be occurring. Identify examples of information processes which are:

- dependent on other information processes,
- independent of other information processes,
- part of a higher level information process, and
- occurring concurrently with other information process.

HARDWARE IN PROCESSING

Processing takes place within the central processing unit (CPU) and for this reason the CPU is a major focus in this section. We consider the CPU and its related components, CPU design factors to increase processing speed, and also various historical and current trends in CPU design.

There are various other hardware tools that influence the effectiveness of the CPU's operation. For example, RAM, secondary storage and also the lines of communication between the CPU and other hardware components. In essence, other hardware must be able to deliver data and instructions to the CPU in sufficient quantities and at sufficient speed if the CPU is to achieve its processing potential. Such issues become more critical when a large amount of data needs to be processed quickly; for example, image, video and audio processing. Finally we examine historical and current trends in processor development.

THE CPU AND ITS RELATED COMPONENTS

In reality there are a large variety of different CPU designs that include different components and different methods of operation, however they are all based on a series of basic components and operational principles. Our aim in this section is to introduce the major components within a simple typical CPU. Throughout later discussions we introduce various modifications used to improve this basic design.

The designs of all CPUs in common usage today are derived from the “stored program concept” originally described by Jon von Neumann in 1945. This concept, as the name suggests, enabled not just data but also program instructions to be stored and hence reused. The “stored program concept” is a logical description of processing, it does not address the physical materials or design required to implement the concept. As a consequence the components within von Neumann's stored program concept are functional components rather than physical components; that is, the components are identified according to the tasks they perform rather than because they are physically separate. So what are these logical or functional components and how do they operate to process data?

Control Unit (CU)

The control unit directs the operation of other components. It interprets instructions and ensures they are performed in the correct sequence and at the correct time. To perform these tasks the control unit includes various temporary storage areas, called registers. The instruction register contains the instruction about to be executed and the program counter contains the address in main memory of the next instruction. The system clock on the motherboard generates equally spaced signals; these signals are used to ensure operations are performed at the correct time. The control unit and the arithmetic logic unit combine to form the central processing unit.

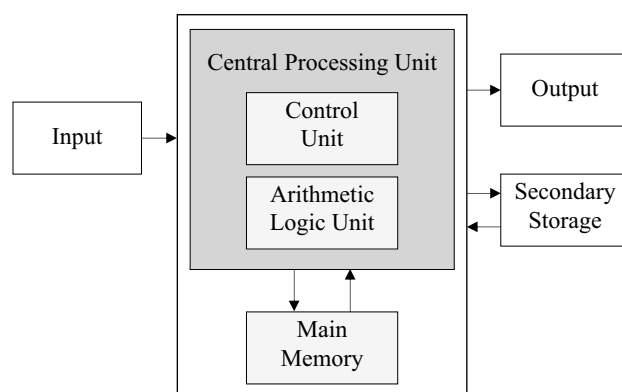


Fig 7.4
Logical components of the von Neumann
stored program concept.

Arithmetic Logic Unit (ALU)

The ALU is where the actual processing of data occurs; in essence the ALU performs all processing information processes. The ALU knows how to execute a relatively small number of instructions; however it only does so when directed by the control unit. There are a variety of general-purpose registers closely associated with and accessible to the ALU. These registers are used to hold data prior to, during and after execution. The accumulator is the most crucial register; it is used during execution and then after execution has completed the new value or result is held in the accumulator.

The word arithmetic refers to basic mathematical operations such as addition and subtraction. The word logic refers to logical operations such as greater than, equal to and less than. Each of these operations is performed on binary data using binary instructions.

Main Memory

Both data and instructions are stored in main memory prior to and after processing. Main memory is primarily RAM, however modern processors also include various types or levels of cache to improve performance; cache is logically part of main memory. Each location in main memory has a unique address. These addresses are used to locate the next instruction to be processed and also to locate data required for processing.

Input/Output

In this course we refer to an input function as a collecting information process and an output function as a displaying information process. Both these functions allow data to enter and exit the system.

Secondary Storage

In terms of processing, secondary storage is used to store and retrieve both data and instructions. The ability to store and retrieve instructions, or programs, in a similar manner to data is the basis of von Neumann's stored program concept. This ability allows computers to easily execute programs multiple times. It is also the reason that computers are multi-purpose machines; that is, they can easily run different programs that solve different problems.



Consider the following:

Text, numeric, audio, image and video data are all processed in binary. Even the instructions used to process data are in binary. Furthermore, both data and instructions are stored identically within both main memory and secondary storage. For example, a word processor file and also a word processor program are both a sequence of binary ones and zeros.



GROUP TASK Discussion

What are the advantages of both data and instructions being stored and processed in binary? Discuss.



GROUP TASK Activity

List specific examples of a binary arithmetic process and a binary logical process. Explain how these processes could be used as part of the processing of text, audio, image and video data.

CPU DESIGN FACTORS TO INCREASE PROCESSING SPEED

There are various techniques that are used to increase the processing speed of CPUs. The most obvious technique is to increase the clock speed, this means the fetch-execute cycle will occur at a faster pace. A second technique is to increase the amount of data processed during each CPU execution cycle. Another possibility is to allow more than one instruction to be executed at any given time. This can be accomplished by having different instructions simultaneously at different points of execution and/or by using multiple processing units. All these techniques increase the amount of data that can potentially be processed by the CPU. For the CPU to realise its processing potential requires that both instructions and data are fetched and the results stored at sufficient speed; the use of various levels of cache aim to solve this problem. In this section, we discuss each of these techniques including their limitations.

Increasing clock speeds

Increasing the speed at which the CPU operates will result in a corresponding increase in the potential speed of processing within the CPU. For example, a CPU processing at a speed of 2GHz has the potential to process data twice as fast as a similar 1GHz CPU. Unfortunately it is not merely the speed of the CPU that affects the amount of processing. There are various other components whose speed of operation is critical if the potential of the CPU is to be realised. Let us consider the relationship between the speeds of these critical processing components.

The system clock is found on the motherboard, its job is to generate equally spaced timing signals. These timing signals determine the speed at which data is moved around the system bus. The system bus is the series of connections on the motherboard that are used to transport data between attached hardware components. *Fig 7.5* shows a typical motherboard from a personal computer; various connectors for the CPU, RAM and other hardware devices can be seen. All the various connections between these devices form the system bus. As the motherboard provides the link between components it makes sense that the speed at which the motherboard operates is critical to the performance of all other components including the CPU.

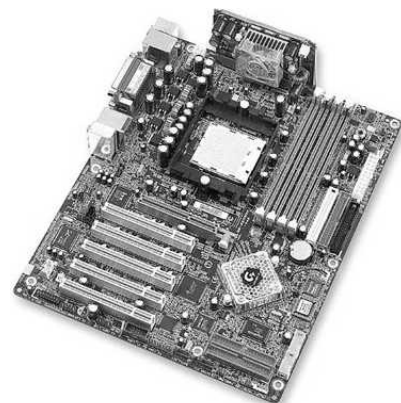


Fig 7.5
Motherboard from typical PC.

Each hardware component is synchronised using the timing signals generated by the system clock. Some components operate at a slower speed and others, such as the CPU, operate at a much higher speed. This means that a multiplier must be used to alter the speed from the system clock so it is more suited to each device. For example, if the system clock is running at 400MHz and a 2GHz CPU is installed then a multiplier of approximately 5 should be used. Essentially, this means data arrives and leaves the CPU at one fifth the speed at which it is processed. If a multiplier greater than 5 were used then the CPU will attempt to process data at a rate greater than that recommended by the manufacturer; this is called “over-clocking”. The result of over-clocking is likely to be processing errors, overheating of the chip and finally crashes and possible damage to the CPU. Similarly, there is little advantage to be



Fig 7.6
CPU fan and heat sink.

gained by merely upgrading to a faster CPU if the speed of the system bus is too slow to supply the faster processor with sufficient data, rather the motherboard, RAM and other components should also be upgraded to faster components.

The speed of CPUs continues to increase and hence so too does the need to dissipate heat. Virtually all CPUs are now cooled using a heat sink together with a fan (see Fig 7.6). A heat sink is commonly a cast aluminium covering containing fins; its job is to radiate heat away from the CPU and into the surrounding air. An attached fan greatly assists this process.



GROUP TASK Activity

Examine the specifications for the motherboard and CPU within your home or school computer. Determine the speed of the system bus and the CPU. Determine the multiplier being used to translate the speed of the system clock to the speed of the CPU.



GROUP TASK Activity

Remove the cover of your home or school computer. Identify the major components, including the motherboard and CPU. Does the CPU have an attached heat sink and fan?

Increasing bus capacity

A bus is a collection of wires used to move data both between and within components. In terms of the system bus on the motherboard more wires means more data can be moved simultaneously. Commonly the size or capacity of a bus is called its width and is expressed in bits. For example, a bus width of 16 bits means there are 16 parallel wires. Each tick of the system clock can therefore move 16 binary digits simultaneously. Clearly for each clock tick a 32-bit bus moves double the amount of data, and a 64-bit bus four times the data. There are different buses used to move data, addresses and instructions both in and out of the CPU, hence many hundreds of connections are required. Fig 7.7 shows an Intel Core 2 Duo CPU containing a total of 478 pins.

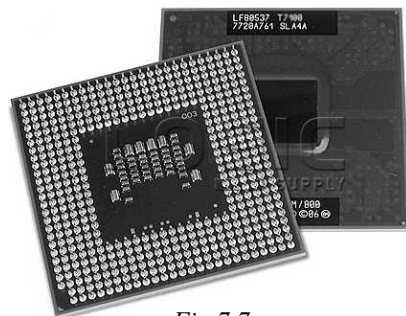


Fig 7.7
There are 478 pins connecting this CPU to the motherboard.

The number of bits processed simultaneously by a CPU is called the word size of the processor. In many cases, the width of the system bus matches the word size of the CPU. The first personal computers used an 8-bit bus and a CPU with an 8-bit word size. At the time of writing (2009), 32 and 64-bit buses and CPUs are common.



GROUP TASK Discussion

Currently there are motherboards available that contain a 32-bit bus but are able to utilise CPUs with a 64-bit word size. Describe why such a design is likely to limit processing performance.



GROUP TASK Research

Research currently available computers to determine the bus capacity of their motherboards and RAM. Do these figures match the word size of the CPU?

Executing more than one instruction at a time

Our discussion of the CPU implied that processing of each instruction must be complete prior to commencement of the next instruction. In reality this is seldom the case. Rather most current CPU chips include multiple processing units and each processor uses a system known as ‘pipelining’. Multiple processors allow different instructions to be executed in parallel. Most current CPU chips include multiple processors and higher end systems include a number of CPU chips on a single motherboard. Pipelining allows multiple instructions to be at different stages of execution at the same time. For example, an Intel Core 2 Duo chip includes two processing cores so both CPUs can be executing different instructions at the same time. In addition, pipelining means that within each CPU multiple instructions can be at different stages of execution at the same time.

Pipelining can be compared to an assembly line. Assembly lines are split into a sequence of stages; each stage performs a specific part of the assembly. For example, a motor vehicle assembly line would contain a stage where the engine is installed, the seats installed, the dashboard installed, and so on. Each car must pass through all stages in sequence before it is completed. The pipeline within each CPU operates in a similar way. Modern CPUs contain around 20 stages within their pipeline hence some 20 instructions are simultaneously being processed at each stage. Each stage of the pipeline completes its task for each tick of the CPU clock. This means a 20-stage pipeline takes 20 CPU clock ticks from when an instruction first enters the pipe until it is completed. However once the pipeline is full instructions are being completed after each clock tick.



Fig 7.8

A car assembly line operates in a similar manner to a CPU pipeline.

There are various problems that arise when using multiple CPUs and pipelining. Firstly, the order in which instructions are executed is often based on the result of a prior instruction. This issue is resolved in two ways. The operating system can allocate a completely different task to each processor using multiple threads and the system can try to guess the correct order of instructions in advance using branch prediction. Secondly, some stages within the pipeline take longer than others to complete. Clearly this would cause a bottleneck, as instructions must wait to enter each of the longer stages. Superscalar architecture is used to overcome such problems.



GROUP TASK Activity

A motor vehicle assembly line has 20 stages and each stage takes 30 minutes to complete. How many cars could be produced during each 40-hour week? What would be the effect of one stage taking 40 minutes to complete? Discuss.

- **Multiple Threads**

Most current software applications are written to support systems that include multiple processors. In simple terms this means different parts of the software application are designed to run as separate independent threads. The operating system can therefore allocate each thread to a different processor in the knowledge that instructions within one thread will not affect the execution of instructions from another thread. Applications that do not include multiple threads are executed on a

single processor (as a single thread). If two such applications are running at the same time then it is likely the operating system would execute each of these applications on a different processor.

- Branch prediction

If an instruction whose result determines the order of execution is yet to be processed or is still within the pipeline then the CPU must make an educated guess as to which instruction sequence is correct. This is known as ‘branch prediction’ and all modern CPUs contain a branch prediction unit to perform this function. Fortunately most computer programs repeat many instruction sequences and therefore most branch prediction units are able to achieve better than 90% accuracy. When an incorrect instruction has commenced execution it must be squashed as soon as the error is detected.

- Superscalar architecture

Each stage of the pipeline takes a particular number of CPU clock ticks to complete. Many stages may take a single tick whilst others take multiple ticks. To avoid bottlenecks a duplicate of the longer stages is used, this is called superscalar architecture. Consider our car assembly line; say installing the engine takes twice as long as each of the other stages. To overcome the bottleneck we add an extra engine installation stage. Now every odd car goes to the first engine install stage and every even car goes to the second engine install stage. The same superscalar system is used within the CPU. Indeed the processors within most CPUs contain multiple ALUs.



Consider the following:

A fictitious CPU chip contains a single processor with a four-stage pipeline. The fetch, decode and store stages take precisely one CPU clock tick to complete, whilst the execute stage takes two CPU clock ticks to complete. *Fig 7.9* describes the progress of a series of instructions, labelled A-J, through the pipeline of this CPU. The first table uses one execution unit and the second table describes the processing using two execution units.

Single Execution Unit	CPU ticks	Fetch stage	Decode stage	Execute stage	Store stage
	1	A			
	2	B	A		
	3	C	B		
	4	D	C	A	
	5	E	D		A
	6	F	E	B	
	7	G	F		B
	8	H	G	C	
	9	I	H		C
	10	J	I	D	

Two Execution Units	CPU ticks	Fetch stage	Decode stage	Execute stage	Store stage
	1	A			
	2	B	A		
	3	C	B	A	
	4	D	C		
	5	E	D	B	A
	6	F	E	C	B
	7	G	F	D	C
	8	H	G	E	D
	9	I	H	F	E
	10	J	I	G	F

Fig 7.9

Table describing the progress of instructions through the CPU pipeline.



GROUP TASK Activity

Work through each of the tables in *Fig 7.9*. What speed improvement is achieved through the use of two execution units?



GROUP TASK Discussion

Our example is vastly simplified, in reality many stages take a different number of ‘ticks’ to complete. How can the number of units required for each stage in the pipeline be calculated?

Adding cache memory

Cache memory is included as an integral part of all modern CPU designs. The aim of cache memory is to provide the CPU with data and instructions more rapidly than would otherwise be the case. As a consequence the CPU is able to process more rapidly.

So what is cache? Cache is a smaller and faster type of storage that is used to improve the speed of access to a larger and slower type of storage. Previously we discussed various implementations of the caching idea. For example, in Chapter 6



Cache

Faster and smaller storage used to accelerate access to slower and larger capacity storage.

(p201) we discussed how web browsers download web pages and other files to the local hard disk. If a requested web page or related file is already on the local hard disk then there is no need to retrieve it from the Internet. Clearly the hard disk has a smaller capacity than the entire Internet, and furthermore it operates much faster than an Internet connection. Hence the local hard disk is being used as a cache between the computer and the Internet. This is but one example; there are many caching subsystems within all computer systems. Many peripherals devices contain their own cache and RAM is used as a cache between the CPU and secondary storage.

In this section we are concerned with caching subsystems between RAM and the CPU. Most modern CPUs contain two levels of built in cache, level 1 (L1) and level 2 (L2). L1 cache has a small capacity, typically just 4 to 16 kilobytes, and operates at virtually the same speed as the CPU. L2 cache is larger, commonly between 128 kilobytes and 1024 kilobytes, and operates at about twice the speed of RAM. Both L1 and L2 cache operate together to speed up access to RAM. At the time of writing a typical personal computer contains around 512 megabytes or more of RAM. This means a 16K L1 cache is used to speed up a 1024K L2 cache, which in turn speeds up more than 512000K of RAM.

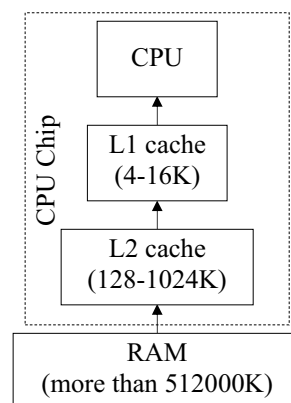


Fig 7.10

The two levels of cache between the CPU and RAM.

How is it possible for such small amounts of memory to accelerate such relatively massive amounts of memory?

Most of the time software executes the same instructions repeatedly. For example, when using a word processor most of the time is spent repeatedly processing input from the user. When we view a video the CPU is decompressing each frame one after the other. Approximately 95% of all instructions are repeated numerous times. These repeated instructions need only be retrieved from RAM once, kept in L1 cache then processed by the CPU multiple times. Furthermore, when instructions (or data) are retrieved from RAM more than just the required instructions are copied to L2 cache. The instructions surrounding those required are also retrieved. As a consequence of the sequential nature of most programs it is highly likely that the next instructions required by the CPU will be within those retrieved.



GROUP TASK Discussion

A program is downloaded from the Internet and executed locally on a personal computer. During this process various different components are used to cache the instructions. Discuss the different levels of cache used and how they interact with each other to accelerate processing.



Consider the following:

Supermarkets use a system similar to cache to maintain stock of items on their shelves. Commonly a small number of each item is actually unpacked on the shelves ready for purchase by customers. Above each set of shelves are boxes containing more of each item. Within the supermarket's storeroom are further boxes. The storeroom is stocked from the supermarket chain's warehouse. And finally the warehouse receives stock from each of the manufacturer's warehouses.



GROUP TASK Activity

Identify aspects of the above scenario that improve the ability of the supermarket to keep their shelves stocked with goods.



GROUP TASK Discussion

Discuss aspects of the supermarket supply system that are similar to caching used within computers and aspects that are dissimilar.

HISTORICAL AND CURRENT TRENDS IN CPU DESIGN

The development of computers is largely based on advances in CPU design. In fact, the underlying technology within the CPU is the most commonly used criteria for categorising computers into various generations. Vacuum tubes being used within first-generation machines (1943-1959), transistors in second-generation machines (1959-1964), integrated circuits in third-generation computers (1964-1972), and finally microprocessors as the defining feature of fourth-generation computers (1972 to present). An example of each of these technologies is shown in *Fig 7.11*. All these components are ultimately switches, where the switch is turned on (or off) using electrical current (or lack of current). Each vacuum tube and transistor is a single switch, whereas an integrated circuit contains many switches and a microprocessor many millions. Today each of the switches within a microprocessor is also called a transistor; an Intel Pentium 4 microprocessor contains approximately 42 million transistors.

Clearly the overriding aim of all CPU design developments is to increase the speed and accuracy of processing. To achieve this aim requires an ever-increasing number of transistors that are able to operate at faster and faster speeds. Unfortunately it takes time for transistors, no matter how small, to perform their job. This problem led to an increase in bus width and then the use of multiple processors to allow many more transistors to perform their tasks simultaneously.

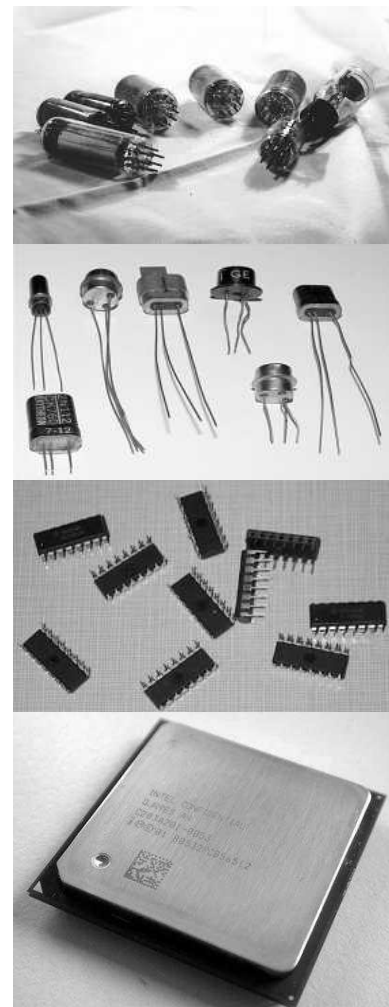


Fig 7.11
From top: vacuum tubes,
transistors, integrated circuits and
microprocessors.



Consider the following historical timeline

1890	Herman Hollerith develops a mechanical and electrically powered tabulating machine for processing US Census data. The design was based on mechanical looms used in the textile industry. Hollerith created the Tabulating Machine Company which in 1924 merged with two other companies to become IBM.
1941	John Atanasoff and Clifford Berry build the first digital computer, called the ABC (Atanasoff-Berry Computer). The ABC was designed specifically to calculate the required trajectories for artillery and missiles.
1943	In January the Harvard Mark I is completed at Harvard University by Howard Aitken. It is the first program controlled calculator and uses paper tape for input. The machine weighs 5 ton and contains 750,000 parts, including more than 3 thousand electrically powered mechanical relays. In December the earliest Programmable Electronic Computer becomes operational. It is called the Colossus and includes approximately 2400 vacuum tubes. The Colossus was designed by Dr Thomas Flowers to decipher German codes during the Second World War.
1946	The ENIAC (Electronic Numerical Integrator and Computer) is completed by John Mauchly and J. Presper Eckert. It weighed some 30 tons, contained approximately 18,000 vacuum tubes and was able to perform some 100,000 calculations per second, an amazing performance in 1946. Unfortunately it required a large room and numerous operators to continually monitor and replace blown vacuum tubes. ENIAC was developed for the military and was used to test theories resulting in the creation of the hydrogen bomb.
1947	William Shockley, John Bardeen and Walter Brattain invent the transistor at The Bell Laboratories in the U.S.A. They later receive the Nobel Prize for this invention. The word transistor is a shortened version of the words transfer and resistor.
1948	Small Scale Experimental Machine (SSEM) or 'Baby' built at Manchester University. This was the first computer to use Von Neumann's stored program concept.
1949	EDVAC (Electronic Discrete Variable Computer) proposed by Von Neumann at Princeton University. EDVAC was completed in 1952 and was the first computer to use magnetic tape for program and data storage.
1951	UNIVAC-1 (Universal Automatic Computer) released commercially by John Mauchly and J. Presper Eckert. UNIVAC was the first general purpose commercially available computer. It was able to process both numeric and text data. The first UNIVAC was sold to the U.S. Bureau of Census during 1951.
1954	Development of FORTRAN, the first high level computer language commences at IBM and is completed in 1957.
1958	The integrated circuit is invented by Jack St Clair Kilby at Texas Instruments. Robert Noyce, who later founded Intel, also worked on the invention.
1960	A variety of high-level programming languages are created during the 1960s including COBOL, ALGOL, APL, PL/1, BASIC and later Pascal. The ability to use high-level languages being a direct consequence of the increased power due to the use of transistors and then integrated circuits.
1971	First microprocessor, the 4004 is developed by Marcian Hoff for Intel. It contained 2300 transistors, had a 4-bit bus width and operated at a speed of 108KHz. The 4004 was used to power a desktop calculator.
1972	Intel releases the 8008 Microprocessor containing some 3,500 transistors and a clock speed of 200KHz.
1974	Intel releases the 8080 microprocessor, an 8-bit processor containing 6,000 transistors and operating at a speed of 2MHz. 8080 chips were used to power the first personal computer called the Altair. The Altair was a kit computer purchased by hobbyists.

1975	First implementation of BASIC by Bill Gates and Paul Allen. Microsoft was formed late in 1975. MOS technologies develops the 6501 and later the 6502 microprocessor. The 6502 chip is later used in the Apple II, Commodore Pet and Commodore 64 computers.
1976	The Apple I is developed by Steve Wozniak and Steve Jobs leading to the founding of Apple Computer. Wozniak and Jobs first used the term 'Personal Computer' to describe the Apple I. The first supercomputer, the Cray 1, is developed. It contained 200,000 integrated circuits.
1977	The Apple II personal computer is released. The Apple II, and its variants, introduced computing to the masses.
1978	Intel develops the 8088 and 8086 microprocessor. The 8086 is a 16-bit processor and the 8088 an inexpensive cut down 8-bit version. The chips contain some 29,000 transistors and can operate at speeds up to 10MHz.
1979	IBM decides to build its own personal computer. They commission Microsoft to develop the operating system. Apparently Bill Gates purchased the rights to an existing operating system from Tim Paterson, which later became MS-DOS. Motorola releases the 68000 microprocessor which is later used in the Apple Macintosh and various other personal computers.
1981	IBM releases the first IBM PC to compete with Apple. The first IBM PC uses an Intel 8088 microprocessor and includes MS-DOS as the operating system.
1982	Intel releases the 80286 microprocessor containing some 134,000 transistors and a clock speed of up to 12.5MHz. Compaq releases their version of the IBM PC. Various other manufacturers enter the market.
1983	IBM XT personal computer released. The machine is based on the 8086 Intel microprocessor.
1984	The Apple Macintosh is released based on the Motorola 68000 microprocessor. The 68000 runs at 8MHz and is able to address 16MB of RAM. IBM AT released based on the Intel 80286 microprocessor.
1985	Intel releases the 80386 microprocessor which contained a 32-bit bus, some 375,000 transistors and was able to run at speeds up to 33MHz. Microsoft releases its first version of Windows. At the time Windows ran on top of MS-DOS and was not widely accepted.
1989	80486 released by Intel. The 486 contained a built-in maths co-processor and approximately 1.2 million transistors. Initial versions ran at 25MHz, however later versions achieved internal speeds up to 100MHz.
1990	Microsoft releases Windows 3.0, 3.1 and finally the popular Windows 3.11.

**GROUP TASK Discussion**

"The development of computers is largely based on advances in CPU design". Do you agree? Discuss using evidence from the above timeline.

**GROUP TASK Activity**

"Hardware and software developments are intimately linked." Identify evidence from the above timeline to support this statement.

**GROUP TASK Research**

Choose one significant individual mentioned in the above timeline. Research this person and present your findings to the class.

SET 7A

1. The actual processing which produces a new value or result occurs in:
 - (A) the control unit.
 - (B) main memory.
 - (C) secondary storage.
 - (D) the ALU.
2. The control unit:
 - (A) directs the operation of the CPU.
 - (B) ensures instructions are executed at the correct time.
 - (C) makes sense of instructions.
 - (D) All of the above.
3. Which of the following is true for most information processes.
 - (A) They operate in isolation.
 - (B) They are composed of other information processes.
 - (C) They dependent on other information processes.
 - (D) Both B and C.
4. Von Neumann's 'stored program concept' essentially means:
 - (A) both instructions and data are stored and reused.
 - (B) instructions are represented differently in main memory compared to data.
 - (C) instructions are performed in a specific sequential order.
 - (D) functional components are not the same as physical components.
5. Which term best describes a software application which can be executed on multiple processors within a CPU chip.
 - (A) Multi-threaded.
 - (B) Multi-tasked.
 - (C) Superscalar.
 - (D) Multi-processor.
6. The ability of a single processor to be executing many instructions at the same time is largely due to the concept known as:
 - (A) caching.
 - (B) over-clocking.
 - (C) pipelining.
 - (D) word size.
7. The accumulator:
 - (A) is located within main memory.
 - (B) contains the address of the next instruction to be executed.
 - (C) is a register used during execution of instructions.
 - (D) is part of level 1 cache.
8. A single processor contains multiple components that perform identical processes. This is likely to be an example of:
 - (A) pipelining.
 - (B) branch prediction.
 - (C) superscalar architecture.
 - (D) parallel processing.
9. Which list of CPU components is correctly ordered from first to fourth generation?
 - (A) transistor, vacuum tube, integrated circuit, microprocessor.
 - (B) vacuum tube, transistor, microprocessor, integrated circuit.
 - (C) vacuum tube, transistor, integrated circuit, microprocessor.
 - (D) vacuum tube, integrated circuit, transistor, microprocessor.
10. Cache memory is:
 - (A) used to speed up access to storage.
 - (B) works best when instructions are often repeated.
 - (C) both faster and smaller than the memory it is designed to accelerate.
 - (D) All of the above.

11. Outline the role of the following functional components:
 - (a) Control unit
 - (b) Arithmetic logic unit
 - (c) Main memory
 - (d) Input/Output
 - (e) Secondary storage

12. A multimedia presentation will be distributed on CD-ROM to promote a new product. The presentation will combine video, images and text into a sequence of slides. The following tasks will be performed during the development of this presentation:
 - I. Images and video clips are collected using a digital still and digital video camera.
 - II. Images are edited, resized and saved as JPEG files using a paint software application.
 - III. The video clips are combined to create a number of compressed MPG files using video processing software.
 - IV. Presentation software is used to create a master slide which includes a vector image of the organisation's logo.
 - V. Text data is entered and formatted on the individual slides.
 - VI. JPEG images are imported onto appropriate individual slides.
 - VII. Links to play the MPG files are created on appropriate slides.
 - VIII. Transitions between slides and navigational elements are added to the presentation.
 - IX. The final presentation is copied onto CD-ROM in preparation for distribution.
 - (a) Identify the significant information process or processes occurring during each of the above tasks.
 - (b) For each information process identified in part (a), outline a "processing" information process that is also occurring.

13. Upgrading to a CPU with a faster clock speed or to a CPU with a larger bus capacity is often seen as a simple way to increase processing speed. However when the CPU is the only component that is upgraded the improvement in processing speed can be less than impressive.
 - (a) Describe how the CPU's clock speed and bus capacity affect processing speed.
 - (b) Identify and describe reasons why upgrading the CPU alone often has little effect on processing speeds.

14. Determine the missing words in each of the following, and then answer the included question:
 - (a) A CPU's _____ contains 20 stages. If the _____ unit is able to correctly determine the next instruction in exactly 90% of cases, then approximately how many instructions will be squashed each second if the CPU processes at a _____ of 2GHz?
 - (b) Frank notices that the first time he runs his word processor after rebooting his computer it takes much longer to start than at other times. You explain to Frank that this is a result of the _____ on the _____. Frank asks you to describe how these components cause this to happen?
 - (c) The CPU retrieves both _____ and _____ from main memory. Main memory includes RAM together with both level 1 and 2 _____. Explain how these components of main memory allow the CPU to process at a greater speed than RAM operates?

15. Research and briefly describe a particular computer from each of the four generations of computers. If possible, locate and print a photograph of each machine.

SOFTWARE APPLICATIONS FOR PROCESSING

In this section we consider examples of software applications used to process data within information systems. Different software applications are available for processing each of the different media types. Furthermore, each of these media types is organised in such a way that it can be most efficiently processed. As a consequence the organising information process is closely linked to the processing information process. Indeed much of Chapter 4 is related not just to the organising information process but also to the processing information process. Hence in this section we refer to software applications originally introduced in Chapter 4. The information in Chapter 4 will be useful when attempting the group tasks in this section.

All data within computers is ultimately processed in binary therefore in all cases processing is altering binary digits. For example, editing the word *dag* to the word *dog* within a word processor is really replacing the binary code for the letter *a* with the binary code for the letter *o*.



GROUP TASK Discussion

Is it possible for data to be edited or updated without first being organised? Similarly, must data be organised after processing? Discuss.

PAINT AND DRAW SOFTWARE FOR IMAGES

Paint software is used to process bitmap images. Hence processing information processes operate on individual pixels. For example, replacing all red pixels with blue pixels. Paint applications provide automated tools that enable the alteration of many pixels based on a single user input. Draw software applications are used to process vector images. That is, they alter the attributes of shapes within the image. For example, the thickness of the line surrounding a circle may be altered or the fill colour changed.



Consider the following:

In Chapter 4, pages 131-135, we discussed the following processes that alter image data:

- Negative function within a paint application.
- Fill operation within a paint application.
- Resizing, stretching or skewing within a paint application.
- Repositioning objects within an image using a draw application.
- Resizing and reshaping objects within a draw application.
- Altering the attributes of multiple shapes at the same time using a draw application.

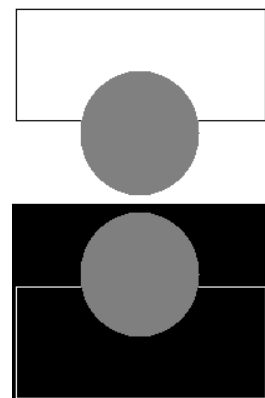


Fig 7.12
Sample image



GROUP TASK Practical Activity

Use the minimum number of steps to create the image shown in Fig 7.12, first within a paint application and then again within a draw application.



GROUP TASK Discussion

Are all the bullet points above examples of processing? In each case, justify your response by identifying the nature of the changed data.

MIXING SOFTWARE FOR AUDIO

Mixing software is used to automate the processing of sound samples. The term mixing refers to the process of combining or adding multiple sounds together; this is the primary task of mixing software. To accomplish this task various other processes are needed. Firstly the volume or level of each sound must be adjusted to suit the requirements of the final mix. Secondly a process is needed to allow parts of a sound to be trimmed or removed. Thirdly a process is needed to remove or filter out unwanted noise. Finally the software should be able to add the component sounds together to form the final sound. This process should include a process for scaling the final sound samples to suit the required range of amplitudes. Each of these processes alters the binary integers used to represent individual sound samples.



Consider the following:

In the above discussion the following processes were mentioned:

- Mixing multiple sounds.
- Adjusting the level of a sound.
- Trimming a sound.
- Filtering out noise from a sound.
- Scaling the amplitudes of a mixed sound.

Each of the above processes was described in Chapter 4, pages 135-138.



GROUP TASK Practical Activity

Use a mixing software application to perform each of the above processes. Explain the changes made to the sound data as a consequence of executing each process.

ANIMATION AND VIDEO EDITING SOFTWARE FOR VIDEO

In most animation and video editing software a project or reference file is first created. This file includes the location of all the various characters, video clips, images and audio clips used within the production. It also includes information in regard to timing and special effects. For example, a typical home movie would likely include a number of video clips that include audio, various titles and also a number of transitions between each video clip. The project file is altered when new items are added or the order of items is changed. Once the project file is complete a separate process is used to create the final animation or video file. The creation of the final file will likely involve altering the format, resolution, frame rate and method of compression.



GROUP TASK Practical Activity

Create a small video using at least three video clips. List and describe the processing steps used to accomplish this task.



GROUP TASK Discussion

Creating a final video based on a project or reference file is an extremely processor intensive process. It takes much longer to perform this process than it takes to play the resulting file. Discuss reasons why this is so.

WORD PROCESSORS, DESKTOP PUBLISHING AND HYPERTEXT SOFTWARE FOR TEXT

The primary purpose of word processors, desktop publishing and hypertext software is to edit and format text in preparation for display. Editing refers to processes that alter the actual words. For example, performing a spell check results in alterations to misspelt words. Formatting refers to processes that alter the way the text is presented, for example changing fonts or adjusting paragraph spacing. Both editing and formatting processes alter the underlying data. Editing alters the actual text data, whilst formatting alters the data describing how the text is to be displayed. Formatting processes can therefore be classified as both processing and displaying information processes. Similarly many editing processes include both analysing and processing information processes, for example performing a search and replace operation.

Word processors place more emphasis on editing processes, whilst desktop publishing software emphasises formatting processes. In fact for most commercial publications the editing is performed using a word processor and then the text is imported and formatted within a desktop publishing application. Similarly hypertext software, such as HTML editors, are used to format text that has first been collected and edited using a word processor. Hypertext or web creation software provides formatting processes to simplify the creation of web pages.



Consider the following processes

1. Copying and pasting a paragraph of text.
2. Performing a spell check.
3. Changing fonts.
4. Opening a text document.
5. Adding a border around a block of text.
6. Specifying spot colours in a desktop publishing document.
7. Correcting the grammar in a sentence.
8. Adding a hyperlink to a word.
9. Rewording a sentence.
10. Placing text into a table using a word processor.
11. Adding page numbers to a document.
12. Counting the number of words in a document.
13. Altering the number of columns on a page.
14. Inserting an image behind text.
15. Inserting a word-processed document into a desktop publishing document.



GROUP TASK Classify

Which of the above processes alter data? Classify these processes as either editing or formatting processes.



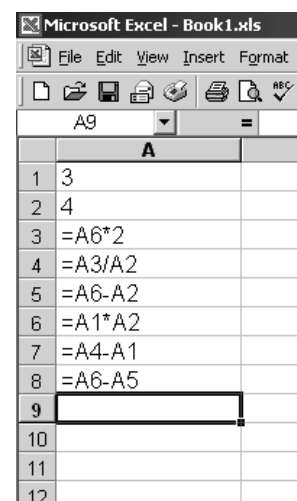
GROUP TASK Discussion

Most of the processes above encompass more than one of the syllabus information processes. Identify the information processes occurring during each of the above processes.

SPREADSHEETS FOR NUMERIC

The primary purpose of spreadsheets is to process numeric data, however most modern spreadsheet applications can also perform various processes on text data. Spreadsheets perform the majority of their processing using formulas entered into cells within the spreadsheet. These formulas include references or addresses of other cells, which in turn contain numeric data known as values, text data known as labels, or further formulas. Clearly editing any of these three data types is a processing task, and furthermore once editing is complete then all cells that reference the altered cell must also be re-evaluated.

The order in which formulas are evaluated is not obvious, or even relevant, to the user. From their perspective all formulas are evaluated immediately. In reality, formulas must be evaluated in a logical systematic order. This recalculation usually occurs automatically each time the data in a cell is altered. Any formulas that either directly or indirectly use the altered data must be recalculated. Older spreadsheets perform recalculation by repeatedly evaluating all formulas in either row or column order until no results change. Currently spreadsheet software is much more efficient, it is able to determine the most logical calculation order based on the references within each formula; this is known as 'natural order recalculation.' For example in *Fig 7.13* the formula in cell A5 contains a reference to cell A6 and cell A6 also contains a formula, therefore cell A6 must be evaluated before cell A5. Natural order recalculation means that no formula is evaluated until all the cells referenced within the formula have first been calculated.



	A
1	3
2	4
3	=A6*2
4	=A3/A2
5	=A6-A2
6	=A1*A2
7	=A4-A1
8	=A6-A5
9	
10	
11	
12	

Fig 7.13
Formulas must be evaluated in a logical order.



GROUP TASK Discussion

Determine the most logical order of evaluation for the formulas in *Fig 7.13*.



GROUP TASK Activity

Create a spreadsheet using the data and formulas in *Fig 7.13*. Alter the values in cells A1 and A2 a number of times. Can you explain the two results displayed in cells A7 and A8?



Consider the following:

Spreadsheet processing does not occur in isolation to other information processes. For example, recalculation takes place as data is entered, or collected. Similarly recalculation occurs automatically when performing 'What-if' analysis. Furthermore, graphs created within spreadsheets change to reflect alterations made to the underlying spreadsheet data.



GROUP TASK Discussion

It is possible to turn off automatic recalculation however this is generally not desirable. Discuss reasons why automatic recalculation is, in most cases, left on. Under what circumstances would it be turned off?

DATABASE MANAGEMENT SYSTEMS FOR TEXT AND NUMERIC

Database management systems (DBMSs) are used to process data contained within databases. Databases can be used to store all the various types of data, however most commonly they hold text and numeric data. Databases contain the data that is processed by DBMS software.

A DBMS provides functions for editing individual data items and also for editing multiple data items based on certain criteria. For example, the price of an individual product may be increased or the price of multiple products may be increased by a certain percentage. From the users perspective data is edited using a data entry form. However behind these forms the software is causing the DBMS to execute a query statement to change the data. Multiple data items can also be edited or updated using queries. The SQL (Structured Query Language) statement `UPDATE Products SET Products.Price = Products.Price*1.1 WHERE Products.Category="p"` when executed by the DBMS results in the price of all products in category p being increased by 10%. In addition, it is also possible to write SQL queries that add or append new records, delete records, or even create new tables and fields.

The editing of data is just one example of the processing performed by DBMS software, there are countless others. Below is a list of some of the DBMS processing tasks we have encountered so far in this text:

1. An ATM confirming your password.
2. Calculating the balance owing on an invoice.
3. Retrieving and displaying a student's timetable.
4. Creating records of all transactions for use as an audit trail.
5. Calculating statistics based on the results of a survey.
6. Formatting dates appropriately in preparation for display.
7. Assigning jobs to particular workers.
8. Calculating tax on employee's weekly salaries.
9. Validating data as it is collected.
10. Maintaining an index within a database table.
11. Searching for websites using an Internet search engine.
12. Synchronising multiple copies of a database using replication.
13. Formulating new information by linking together multiple databases.
14. Scaling marks within a school's database.
15. Controlling simultaneous access to records by multiple users.
16. Assigning data access permissions to users.
17. A database server receiving and responding to client requests.



GROUP TASK Discussion

Identify the data that is being altered by each of the above processes. Which processes may or do cause data held in the database to change?



GROUP TASK Practical Activity

All of the above DBMS processes include various other information processes. Either manually or using a networked database, each student is to record the information processes they think are occurring during each of the above DBMS processes. Analyse this information to assist the class to reach agreement on the information processes occurring.

NON-COMPUTER TOOLS FOR DOCUMENTING PROCESSING

In previous chapters, we considered non-computer tools that are or were once used to perform information processing. Prior to computers, humans performed most processing, with perhaps some assistance being provided by simple tools such as slide rules and calculators. In fact, it is people that have designed the processing performed by computers and it is people who must initiate and manage computer-based processing. In this section we examine techniques for documenting the processing performed by information systems. We first consider documentation designed to assist system participants to direct the processing within information systems. We then consider techniques for representing the processing occurring within information systems.

DOCUMENTING PROCEDURES TO FOLLOW WHEN PROCESSING

Participants are critical elements of all information systems. They initiate and direct the various computer-based tools within the system to perform their processing in the correct order and at the correct time. Clearly it is necessary for participants to understand the procedures or series of steps required to complete tasks within the information system. Documentation specifying these procedures is, not surprisingly, known as ‘procedural documentation’.



Procedure

The series of steps required to complete a process successfully.

As procedural documentation is intended for participants it should be structured in terms of the processes or tasks performed by people. Each task should have a clearly defined purpose. For example, if a user commonly needs to generate and fax statements to individual clients then the procedure necessary to perform this task should be included within the procedural documentation of the system. Such a task is likely to involve initiating a number of system processes, many of which are also used to perform other tasks. Hence procedural documentation is not a description of each isolated process but rather a description of how these processes are used to perform particular tasks.

Procedural documentation for particular information systems is often provided in written form, either as a printed manual or its electronic equivalent. Procedural documentation is also included as part of the online help system within many software applications. In either case it is vital that procedural documentation be continually updated to reflect changes in the information system.

For each identified task procedural documentation should include:

- What the task is and why it is required. In essence a general statement describing the overall process and its purpose. For example, a particular task may be “How to generate and fax individual client statements”. This task is required because individual clients regularly request statements at different times.
- How the task relates to other tasks within the system. For example, commonly orders do not appear on a client’s account until the goods have been despatched. A user preparing client statements must be aware of this.
- Who is responsible for the task and who performs the task? Each task is assigned to a particular participant or group of participants. For example, performing backups may be the responsibility of the system administrator, however an end user performs the actual task.

- When the task is to be completed. Many tasks must be completed at a particular time or under particular circumstances. For example, overdue accounts maybe generated every 30 days, or virus software should be installed prior to new computers being added to the network.
- How to complete the task. This section describes the steps the user must perform to complete the task. In most cases this is the major part of each entry.



Consider the following:

Accounts: Creating a new customer account		
Related tasks: Creating an order. Updating credit limits.	Officer responsible: Accounts Manager.	Frequency: As required.
Task notes: Potential new customers are frequently indicated when no account number is present on a purchase order received via fax, email or mail. A new customer account must be created for all new customers. Cash customers are assigned a zero credit limit, which causes the system to demand prepayment of orders prior to goods being dispatched. Often cash clients are unaware that an account is maintained in their name and hence do not quote their account number. Credit is only made available to customers once supplier references have been confirmed or a history of past cash orders is present.		
Procedure: - Determine that the order is in fact from a new customer. - Enter the customer details via the <i>new account</i> option on the accounts menu. This process will create a new account number for the new customer. - Select <i>find matches</i> on the new accounts screen. This function looks for similar customer details based on phone, fax and address details. - If a match is found then contact the existing customer to resolve the issue. If no clear resolution is determined then the matter is referred to the accounts manager. - If no match then write down the account number and save the record. (Credit limit must be 0). - Contact the new customer by phone. - Inform customer that the order has been received. - Determine if a credit account is required. - If no credit required then redirect call to an orders clerk. Supply the order clerk with the new account number prior to connecting the customer. End of procedure. - If credit is required then go to step 3. - Initiate credit account application. - Explain requirements for opening a credit account as listed on the <i>Credit Account Application</i>. - Write account number on <i>Credit Account Application</i> and forward to customer. - Inform client that current order cannot be processed without either prepayment or waiting for credit approval. - If prepayment is desired for current order then redirect call to an orders clerk. Supply the order clerk with the new account number prior to connecting the customer. - If waiting for credit approval is desired then write the account number and date on the original order together with the words "Awaiting credit approval". When, and if, the application is approved the order is forwarded to an orders clerk. - When the completed Credit Account Application is received follow the procedures described in Accounts: Updating Credit Limits.		



GROUP TASK Discussion

Why is it desirable to have step-by-step descriptions like the one above?
Discuss.



GROUP TASK Activity

Identify procedural aspects of help systems present in a variety of software applications. Is it necessary for organisations to develop their own procedural documentation if they are using these applications as part of their information systems? Discuss.

DIAGRAMMATIC REPRESENTATION OF PROCESSING

Diagrams are able to communicate information more efficiently than text alone. They enable the complex relationships between processes within a system to be understood. Furthermore, diagrams are able to detail the steps required to perform processing more efficiently than a written description.

Throughout the course we have used various techniques for diagrammatically representing information systems. In particular, we have used data flow diagrams to detail the processes occurring and the movement of data. In this section we formally introduce data flow diagrams.



GROUP TASK Discussion

Flick through the text and identify pages that contain diagrams, in particular data flow diagrams. Discuss the essential features of each type of diagram found.

Data Flow Diagrams (DFDs)

DFDs do not attempt to describe the step-by-step detail of individual processes within a system. Rather they describe the movement and changes in data between processes. As all processes produce a new value or result then the data leaving or output from a process must be different in some way to the data that entered or was input to that process. The aim of DFDs is to represent systems by describing each process in terms of the input data and the new data produced or output by the process. For example, a process that adds up numbers receives various numbers as its input and outputs their sum. There is no attempt to describe how the numbers are summed. On a DFD the emphasis is on where the numbers come from and where the sum is headed.

To represent the data moving between processes we use labelled arrows. The label describes the data and the direction of the arrow describes the movement. Processes are represented using circles. The label within the circle describes the process. As all processes perform some action that produces new data then the labels used to describe data leaving a process should be different to those used for data entering the process.

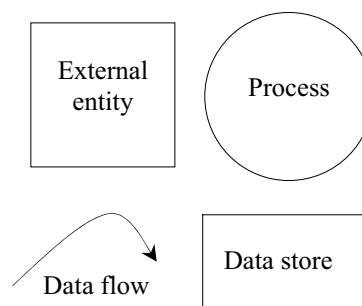


Fig 7.14

Symbols used on data flow diagrams.

In all systems data must enter from outside the system and at some stage the processed data or information must exit the system. The source of data and the destination for information are known as external entities and are represented using labelled squares. External entities are not part of the system; rather they are within the system's environment. External entities are commonly people or organisations that provide data to the system or receive data from the system. An external entity that provides data is known as a source and an external entity that receives processed data is called a sink. It is possible for a single external entity to be both a source and a sink.

The final symbol used on DFDs represents data stores. A data store is where data is maintained prior to and after it has been processed. In most cases a data store will be a file or database stored on a secondary storage device, however it could also be some form of non-computer storage such as a file within a filing cabinet. An open rectangle together with a descriptive label is used to represent data stores.

Commonly a series of DFDs are used to represent an information system. Firstly a context diagram (also known as a Level 0 DFD) is constructed. A context diagram includes a single process together with all the external entities. The single process represents the entire information system. This single process is expanded into a Level 1 DFD containing multiple processes. Each individual process on this DFD can then be broken down into component processes to form a series of even more detailed Level 2 DFDs. Each new level of DFD progressively describes more and more detail.



Consider the following:

Wilbur's Watches employs a buyer and a stores clerk. The buyer consults watch catalogues from various suppliers before sending orders to the suppliers. A copy of each order sent is retained in the purchase order book. The stores clerk takes deliveries of watches, consulting the purchase order book to check that the watches listed on the delivery note have been ordered, and checking the watches themselves against the delivery note. Once all checks have been completed successfully, the stores clerk initials the delivery note, stores it in the deliveries file and forwards the watches to the sales staff for display.

Let us create a series of DFDs to represent Wilbur's Watches purchasing system.

To create a context diagram we need to determine the elements within the scenario that are outside the control of the purchasing system yet provide data to or receive data from the system. The suppliers provide catalogues, receive orders and then deliver the ordered watches. The sales staff receive the watches for display. All these processes are outside the control of the purchasing system and hence should not be included on the DFDs. Rather, the suppliers and sales staff are included as external entities to the system. The catalogues, orders and watch deliveries being the data flowing between the suppliers and the system. The actual watches being the data moving to the sales staff. The completed context diagram (or Level 0 DFD) is shown in *Fig 7.15*.

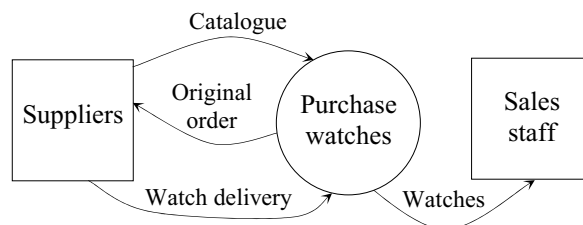


Fig 7.15

Context diagram for Wilbur's Watches.

Now we consider the purchase watches process as a system. We need to determine the processes occurring to complete the purchase watches process, together with any data stores, and data flows. There are essentially two general processes described. One involves the buyer generating orders and the other involves the store clerk taking deliveries. Let us construct a DFD based on these two processes. Notice that the orders are sent to the supplier but they are also filed in the purchase order book. The purchase order book is later used when taking deliveries. As both processes require access to the purchase order book we include it on this DFD as a data store. The purpose of this data store is to allow processing to stop whilst the supplier fills an order. This is often the purpose of data stores, to allow different processes to operate at different times using the same

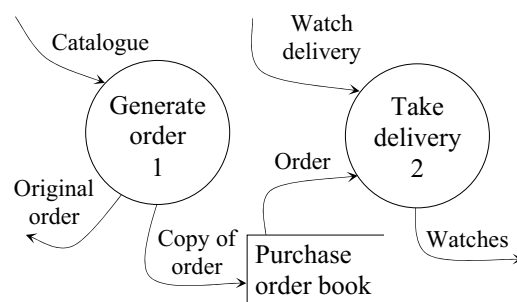


Fig 7.16

Purchase watches DFD.

data. Fig 7.16 shows the completed DFD for the Purchase watches process. Notice that there is no need to show the external entities and that the data flows entering and leaving the DFD are identical to those entering and leaving the Purchase watches process in the context diagram. Each process is numbered to improve readability; the numbers have no meaning in regard to the order of processing. The DFD in Fig 7.16 is called a Level 1 DFD (the initial context diagram was a level 0 DFD). The level is increased for each series of DFD created.



GROUP TASK Discussion

The two processes in Fig 7.16 have both inputs and outputs. Do you think it is necessary for all processes on a DFD to contain both inputs and outputs? Discuss.

The DFDs shown in Fig 7.17 and Fig 7.18 are expansions of the Generate order and Take delivery processes. Each of the sub-processes is given a unique number that identifies the process as belonging to a parent process. For example, the number 1.2 means this is the second sub-process within process 1. Each of these processes could be broken down further. For example, process 1.1 could be broken down into sub-processes 1.1.1, 1.1.2 and 1.1.3.

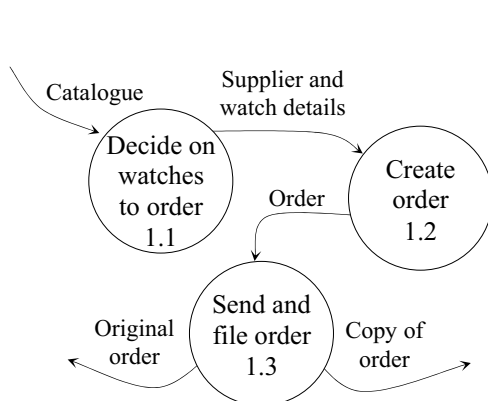


Fig 7.17
Generate order DFD.

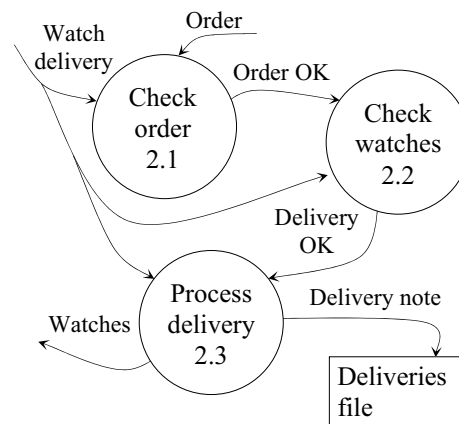


Fig 7.18
Take delivery DFD.



GROUP TASK Activity

The details of the data flows present on DFDs are usually specified within an accompanying data dictionary. As a minimum such a dictionary should include the data type together with a brief description of each data item. Create a possible data dictionary for the above set of DFDs.



GROUP TASK Activity

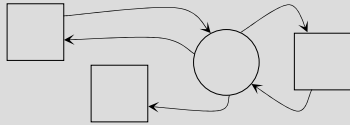
Create a further level DFD for process number 2.3. Notice that the Deliveries file is written to but is not read. Discuss other possible uses for the data within this file.



GROUP TASK Discussion

The system described above does not include computer-based technologies. Discuss aspects of the above system that could be computerised.

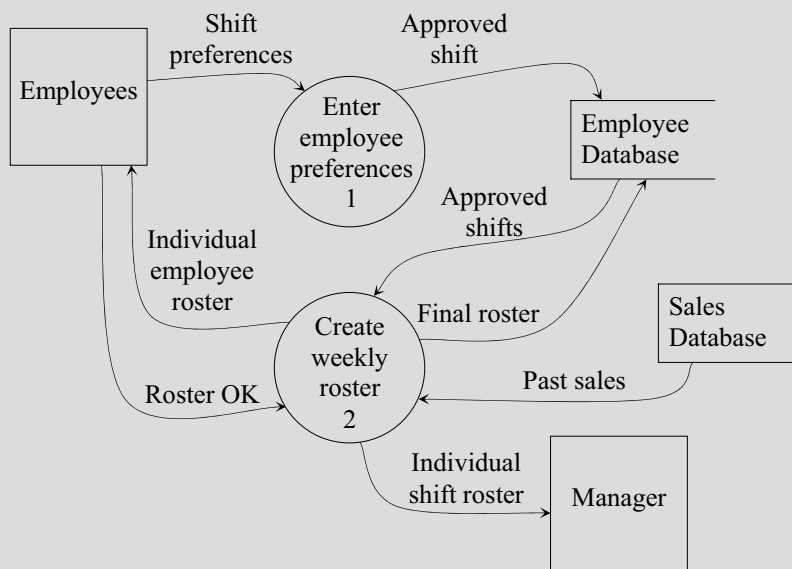
SET 7B

1. Moving a line to a new position within an image is easiest when using a:
 - (A) paint software application.
 - (B) mixing software application.
 - (C) draw software application.
 - (D) animation software application.
2. Applications that process text and numeric data do not often include compression functions because:
 - (A) the data is already tightly packed, so little can be gained by compression.
 - (B) the data cannot be compressed without corruption.
 - (C) there is no standard file formats for storing text and numeric data.
 - (D) it is rare for such files to be of sufficient size to warrant compression.
3. Compression functionality is included within most image, audio and video applications because:
 - (A) without compression the size of such files would almost always be large.
 - (B) decompressing the data as it is displayed is often faster than retrieving and displaying the uncompressed equivalent.
 - (C) such data is more efficiently compressed and decompressed using techniques specific to the type of data.
 - (D) All of the above.
4. Mixing is a process that:
 - (A) adjusts the volume or level of a sound.
 - (B) combines multiple sounds.
 - (C) removes parts of a sound.
 - (D) is used to copy and paste sounds.
5. In regard to the processing of text data, the essential difference between editing and formatting is:
 - (A) editing alters the actual words whereas formatting alters how the text will be displayed.
 - (B) formatting alters the actual words whereas editing alters how the text will be displayed.
 - (C) editing is a collecting information process whereas formatting is part of a displaying information process.
 - (D) editing alters data whereas no data is altered during formatting.
6. Within a spreadsheet cell A3 is edited so it contains the formula $=A1+A2$. If both A1 and A2 also contain formulas then after editing is complete:
 - (A) A3 must always be evaluated first, followed by A1 and A2.
 - (B) only A3 needs to be evaluated as A1 and A2 will already be correct.
 - (C) A1 and A2 must be evaluated before A3 as it is possible that these formulas may indirectly reference A3.
 - (D) no cells need to be recalculated.
7. SQL statements are able to:
 - (A) retrieve records.
 - (B) add or delete records.
 - (C) edit existing records.
 - (D) All of the above.
8. Within a data flow diagram, the symbols used to represent external entities, processes, and data stores respectively are:
 - (A) circles, open rectangles and squares.
 - (B) squares, open rectangles and circles.
 - (C) open rectangles, circles and squares.
 - (D) squares, circles and open rectangles.
9. The diagram below is best described as a:
 

The diagram is a data flow diagram (DFD) with three external entities represented by squares and one process represented by a circle. The process is in the center. Arrows show data flow: one arrow from the top-left square to the process, one from the top-right square to the process, one from the bottom square to the process, and one from the process to the top-right square. There is also a self-loop arrow on the process circle.

 - (A) data flow diagram.
 - (B) systems flowchart.
 - (C) context diagram.
 - (D) data dictionary.
10. On a particular DFD a process is labelled 4.2.5, which of the following is true?
 - (A) This is a level 3 DFD containing at least 5 processes and the associated level 1 DFD includes at least 4 processes.
 - (B) This is a level 3 DFD containing at least 4 processes and the associated level 1 DFD includes at least 5 processes.
 - (C) This is a level 4 DFD containing at least 5 processes and the associated level 1 DFD includes at least 2 processes.
 - (D) This is a level 4 DFD containing 5 processes and the associated level 1 DFD includes at least 4 processes.

11. Describe the changes made to the underlying data when the following processing is performed.
 - (a) A word is removed from a sentence using a word processor.
 - (b) Two sampled sounds are mixed together.
 - (c) A bitmap image is enlarged.
12. Identify appropriate software applications for performing the following:
 - (a) Removing a person from a photograph.
 - (b) Creating personalised invitations to a party.
 - (c) Adding an audio track to a video.
 - (d) Retrieving a list of overdue invoices.
13. Create a step-by-step procedure describing how to send an email message using either your home or school computer. Assume the person who will use your procedure has no knowledge of how to use computers.
14. Refer to the data flow diagram below when answering the questions that follow.



- (a) Construct a context diagram based on the above DFD.
- (b) Explain, in words, the purpose and operation of the system described in the DFD.
- (c) The 'Create weekly roster' process involves the following:
 - Deciding on the number of employees required for each shift based on past sales data.
 - Assigning employees to each shift.
 - Generating and distributing rosters to individual employees. If an employee is not happy with the roster then it may be revised.
 - Generating the final and individual shift rosters.

Develop a DFD to describe the 'Create weekly roster' process.
15. Create a Level 1 DFD to describe the operation of the following bus lane monitoring system:

There are two cameras installed approximately 200 metres apart. Each camera photographs every vehicle travelling in the bus lane. The CPU within each camera analyses every photo to determine the vehicle's registration plate number. Both cameras communicate to determine if a photo of the same car has been taken by the first camera and then by the second camera. If this occurs then the photos and registration number are stored on a hard disk within the first camera. An RTA officer manually replaces the hard disks about once a week. At the RTA office the contents of the hard disk is imported into the RTA database. Then each vehicle's owner and address is determined by querying the RTA database using the registration numbers. Finally fines are generated and posted to vehicle owners.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH PROCESSING

It is people that design, initiate and respond to the results of processing. As a consequence, social and ethical issues arise concerning:

- who can perform processing?
- what type of processing they can perform?
- who owns and has the right to view processed data?
- are the processes that are being performed legitimate?

There are various approaches to dealing with responses to such questions. One extreme is to totally secure the system and only allow authorised personnel to perform a predefined set of processes. The other extreme involves total flexibility, whereby participants are able to design and initiate their own processes as they see fit. Clearly the approach used depends on the individual information system, and in particular the sensitivity of the data being processed.

In this section, we examine some general areas that are often the reason social and ethical issues arise and thus require consideration for most information systems.

OWNERSHIP OF PROCESSED DATA

Generally the creator or collector of data is considered to own the data. However, what if significant processing is performed on the data, should ownership of the resulting data remain with the original owner? The answer to this question is unclear, as it varies depending on the nature of the data and the effort expended to process the data. For example, if a movie on videotape is converted to DVD format then clearly ownership should remain with the original creator. The effort to actually create the movie far outweighs the effort required to alter its format, and furthermore copyright laws cover movies. However, what if data is obtained from the Bureau of Statistics and is subsequently processed to create various statistical graphs for an organisation? In this case, one could argue that significant effort is involved in processing the data and therefore the resulting graphs should be owned by the organisation that performed the processing.



Consider the following:

- You take a photograph and email it to your friend. Your friend then uses the photograph to create a simple animation.
- Your teachers process all marks in all assessment tasks to create a total assessment mark for each student in each course.
- The RTA collects data on all registered vehicles. This data is forwarded to the police who are then able to link this data to an individual's police record.
- A sound technician records a variety of different animal noises. A small selection of these sounds is later used within a computer game.



GROUP TASK Discussion

Identify the unprocessed data and the processed data in each of the above scenarios. In each case, discuss who you think should own the processed data. Justify your responses.

BIAS IN THE WAY DATA IS PROCESSED

In Chapter 3, we defined bias to be ‘an inclination or preference towards an outcome.’ and as something that ‘unfairly influences an outcome’. Bias can influence data as it is collected, and it can also affect the way data is analysed and processed. The aim is to exclude any form of bias from all information processes.

Bias is a natural human condition. We all have inclinations or preferences that cause us to prefer certain outcomes over others. It is often true that we expect a certain result when processing data. Such expectations should not be allowed to influence the processing taking place.



GROUP TASK Discussion

Computers are just machines, therefore computer-based processing cannot possibly be affected by bias. Do you agree? Discuss.



GROUP TASK Discussion

List and describe at least three examples of information processing where a particular result is expected. Do you think the expectation has any effect on the processing taking place? Discuss.

CENTRALISED VERSUS DISTRIBUTED PROCESSING

Centralised processing uses a central computer to perform much of the processing, this greatly simplifies the control of processing tasks. The central computer can be used to restrict the processing which can be performed by users. Furthermore, access to data can be strictly controlled. Professional information technology personnel manage most centralised systems; a major part of their job is ensuring the correct and secure operation of the system. This involves installing, and perhaps developing or updating software as well as assigning permissions to users. These people are more likely to understand the technical requirements needed to ensure security is maintained. However they are unlikely to have a deep understanding of the specific needs of particular groups of users. Distributed systems aim to empower users by allowing them to process data on their own computers to meet their specific needs.

Unfortunately as security increases flexibility decreases. Hence many of the advantages of centralised systems are also disadvantages. Similarly the flexibility inherent in distributed systems makes securing such systems difficult. If a group of users within a large centralised system wish to perform some new type of processing they must work through the organisation’s IT department. Any software tools needed must be approved, perhaps developed and then finally installed on the central computer. Within large organisations the IT department receives numerous requests. These requests must be prioritised, and those that are not critical may take months or even years to implement. Clearly most users would not bother working through such a procedure unless the need is both critical and long term. Distributed systems allow users the flexibility to develop, modify and install their own software tools. Hence even simple one off processing tasks that are needed at short notice can be performed.

Our discussion implies that all distributed systems provide total flexibility with poor security and all centralised systems are secure but inflexible. In reality, this is not the case, rather degrees of security and flexibility are possible in both types of systems. Centralised processing naturally facilitates security of processing and data access, whilst distributed processing naturally facilitates flexibility of both processing and data access. Securing a distributed system or providing flexibility within a centralised system is of course possible, it is just more difficult to implement.



Consider the following:

Client-server processing is a form of distributed processing, however it allows strict control over the processes performed by the server. Each of the following is an example of client-server processing:

- A DBMS running on a server provides and controls access to data for a variety of different software applications running on different client machines.
- A web server sends web pages across the Internet to client machines. Access to some web pages requires the end user to enter a valid user name and password.
- A local area network (LAN) within a home has a single modem. This modem is normally shared with all other machines on the LAN.



GROUP TASK Discussion

Security is just one of the reasons that servers are used. Identify and describe other reasons for a server being used for processing in each of the above scenarios.



GROUP TASK Discussion

In each of the above systems a server provides resources to other client machines. One could argue that these servers result in the system being as inflexible as centralised systems. Do you agree? Discuss.



Consider the following:

Large centralised computer systems are commonly contained within secure air-conditioned premises. Such facilities require protection against:

- Deliberate acts such as theft and vandalism by both authorised and unauthorised personnel.
- Disruption or failure of power or communication into and out of the premises. Even the failure of power to secondary systems, such as air conditioning and alarms has the potential to halt the complete system.
- Breakdown of components. Redundancy should be built into the system and replacement components should be readily available. Old and specially built components are particularly at risk.
- Natural causes such as lightning, water and fire.

Distributed systems counteract such problems by storing and processing data in many different physical locations.



GROUP TASK Research

Research strategies and techniques used by large organisations to protect their central computer facilities from the above threats.



GROUP TASK Discussion

The above information seems to indicate that centralised systems are not as secure as distributed systems. Discuss.

CHAPTER 7 REVIEW

1. The most identifiable characteristic of a processing information process is:
 - (A) new values are changed into information.
 - (B) new values or results are produced.
 - (C) data is rearranged.
 - (D) results are represented differently.
2. The processing information process:
 - (A) edits data.
 - (B) updates data.
 - (C) alters data.
 - (D) can perform all of the above.
3. Execution of an instruction by the CPU:
 - (A) does not always alter data.
 - (B) always alters data.
 - (C) only alters the input data.
 - (D) only alters data used to control or direct the system.
4. Which of the following is true for all information processes?
 - (A) Each information process occurs in isolation to each of the other information processes.
 - (B) The actual collected data is always altered.
 - (C) They are all examples of processing information processes.
 - (D) When examined in sufficient detail they are ultimately composed of processing information processes.
5. The component commonly used to define second generation computers is the:
 - (A) integrated circuit.
 - (B) transistor.
 - (C) vacuum tube.
 - (D) microprocessor.
6. During execution of a software application, instructions that are about to be executed are stored in:
 - (A) the ALU.
 - (B) secondary storage.
 - (C) main memory.
 - (D) the system bus.
7. Removing a shape from an existing image is easiest if the image is stored as a:
 - (A) bitmap.
 - (B) JPEG file.
 - (C) vector file.
 - (D) sequence of slides.
8. Most CPUs process at greater clock speeds than the system clock. One component that makes this possible is called:
 - (A) the ALU.
 - (B) cache memory.
 - (C) main memory.
 - (D) the system bus.
9. The ability of CPUs to allow multiple instructions to be at different stages of execution at the same time is a result of:
 - (A) increasing clock speeds.
 - (B) increasing bus capacity.
 - (C) pipelining
 - (D) parallel processing.
10. A user removes results they did not expect to be present from the data before calculating the average. This is an example of what type of issue?
 - (A) bias.
 - (B) ownership.
 - (C) security.
 - (D) privacy.
11. Define the following terms.
 - (a) CPU
 - (b) RAM
 - (c) bus capacity
 - (d) clock speed
12. Only processing information processes produce new values or results, yet all information processes are composed of various processing information processes. Explain how this apparent contradiction is resolved.
13. Describe the operation of:
 - (a) branch prediction
 - (b) cache memory
 - (c) pipelining
14. Create a table describing the purpose of each symbol used on data flow diagrams
15. An increase in processing speed often results when extra RAM is added to a system. Discuss reasons why this speed increase occurs.

In this chapter you will learn to:

- differentiate between the requirements for a local area network and a wide area network
- transfer numeric, text, image, audio and video data and discuss the time to transfer and required bandwidth
- describe concepts of downloading, uploading and streaming
- demonstrate sending and receiving mail, with attachments, over an e-mail system
- select a relevant technology for a given situation to allow computers to transmit and receive data or information
- compare and contrast computer and non-computer based communication systems
- describe and employ net-etiquette when using the Internet
- predict and discuss possible future trends in communications and the impact they are likely to have on the transmitting and receiving of data/information

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams

In this chapter you will learn about:

Transmitting and receiving – the process that transfers information and data within and between information systems

Hardware for transmitting and receiving

- communications within a computer between peripheral devices and the CPU via buses
- the role of modems, including modulation demodulation
- local area networks and wide area networks

Software for transmitting and receiving

- communications packages
- transmitting and receiving text, numeric, image, audio and video
- electronic mail and its operation

Non-computer tools for transmitting and receiving, such as:

- mail, phone and fax
- radio and television (transmit only)

Social and ethical issues associated with transmitting and receiving

- accuracy of data received from the Internet
- security of data being transferred
- net-etiquette
- acknowledgment of data source
- global network issues, time zones, date fields, exchange rates
- changing nature of work for participants, such as work from home and telecommuting
- current developments and future trends in digital communications, radio and television
- the impact of the Internet on traditional business

TOOLS FOR INFORMATION PROCESSES: TRANSMITTING AND RECEIVING

The transmitting and receiving information process transfers information and data within and between information systems. This transfer of data occurs between components within a single computer, such as the transfer of data between RAM and the CPU. It occurs whenever peripheral devices are used, such as keyboards, printers and modems. It also occurs between computers when communicating using local areas networks and wide area networks such as the Internet.

For communication to take place both transmitting and receiving must occur successfully. Transmitting involves the sender encoding the message and transmitting it over the medium. Receiving involves the receiver understanding the organisation of the encoded message and decoding it into a form suitable for its use. In essence both encoding and decoding are organising information processes. Encoding organises the data into a form suitable for transmission along the communication medium. Decoding changes the organisation of the received data into a form suitable for subsequent information processes. For example, in Chapter 3 we examined how keyboards encode each keystroke into an electrical signal representing the scan code of the key pressed. This signal is transmitted down the interface cable. The computer receives the signal and decodes it into its corresponding ASCII code.

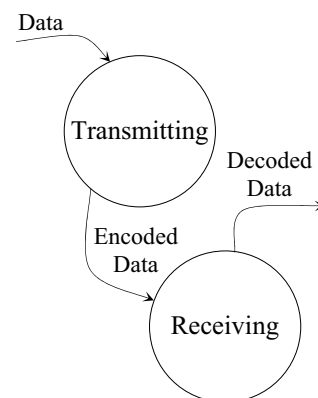


Fig 8.1
Transmitting encodes data
and receiving decodes data.

Transmitting and receiving information processes are an integral part of all other information processes. Whenever communication between hardware components occurs transmitting and receiving information processes are also occurring. As all information processes are performed using a variety of different hardware tools then it follows that transmitting and receiving processes must also be occurring.



Consider the following:

- Scanning an image using a flatbed scanner.
- Creating a graph using a spreadsheet.
- Backing up data to magnetic tape.
- The CPU executing a machine language instruction.
- Surfing the web using an Internet browser.



GROUP TASK Discussion

Identify and discuss transmitting and receiving information processes occurring during each of the above processes.

COMMUNICATION CONCEPTS

Successful communication requires that both sender and receiver agree on the method of data transmission. The agreed method must also work for the medium being used. There are many different characteristics used to determine the method of data transmission. Often the characteristics of one type of communication appear to contradict the characteristics of another. For example, a 32-bit system bus contains 32 physical connections and always transfers exactly 32 binary digits at any one time. However, broadband communication mediums use a single physical connection yet they are able to transfer many bits simultaneously. Understanding the underlying concepts is necessary to resolve such apparent anomalies.

In this section we explain concepts central to understanding the communication process, namely:

- Uploading and downloading
- Serial and parallel
- Simplex, half duplex and full duplex
- Synchronous and asynchronous
- Measurements of speed (bps, baud and bandwidth)

Each concept describes a characteristic that must be agreed upon for successful communication to occur. These concepts underpin the operation of hardware and software used for transmitting and receiving.

UPLOADING AND DOWNLOADING

Upload and download refer to the direction in which a transmission occurs. Commonly these terms refer to a little/big scenario, such as transfers between a single computer and a larger system such as a server or network. Uploading is the transmitting of data from the single (little) computer to the (big) server or network. Downloading occurs in the opposite direction – the little computer receives data from a bigger system. In both cases it is generally the little computer that initiates the transfer. For example, the files required for a webpage are uploaded from a personal computer to a web server on the Internet. Or a PDF file is downloaded from a file server to a personal computer. In each example the personal computer initiates the transfer however the files are transferred in different directions.

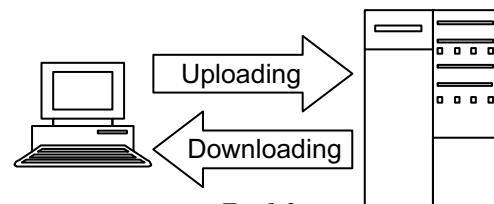


Fig 8.2
Uploading and downloading.

Many connections to the Internet, particularly home plans, have different speeds for downloads compared to uploads. For example, a 1500/256 ADSL connection means the maximum download speed is 1500kbps and the maximum upload speed is 256kbps. The A in ADSL means asymmetrical which means different download and upload speeds. The terms are also used to measure the amount of data transferred. Figs 8.3 shows the breakdown of download and upload usage for a Bigpond user's account.

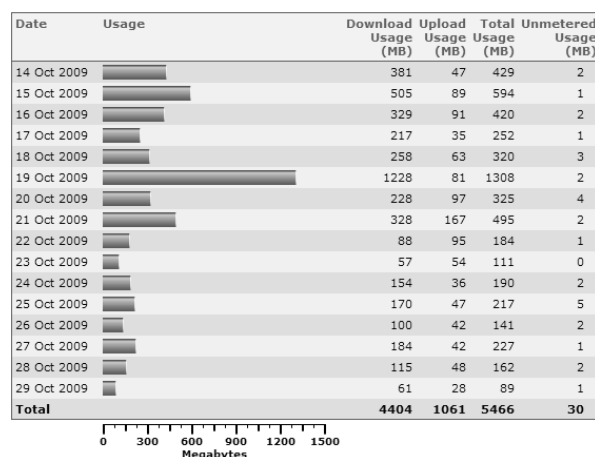


Fig 8.3
Download and upload chart for a Bigpond user.

SERIAL AND PARALLEL

Serial transmission can be likened to a single lane road. Cars travelling in the same direction are all one behind the other. Parallel transmission is more like a multi-lane road. Cars travelling in the same direction can be side by side. Our parallel analogy is even more accurate if we include the rule that cars must travel side by side and furthermore this is to be accomplished by everyone travelling at the speed limit rather than observing cars on either side. In our analogy the lanes represent the channels used for transmission and the cars represent the individual units of data, commonly individual binary digits.

Some forms of serial transmission use a single channel for both transmitting and receiving, whilst others use two separate channels. Now what do we mean by a channel? A channel could be an individual physical connection, such as a wire or an optical fibre, or it could also be a particular range of frequencies within the physical connection. For example, a older RS232 serial ports support a single device and use a single connection for sending and a separate connection for receiving, each of these connections is a single channel. On the other hand a USB (Universal Serial Bus) port can support up to 127 devices. USB ports contain just two data wires, which together form just one communication channel. Messages from and to all connected devices are transferred in both directions using this single channel.

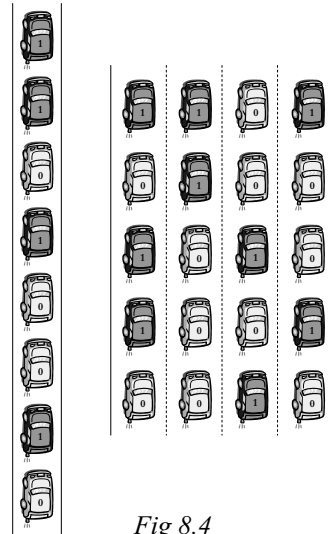


Fig 8.4

Serial transmission is like a single lane road (left) whilst parallel transmission is like a multi-lane highway (right).



Fig 8.5

Connectors for an RS232 serial port (left) and a USB port (right).



GROUP TASK Discussion

The RS232 connector in Fig 8.5 has 9 pins and the USB connector contains 4 contacts. Two wires are required to complete the circuit for each data connection. What is the purpose of the other wires?

Today parallel transmission is seldom used for communication outside of an individual computer. In fact, the majority of computer systems only use parallel communication between components that are either part of the motherboard or are connected directly to the motherboard. In the past most computer's included a parallel port used to connect printers and scanners, today these connections have been largely replaced by serial USB connections.

Intuitively it would seem that parallel transmission should be significantly faster than serial transmission. For instance, in our road example it seems clear that cars will arrive at their destination much faster on a multi-lane road

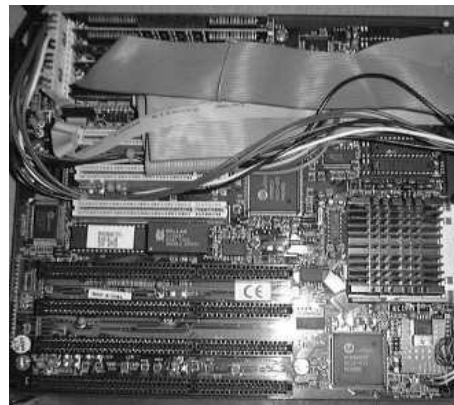


Fig 8.6

Parallel communication is used between components on the motherboard.

than on a single lane road. However, in reality, parallel transmission is only used over small distances and even then the transmission medium must be of extremely high quality. Let us examine reasons for this apparent contradiction.

- Obviously, many more wires are required for parallel transmission than for serial. The extra cost becomes more and more significant as the distance increases. Over short distances, such as within a computer, the extra cost is minor compared to the increase in performance. Over larger distances, such as for networks, the extra cost is not justified and furthermore accurate data transfer is difficult to achieve.
- Data must be assembled into groups equal to the number of parallel wires. For example a standard parallel port has 8 data wires. If the system bus is 32-bits wide then each 32-bit word needs to be split into four groups of 8-bit bytes. Only once a complete 8-bit byte has been assembled can it be transmitted. When using serial transmission each bit can be fired off more rapidly.
- In our multi-lane road example, as the length of the road increases it becomes more and more difficult for the cars in each group to remain precisely side-by-side. The same problem occurs with parallel transmission. As bits travel down the wire they are influenced by external environmental factors. The result being that not all bits arrive at the same time, this phenomenon is known as ‘data skew’.



GROUP TASK Discussion

“The problems with parallel transmission become more and more significant as the distance and speed of data transfer increases.”
Do you agree? Justify your response.

SIMPLEX, HALF DUPLEX AND FULL DUPLEX

Simplex means communication occurs in one direction only. Television and radio are examples of simplex communication. The TV or radio station is always the sender and your television or radio is always the receiver. A single channel is used to transfer data in a single direction. If a connection is purely simplex then the receiver can never provide feedback to the sender. In essence the sender is blindly transmitting data with no way of knowing if the data was in fact received.

In terms of the information processes present within an information system, both collecting and displaying include essentially simplex transmitting and receiving processes. For example, when collecting data using a keyboard, the data being collected is transmitted from the keyboard to the receiving computer. Similarly when displaying, the data to be displayed is transmitted from the computer to the receiving display device. In reality a small amount of data is often returned in the other direction. This data is used to control the operation of the collection or display device or to ensure the accuracy of the data transfer. For all practical purposes collecting and displaying utilise simplex communication.

Half duplex means communication can occur in both directions, but never at the same time. A walkie-talkie or two-way radio operates in half duplex using a single channel. Only one person is able to transmit at a time, either data is travelling from person A to person B or it is travelling from person B to person A. In essence, person A and person B must take it in turns to speak.

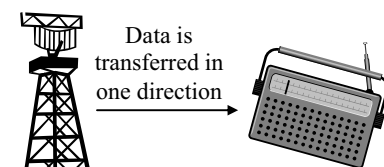


Fig 8.7
Radio is an example of simplex transmission.

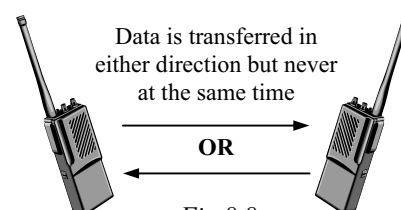


Fig 8.8
Walkie-talkies use half-duplex transmission.

Walkie-talkies are designed to be either transmitting or receiving; they physically cannot do both at the same time. However, between computers half duplex communication is commonly used even when simultaneous two-way communication is possible. For example, client-server processing utilises half duplex; a client sends a request and then the server sends a response. Either the client is transmitting or the server is transmitting but not at the same time. The physical connection and the network software may well support simultaneous communication in both directions, however individual processes performed by software applications commonly use a half duplex mode of communication.

Full duplex, or just duplex means data can be transferred in both directions at the same time. Telephones operate in this way using a single channel; it is possible for both parties to talk at the same time. Full duplex communication requires either a single channel that can represent transmissions in both directions, or two separate channels, one for each direction.

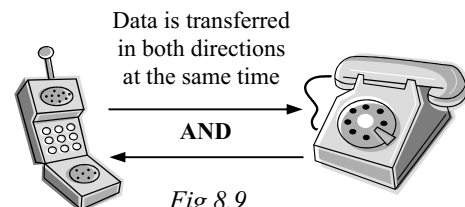


*Fig 8.9
Telephones use full-duplex transmission.*

In theory a full duplex link should be able to transfer double the amount of data compared to a similar half duplex link. For example, a 100Mbps half duplex LAN connection is capable of transferring approximately 100 million bits every second, however a 100Mbps full duplex connection can transfer 100 million bits per second in both directions, hence in total some 200 million bits can be transferred per second.



Consider the following:

Each of the following processes includes the transmission of data:

- A courier delivering a parcel.
- The control unit retrieves the next instruction from RAM.
- A browser requests and receives a web page.
- Mouse movements being sent to a computer.
- Downloading email from a mail server.
- Playing an interactive game over the Internet with many other players.



GROUP TASK Discussion

Describe the transmission used in each of the above scenarios as simplex, half duplex and/or full duplex. Discuss your answers.

SYNCHRONOUS AND ASYNCHRONOUS

The term synchronous means events happen or occur in real time, however a more technical meaning requires that events occur in time with each other. Asynchronous is the opposite; it implies pauses occur between events, such as between the sending and receiving processes. Or more technically, asynchronous means sending and receiving do not occur in time with each other. There are two common uses of the terms corresponding to these slightly different meanings. The first is from the user's perspective and is not restricted to the transfer of digital binary data. The second, rather more technical usage refers to sending and receiving occurring in time with each other and is relevant only to the transfer of binary data.

Synchronous and asynchronous from the user's perspective

From the user's perspective synchronous is commonly used to describe communication between people that occurs in real time. For example, a chat room, instant messaging system or even a telephone call, provide simultaneous communication between two or more users. The users participate in a real time conversation, each user immediately sees or hears what other users are typing or saying. Asynchronous communication between users includes pauses between sending and receiving. Email and traditional mail are both examples, one user sends or posts their message and at some later time the receiver views the message. The communication between users is not continuous; rather the receiving user views the message some time after the message was transmitted.



Consider the following:

Various forms of communication are used within a typical classroom. The teacher may stand up the front and lecture the class, they may ask students questions or perhaps initiate a class discussion. Some classes may be split into groups, where each group works independently to complete a task. Other classes may require students to work on tasks alone, such as completing an examination.



GROUP TASK Discussion

Identify and describe examples of synchronous and asynchronous communication occurring within classrooms.

Synchronous and asynchronous transfer of binary data

To transmit binary data requires at least two physical states. At this stage we restrict our discussion to exactly two states representing the binary digits 1 and 0. For example, +5 volts may represent a 1 and –5 volts a 0. The signal changes between these two states. To accurately decode the signal requires the receiver to sample the signal using precisely the same timing used by the sender during encoding. If both the sender and the receiver use a common clock then transmission can take place in the knowledge that sampling is almost perfectly synchronised with transmitting. This is the most obvious method of achieving synchronous communication. For example, the system clock is used during synchronous communication between components on the motherboard. Unfortunately, the use of a common clock is rarely a practical possibility. As a consequence, other techniques must be used in an attempt to bring the receiver into synch with the transmitter. Asynchronous and synchronous are terms describing two strategies for resolving this problem.

Before we discuss the detail of asynchronous and synchronous techniques we need to understand the nature of a typical binary signal as it arrives for sampling by the receiver. In *Fig 8.10* the solid lines represent signals, and the dotted lines represent the points where the receiver examines the signal. Both signals shown are supposed to represent the same sequence of bits, however the bottom signal is not received correctly. The top signal shows the perfect situation, each bit in the signal is equally spaced and all bits are examined precisely in the centre. The bottom signal includes

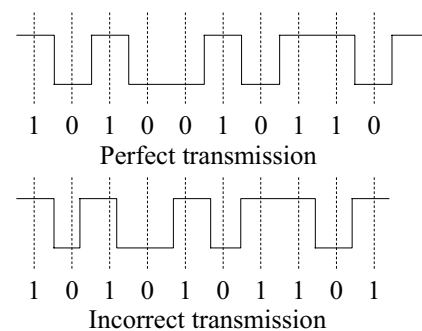


Fig 8.10

A perfectly synchronised and incorrectly synchronised transmission.

variation in the distance between bits, the receiver has not compensated for these variations and consequently the data received contains errors. However the errors only begin to emerge after a number of bits have indeed been received correctly. Our example shows an extreme variation in spacing; in reality such errors are unlikely to occur quite so rapidly. Both asynchronous and synchronous techniques aim to overcome these problems using quite different strategies.

Asynchronous transmission does not try to synchronise the receiving clock with the transmitting clock at all, rather it just detects the start of the data and hopes for the best. Because of this ‘hope for the best’ strategy asynchronous transmission only works successfully when small amounts of data are being transferred at relatively low speeds.

In practice most asynchronous communication transfers single bytes of data, which commonly correspond to individual characters. A single transfer usually contains just 10 bits comprised of a start bit, 8 data bits, and a final stop bit. However an extra stop bit and a parity bit for error correction are often included. The receiver detects the change in signal caused by the start bit and activates its clock. It then commences

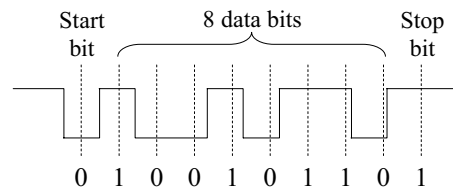


Fig 8.11

Asynchronous communication using 10 bits to transfer each byte of data.

receiving the data, the stop bit indicates the end of the data and returns the signal to its original idle state. The clock rate of the receiver need only be approximately equal to the clock rate of the transmitter. In Fig 8.11 the receiving clock is slightly slower than the transmitting clock, yet all 8 bits are correctly received. If the data had been much longer than 8 bits in length then errors would have begun to occur.

Asynchronous communication is also known as ‘start stop’ communication due partially to the inclusion of start and stop bits but also because the transmission literally starts for each character and stops between each character. In the past, data transferred asynchronously was primarily ASCII text, however asynchronous communication is now used for all types of binary data. The link between your computer and dial-up modem is most likely asynchronous. Links between dial-up modems are asynchronous only for slow speeds up to 1200bps. Faster connections use synchronous communication.



GROUP TASK Discussion

As the speed of data transfer increases an asynchronous link experiences more and more errors. How can this be explained? Discuss.

Synchronous communication does not transfer bytes individually, rather it transfers large data blocks known as frames. Frames vary in size depending upon the individual implementation. 10baseT Ethernet networks use a frame size of up to 1500 bytes and frame sizes in excess of 4000 bytes are used on many high speed dedicated links.

There are two elements commonly used to assist the synchronising process. A preamble can be included at the start of each frame whose purpose is initial synchronisation of the receive and transmit clocks. The second element is included or embedded within the data and is used to ensure synchronisation is maintained throughout the frames transmission. Let us consider each of these elements.

Firstly each frame commences with a preamble. On 10baseT Ethernet networks the preamble is 8 bytes (64 bits) long and is simply a sequence of alternating 1s and 0s that end with a terminating pattern (commonly 1 1) called a frame delimiter. The receiver uses the preamble to adjust its clock to the correct phase as the transmitting clock (see Fig 8.12). A frame delimiter is needed at the end of the preamble because

the receiver may lose some bits during clock adjustment so these delimiting bits act as a flag indicating the start of the actual data.

The preamble is followed by the signal that needs to be received. The representation of the bits within the signal provides the second element used to maintain synchronisation. Commonly bits are represented not as high or low signals but using the transitions between these states. An example of such a system is Manchester Encoding used within 10baseT Ethernet networks. Using this system a low to high transition represents a 1 and a high to low transition represents a 0. As the clocks are initially synchronised then the location of the transitions representing the bits is known. The receiver detects each transition, if they are slightly out of synch then the receiving clock adjusts accordingly, hence Manchester Encoding is an example of a self-clocking code. As can be seen in *Fig 8.13*, two frequencies are needed to implement such a system; a base frequency and a frequency that is precisely double the base frequency. Data is transmitted at the same rate as the base frequency. For example, 10baseT Ethernet transfers data at 10 megabits per second and therefore a base frequency of 10 mega hertz is used.

Synchronous transmission systems, such as Manchester Encoding are technically much more difficult to implement than simple asynchronous systems. They were once used solely for high speed, high quality links between large mainframe and mini-computers. Today synchronous transmission systems have largely replaced most asynchronous links.



Consider the following:

Asynchronous communication requires an additional start bit and at least one stop bit for every 8 bits of data. These extra bits add at least 25% to the amount of data actually transferred. Furthermore, the data must be split into individual bytes. In contrast synchronous transmission requires a negligible overhead and each frame contains thousands of bytes.



GROUP TASK Discussion

The discussion above implies that synchronous communication is clearly superior to asynchronous communication. In fact this is not always the case. Identify and describe applications where asynchronous communication is superior.

MEASUREMENTS OF SPEED (bps, Baud and Bandwidth)

Bits per second (bps), baud rate and bandwidth are all measures commonly used to describe the speed of communication. Unfortunately many references use these terms incorrectly. The most common error is to use all three terms interchangeably to mean bits per second. In this section we consider the correct meaning of each of these measures, together with their relationship to each other.

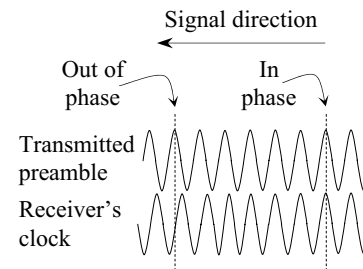


Fig 8.12

The preamble is used to synchronise the phase of the receiver's clock to match the transmitter's clock.

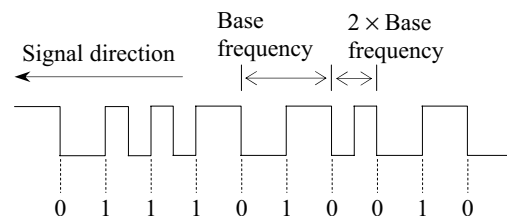


Fig 8.13

Manchester encoding uses the transitions between high and low to represent bits.

Bits per second is the rate at which binary digital data is transferred. For instance a speed of 2400bps, means 2400 binary digits can be transferred each second. Notice bps means bits per second not bytes per second. If a measure refers to bytes a capital B should be used, and if it refers to bits then a lower case b should be used; for example kB means kilobyte and kb means kilobit, similarly MB means megabyte whilst Mb means megabit. It is customary to refer to bits when describing transmission speeds.



Bits per second (bps)

The number of bits transferred each second. The speed of binary data transmission.

Baud rate is a measure of the number of distinct signal events occurring each second along a communication channel. A signal event being a change in the transmission signal used to represent the data. Technically each of these signal events is called a baud, however often the term baud is used as a shortened form of the term baud rate.

In our discussion on asynchronous and synchronous transmission we assumed each signal event or baud represented a single bit, this need not be the case. For example, a connection could represent 2 bits within each baud by transmitting say +12 volts to represent the bits 11, +6 volts for 10, -6 volts for 01, and -12 volts for 00. If this connection were operating at 1200 baud then 2400bps could be transmitted. This example is trivial, in reality various complex systems are used where up to 4, 6, 8 or more bits are represented by each baud. In these situations different waveforms or symbols are needed to represent each bit pattern. The number of different symbols required doubles for each extra bit represented, for example to represent 4 bits requires $2^4 = 16$ different symbols whilst 5 bits requires $2 \times 16 = 32$ different symbols. Altering or modulating the amplitude, frequency and/or phase of the signal produces these different symbols; Fig 8.14 shows these modulation techniques separately. As most high-speed data communication is restricted to a particular range of frequencies, most encoding systems use a combination of amplitude and phase modulation.



Baud (or baud rate)

The number of signal events occurring each second along a communication channel. Equivalent to the number of symbols per second.

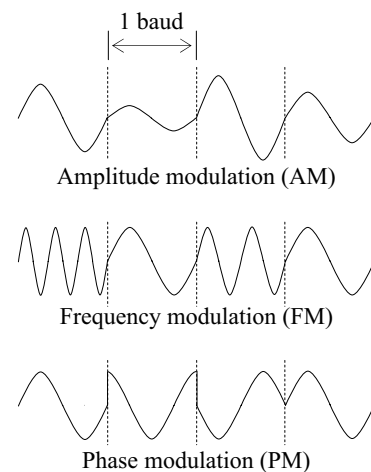


Fig 8.14
Examples of amplitude, frequency and phase modulation.



Consider the following:

Modems transferring data at 28.8kbps usually communicate at 3200 baud where each signal event represents 9 bits of data, that is, a rate of 9 bits/baud. However 28.8kbps can also be achieved by using 2400 baud or even 1200 baud.



GROUP TASK Activity

Make up a table with the column headings: baud rate, bits/baud and number of different waveforms. Complete the table so that all rows generate a data transfer rate of 28.8kbps.

**GROUP TASK Activity**

Design a series of modulated waveforms that could be used when transferring 1200bps at 300baud. Use the minimum number of different amplitude and phase alterations.

**GROUP TASK Discussion**

The term 'baud rate' is seldom used anymore, as bps is a more meaningful measure of speed. Do you agree? Discuss.

The term **bandwidth** is often used incorrectly, people make statements such as “video requires much more bandwidth than text” or “my bandwidth decreases as more people use the Internet”. Statements such as these are incorrect, they are using bandwidth when they really mean speed or bps. Bandwidth is not a measure of speed at all, rather it is the range of frequencies used by a transmission channel. Presumably misunderstandings have occurred because the theoretical maximum speed does increase as the bandwidth of a channel increases. However, it is simply impossible for the bandwidth of most channels to change during transmission. Each channel is assigned a particular range of frequencies when it is first setup, unless you run a high-speed Internet company or are creating your own hardware transmitters and receivers, then altering bandwidth is really beyond your control.

**Bandwidth**

The difference between the highest and lowest frequencies in a transmission channel. Hence bandwidth is expressed in hertz (Hz), usually kilohertz (kHz) or megahertz (MHz).

So what is bandwidth? It is the difference between the highest and the lowest frequencies used by a transmission channel. Frequency is measured in hertz (Hz), meaning cycles per second. Each cycle being a complete wavelength of an electromagnetic wave, so 20Hz means 20 complete wavelengths occur every second. As frequency is expressed in hertz then so is bandwidth. For example, standard telephone equipment used for voice operates within a frequency range from about 200Hz to 3400Hz, so the available bandwidth is approximately 3200Hz. As high-speed connections routinely use bandwidths larger than 1,000Hz or even 1,000,000Hz, bandwidth is usually expressed using kilohertz (kHz) or megahertz (MHz). For example 3200Hz would be expressed as 3.2kHz.

**GROUP TASK Discussion**

In Chapter 2 (p60), we discussed the audio media type. During this discussion we stated that the human ear can discern frequencies in the range 20 to 20,000Hz, yet telephones use a range of 200 to 3400Hz. What are the consequences of these differences? Discuss.

All signals need to be modulated in such a way that they remain within their allocated bandwidth. This places restrictions on the degree of frequency modulation that can be used. As a consequence most modulation systems rely on amplitude and phase modulation. For example, most current connections to the Internet use Quadrature Amplitude Modulation (QAM), this system represents different bit patterns by altering only the amplitude and phase of the wave. 16QAM uses 16 symbols to represent 4 bits/ baud, 64QAM uses 64 symbols to represent 6 bits/ baud and 256QAM uses 256 symbols representing 8 bits/ baud.

Amplitude, phase and frequency are related, altering one has an effect on each of the others. Increasing the available frequency range (bandwidth) results in a corresponding increase in the total number of unique amplitude and phase change combinations (symbols) that can accurately be represented and detected. In general, it is true that the speed of data transfer increases as the bandwidth is increased.

It is difficult to discuss bandwidth without mentioning the related term 'broadband'. Broadband, is a shortened form of the words broad and bandwidth. As is the case with numerous computer related terms there are various accepted meanings. In common usage broadband simply refers to a communication channel with a large bandwidth. However, the term is also used in reference to a physical transmission medium that carries more than one channel. In essence, the total bandwidth is split into separate channels that each use a distinct range of frequencies. Using either meaning, ADSL (Asymmetrical Digital Subscriber Line), cable and 3G HSPA (High Speeds Packet Access) mobile networks are all examples of broadband technologies. They deliver high data rates (theoretically in excess of 5Mbps) by splitting the total bandwidth into channels. The opposite of broadband is narrowband. Narrowband connections include 56kbps dial-up modem links and 128kbps ISDN links.



Consider the following:

Cable modems are used to connect individual homes to their local cable Internet service provider. The following specifications relate to the Motorola Surfboard SB4200 cable modem:

DOWNSTREAM		UPSTREAM	
Modulation	64 or 256 QAM	Modulation	16 QAM or QPSK
Maximum Data Rate	64QAM 27 Mbps	Maximum Data Rate	10 Mbps
Maximum Data Rate	256QAM 38 Mbps	Bandwidth	200 kHz, 400 kHz, 800 kHz, 1.6 MHz, 3.2 MHz
Bandwidth	6 MHz	Symbol Rates	160, 320, 640, 1280 and 2560 ksym/s
Symbol Rate	64 QAM 5.069 Msym/s	Operating Level Range	+8 to +55 dBmV (16QAM) +8 to +58 dBmV (QPSK)
Symbol Rate	256 QAM 5.361 Msym/s	Output Impedance	75 . (nominal)
Operating Level Range	-15 to +15 dBmV	Frequency Range	5 to 42 MHz (edge to edge)
Input Impedance	75 . (nominal)		
Frequency Range	88 to 860 MHz		

(Source: www.motorola.com/broadband)



GROUP TASK Discussion

Each cable TV channel occupies a bandwidth of 6MHz. How many TV stations can be accommodated within the downstream frequency range, and why do you think the downstream bandwidth of the modem is equal to the bandwidth of a TV channel? Discuss.



GROUP TASK Research

The above specifications indicate that the symbol rate (Msym/s) is always slightly less than the bandwidth. For example, just over 5Msym/s on a 6MHz channel downstream. Examine the specifications for a variety of modems. Is the symbol rate always just less than the bandwidth?

SET 8A

1. Many bits are transmitted simultaneously over which type of transmission link?
 - (A) asynchronous
 - (B) synchronous
 - (C) serial
 - (D) parallel
2. 'Data skew' can only occur over which type of transmission link?
 - (A) asynchronous
 - (B) synchronous
 - (C) serial
 - (D) parallel
3. A communication link that can be used to either send or receive is best described as a:
 - (A) simplex link.
 - (B) half duplex link.
 - (C) full duplex link
 - (D) duplex link.
4. Which of the following best defines the term bandwidth?
 - (A) The speed of data transfer.
 - (B) The number of signal changes per second.
 - (C) The difference between the highest and lowest frequency used by a channel.
 - (D) The technique used to modulate a digital signal in preparation for transmission
5. Start-stop communication is:
 - (A) the same as asynchronous communication.
 - (B) the same as synchronous communication.
 - (C) an example of asynchronous communication.
 - (D) an example of synchronous communication.
6. Full duplex transmission requires:
 - (A) two communication channels.
 - (B) a single bi-directional channel.
 - (C) simultaneous bi-directional transfer.
 - (D) a telephone line.
7. Encoding takes place during:
 - (A) the transmitting information process.
 - (B) the receiving information process.
 - (C) actual transmission.
 - (D) digital to analog conversion.
8. A polite conversation could best be described as:
 - (A) serial and full-duplex.
 - (B) serial and half duplex.
 - (C) parallel and full-duplex.
 - (D) parallel and half duplex.
9. What is the purpose of a preamble prior to transmission of an Ethernet frame?
 - (A) To alert the receiver that data is about to be transmitted.
 - (B) To synchronise the phase of the receiver's clock with the transmitters.
 - (C) To ensure both receiver and transmitter's clocks remain in synch.
 - (D) To indicate the destination address for the frame of data that follows.
10. Transitions between high and low, and low and high are used to represent bits because:
 - (A) transitions make up the majority of the signal.
 - (B) transitions are easy to detect compared to detecting high or low voltages.
 - (C) transitions allow the receiver to adjust its clock during the transmission.
 - (D) such systems do not require the use of start and stop bits.
11. Compare and contrast the terms:
 - (a) serial and parallel
 - (b) simplex, half duplex and full duplex
 - (c) synchronous and asynchronous
 - (d) uploading and downloading
12. Is Baud rate and symbols per second always equal? Justify your response.
13. Discuss reasons why parallel communication is rarely used apart from within a single computer.
14. Describe how synchronous communication is achieved:
 - (a) between components on the motherboard.
 - (b) during transmission of Ethernet frames.
15. Define the terms bps, Baud rate and bandwidth. Explain their relationship to each other.

HARDWARE IN TRANSMITTING AND RECEIVING

Transmitting and receiving information processes occur within a computer, between the computer and its peripheral devices and also between computers using modems and networks. In this section we examine the different hardware tools used to transfer data in each of these areas. Remember, at this stage, we are interested in the hardware components. Hardware means physical devices, however in regard to transmitting and receiving we also include electro-magnetic waves (used to represent data) as physical components of the system. Hence a wire is a physical component and so to is the signal moving down the wire.

Before we commence, let us consider a general overview of the communication hardware present in a typical system. To simplify our overview, we use an example home network that includes a connection to the Internet (see *Fig 8.15*). This example could easily be expanded to represent much larger system configurations.

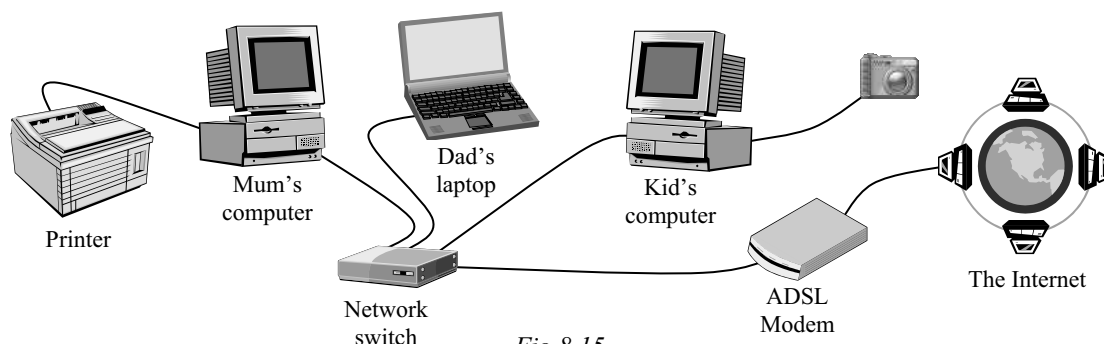


Fig 8.15

A typical home network connected to the Internet.

Firstly transmitting and receiving processes are occurring between components within each computer. For example Mum's computer is transferring data from RAM to the CPU whilst she surfs the net. Therefore we examine the various busses on the motherboard. Secondly, Mum's computer sends data to the printer and the Kid's computer communicates with the digital camera. Communication with external devices occurs via expansion slots and ports, hence we examine examples of these components. Thirdly modems are used to connect to other remote computers (usually over the Internet). In *Fig 8.15* the modem connects the LAN (Local Area Network) to the Internet. Modems come in various types; for example cable, satellite, ADSL and dial-up modems. We examine the processes modems perform, in particular the modulation and demodulation of signals. Finally we discuss network hardware for local area networks (LANs) and wide area networks (WANs). Our discussions in regard to networks centre on the hardware tools required rather than their detailed operation.

In summary, we examine hardware for transmitting and receiving in regard to:

- Communication within a computer (the system or internal bus).
- Communication with external devices (external buses)
- The role of modems.
- Local area networks (LANs).
- Wide area networks (WANs).



GROUP TASK Activity

Many of the hardware tools for communication have already been mentioned throughout the text. Identify and list the hardware tools used for each of the above dot points.

COMMUNICATION WITHIN A COMPUTER (THE SYSTEM OR INTERNAL BUS)

Within a computer all the various devices and components communicate using an intricate network of wires called ‘the bus’. The motherboard is a circuit board containing the bus together with various chips; if you look closely at a motherboard you can see some of the bus lines (see *Fig 8.16*). There are multiple layers sandwiched together on a modern motherboard, hence only the top and bottom layers are visible. The bus lines are copper traces, essentially wires for moving data between different components.



Fig 8.16
Bus lines on a
motherboard.

The computer’s bus is made up of the system bus (or internal bus) and a collection of external busses. In this section we examine the characteristics and operation of the system bus. The system bus is used to transfer data between the CPU and main memory, and also between the CPU and the input/output (I/O) systems. Therefore the system bus corresponds to von Neumann’s “stored program concept”; that is, each of the arrows shown on *Fig 7.4* (p250) corresponds to parts of the system bus.

The system bus is made up of the data bus, the address bus and the control bus; data is sent down these wires in parallel. All wires within the system bus originate at the CPU and are connected to both main memory and also to each I/O system. Each wire is used to transmit electrical signals representing the bits being sent. Usually a voltage above 2 volts represents a binary 1 and below 0.8 volts represents a binary 0. Most motherboards operate using voltages ranging from 2.8 to 5 volts. For example, a motherboard with a core supply voltage of 2.8 volts would represent a binary 1 using 2.8 volts and a binary 0 using a low voltage close to zero. Two layers on the motherboard are used to supply power to each component; one supplies the core voltage and the other is a ground connection used to complete the circuit.

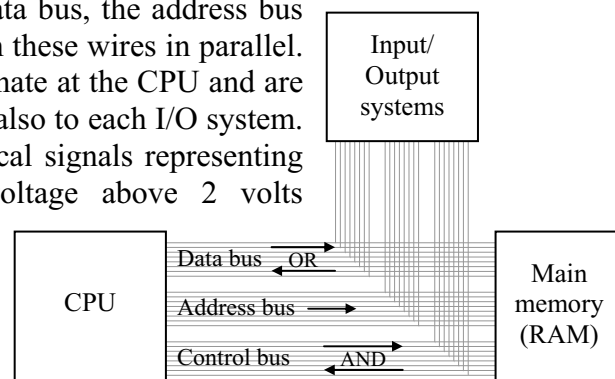


Fig 8.17
The system (or internal) bus contains the data,
address and control bus.

The data bus on most current motherboards contains 64 parallel connections, thus it is able to transfer 64 bits simultaneously. It is the size of the data bus that is used when the bus capacity of a computer is quoted. We discussed the relationship between bus capacity and processing speed in Chapter 7 (p253). The data bus is used to transfer data from the CPU or to the CPU. Notice we said ‘or’, hence the data bus operates in half-duplex. Each wire within the data bus connects the CPU to both main memory and the I/O systems. Therefore, all data placed on the data bus is available to all components. The intended receiver is determined by the data that is present on the address and control buses.

The address bus is used to transmit memory locations from the CPU to both memory and the I/O systems. The data always travels from the CPU, hence the address bus is an example of simplex transmission. The width of the address bus determines the theoretical maximum number of memory locations possible, where each memory location holds 1 byte (8 bits) of data. Current Intel CPU based systems contain a 36-bit or 40-bit address bus. Theoretically a 36-bit address bus can address up to 2^{36} bytes or 64GB of memory (primarily RAM) and a 40-bit address bus up to 1024GB.

When the CPU places an address on the address bus all memory modules and I/O systems receive the address. If the address corresponds to one of their own addresses then the action detailed on the control bus is performed. For example, if the address is a main memory address and the control bus specifies a read operation then the data at that address in memory is retrieved and placed on the data bus. Similarly if the control bus specifies a write operation then the data on the data bus is stored in the memory location specified on the address bus.

Most CPU designs use a single address bus for both memory operations and I/O operations. These CPUs utilise two different address spaces, one for memory and another for I/O. On Intel CPU systems just 16 of the 36 (or 40) address bus wires are used for I/O operations. Hence just 65,536 (2^{16}) unique I/O addresses are possible ranging from 0 to 65,535. Each I/O system is allocated one or more of these addresses. For example in *Fig 8.18* the virtual serial port COM1, is using addresses 1016 to 1023 (03F8 to 03FF in hexadecimal or base 16). If data is to be transferred to or from COM1 then the CPU places an address in the range 1016 to 1023 on the address bus.



Consider the following:

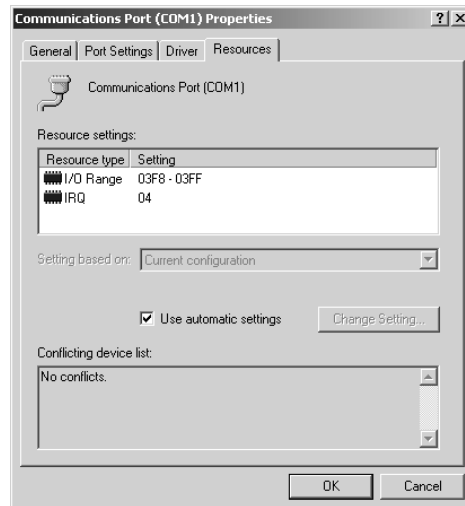


Fig 8.18
I/O addresses and IRQ line used by COM1 to communicate with the CPU.

Systems with a 36-bit address bus can have up to 68,719,476,736 (2^{36}) different memory locations that each corresponds to 1 byte of actual storage. Therefore memory addresses range from 0 to 68,719,476,735. However I/O addresses range from 0 to 65,535. This means addresses from 0 to 65,535 are present in both the memory and I/O spaces. As a single address bus is used, how can the difference between memory and I/O addresses be determined? Using a signal placed on the chip enable wire within the control bus. If a 0 (low voltage) is present on the chip enable wire then the address is a memory address. Conversely if a 1 (higher voltage) is present then the address bus holds an I/O address.

What about devices that contain large amounts of their own memory? For example video cards commonly contain more than 32MB of VRAM, more than 32 million addresses are needed yet only 65,535 I/O addresses are available. To overcome such problems I/O systems pretend that they are part of main memory; in essence they use some of the memory address space.



GROUP TASK Discussion

Explain how the data on the chip enable wire and the address bus can be used to allow the CPU access to large amounts of memory present on video cards (and other I/O systems).



GROUP TASK Activity

Examine the I/O addresses and memory addresses used by the various I/O systems on your home or school computer. Confirm that these addresses are within the ranges mentioned above.

The control bus is a collection of parallel wires used by the CPU to control the operation of main memory and the various I/O systems. It is the design of the CPU that determines the individual control bus wires present on a particular motherboard. In general, each wire within the control bus connects to the control unit within the CPU. Let us consider the purpose and operation of some of these wires.

The system clock is located on the motherboard. The signal generated is transmitted along a wire within the control bus and hence is available to all devices connected to the system bus. The clock signal continuously alternates between high (1) and low (0) at a constant pace. The transitions from high to low and in some instances also from low to high are used to synchronise the transfer of data and also the operation of all components connected to the system bus. Communication along the system bus is therefore an example of synchronous transmission.

A high to low transition is known as a falling edge whilst a low to high transition is known as a rising edge. For example, *Fig 8.19* describes a system clock operating at 100MHz, this means a falling edge occurs every 10ns ($1/100,000,000$ sec) and some sort of transition occurs every 5ns. A nanosecond (ns) is 1 billionth of a second, consequently processes are occurring at an amazing speed. In fact, it is the processes within each device that take most of the time; the actual transfer of data along the system bus takes virtually no time at all! Once the voltage is altered in a copper wire it travels at close to the speed of light. Therefore, transitions present in the clock wire occur at virtually the same instant at all points along the wire.

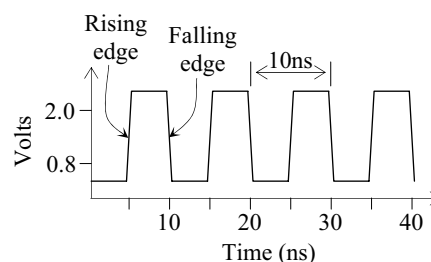


Fig 8.19

Signals from the system clock are transmitted along one wire within the control bus.



Consider the following:

How can a change of voltage take virtually no time to arrive at the other end of a wire? Imagine the wire is a garden hose. If the hose is empty then it will take a few seconds from when the tap is turned on until water starts to flow out of the hose. However, if the hose is initially full of water then water will begin flowing out the end of the hose virtually as soon as the tap is turned on.

Now imagine the water in the hose are electrons in a wire. Essentially, all wires are full of electrons, just like the hose when full of water. Also imagine the pressure released from the tap is the voltage placed on the wire. As soon as the voltage is placed on the wire it is almost instantly present at all places on the wire in the same way that the water pressure is present within the hose.

Our hose analogy is not quite accurate, water pressure waves travel at around the speed of sound (approximately 330m/s). Electromagnetic waves are much faster, within a vacuum they travel at the speed of light (3×10^8 m/s), through copper wire this speed is around 2×10^8 m/s or two thirds the speed of light.



GROUP TASK Discussion

Imagine the hose is turned on for two seconds and off for two seconds repeatedly. Identify aspects of this scenario that are similar to the signal generated by the system clock and described in *Fig 8.19*.

During our discussion we alluded to a number of different wires within the control bus. We stated that the data bus is used to both send and receive data, hence the control bus contains a read/write wire. A low voltage (0) means write and a high voltage (1) means read. We also mentioned the chip enable wire, this wire determines if the address refers to memory (0) or to the I/O system (1). For example, to receive data from a modem connected to say COM1, the CPU must place a 1 on both the read/write and chip enable lines together with the appropriate COM1 I/O address on the address bus. When COM1 responds it places the requested data on the data bus and sets the acknowledge wire to high. The acknowledge wire, which is also part of the control bus, is returned to low once the CPU has the data.

There are many more elements of the control bus that perform various different tasks. At this stage we restrict our discussion to a brief mention of just two of these elements, interrupt request lines (IRQs) and direct memory access lines (DMAs).

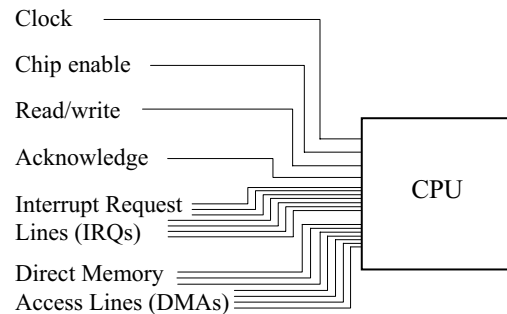


*Fig 8.20
Some of the wires present within
a typical control bus.*

An IRQ is a direct line from a device to the CPU. Intel CPU based systems contain 16 IRQ wires labelled IRQ0 through to IRQ15. Devices use their IRQ line to get the attention of the CPU. For example, in Fig 8.18 COM1 port is assigned IRQ4. If COM1 has data for the CPU then it sets the IRQ4 wire to 1. The CPU recognises this, stops what it is doing and commences communicating with COM1. Once finished the CPU returns to its previous task. In essence each interrupt request wire can be thought of as a hotline from a device to the CPU.

Direct memory access (DMA) allows devices to transfer data to and from main memory without the assistance of the CPU. Therefore DMAs are used between most secondary storage devices and also video cards, however they can be used for various other purposes. A DMA controller is installed between the I/O systems and the system bus. When a device wishes to transfer data directly with main memory the DMA controller informs the CPU using one of the DMA channels on the control bus. When the CPU is ready it relinquishes control of the system bus to the DMA controller. The DMA controller then takes over to facilitate the transfer of data directly to or from main memory. The aim of DMA systems is to free the CPU from any involvement in simple data transfer operations.



Consider the following:

Consider a typical store operation: as the falling edge of the clock is detected, the CPU simultaneously places an address on the address bus, data on the data bus, and a 0 on both the read/write and chip enable wires of the control bus. Main memory recognises the address and stores the data into that memory location. Main memory then places a 1 on the acknowledge wire. The CPU detects this and returns the acknowledge wire to 0.



GROUP TASK Discussion

Develop a description of a typical read operation similar to the above description of a store operation.

COMMUNICATION WITH EXTERNAL DEVICES (EXTERNAL BUSSES)

External busses are used to transfer data between the system bus and other hardware devices. Essentially they are the other side of the I/O systems present on the system bus. Hardware devices connected to an external bus include hard disks, floppy disks, display adaptors, network cards, modems, printers, scanners, digital still and video cameras, etc. In fact, all hardware except the CPU and main memory is connected to an external bus. Some devices connect via an expansion slot on the motherboard and others connect to ports either on or connected to the motherboard. Many modern motherboards include complete peripheral devices embedded on the circuit board: these embedded peripherals still connect to the system bus via an external bus.

So what is an external bus and what is its purpose? The purpose of an external bus is to transfer information between the system bus and some hardware device. To do this requires the signals on the system bus to be converted into signals that can be transferred and understood by the attached hardware device, and vice versa. Therefore, external buses provide an interface between hardware devices and the system bus. All such interfaces have two significant components, a controller that performs the encoding and decoding together with the physical communication lines. A similar selection of interfaces is present on most computers, consequently their controllers have been combined within a single chip. This chip is often referred to as a 'chipset', which indicates its functionality was once performed by a set of different chips.

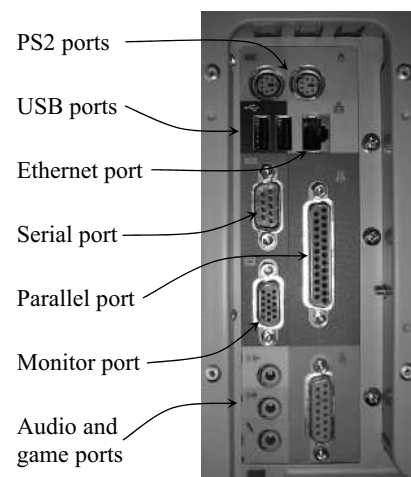


*Fig 8.21
Ports on the rear of a typical
personal computer.*



*Fig 8.22
Major components on a typical motherboard.*

There are many different types of external bus and each uses a different set of rules and communication lines to transfer data. *Fig 8.21* and *Fig 8.22* show the expansion slots, connectors and ports on a typical personal computer. In this section we restrict our discussion to two common examples: the PCI bus and USB ports.

PCI Bus

The Peripheral Component Interconnect (PCI) bus was first developed in 1992. More recently (from 2004) a series of PCI Express standards have been introduced. PCI Express or simply PCIe operates differently from conventional PCI. From 2007 PCI Express has been widely used to connect high performance video cards to most consumer computers. PCI Express uses multiple serial connections whilst conventional PCI is a parallel communication standard. Currently (2009) most personal computers include both conventional PCI and PCI Express slots. In this section we consider conventional PCI buses.



Fig 8.23
Belkin's 32-bit PCI wireless network interface card.



GROUP TASK Research

Determine the number and type of PCI and PCIe slots present on your home or school desktop computer. Are any of these slots being used?

The original PCI specification operated at clock speeds up to 33MHz using a 32-bit data bus, data transfer rates up to 133Mbps were possible. The PCI standard, administered by the PCI Special Interest Group, has been revised regularly over the years. Currently the conventional PCI standard provides for clock speeds up to 533MHz using bus widths of either 32 bits or 64 bits. Speeds up to 4.3Gbps are possible, some 32 times faster than the original PCI standard. As the principles underlying PCI communication have not changed significantly older PCI expansion cards remain compatible with conventional PCI slots on newer motherboards. Fig 8.23 shows a PCI wireless network expansion card and Fig 8.22 shows a motherboard containing conventional PCI expansion slots.

Conventional PCI is a parallel synchronous standard. This is clearly evident within the above description; when clock speeds are mentioned it is reasonable to assume transfer is synchronised to the clock, and when bus widths are mentioned then parallel transmission is being used.

The original PCI standard was the first to introduce PnP (Plug and Play). PnP requires that all PCI expansion cards contain permanent registers specifying a unique identifier together with details of the card's system requirements. This data allows newly installed expansion cards to be automatically allocated system resources, such as IRQs, DMAs and memory addresses, without user intervention. PCI is not a CPU specific standard, in fact most modern motherboards for most types of CPU include PCI expansion slots. Unfortunately the same cannot be said for many PCI expansion cards. In reality most are compatible with particular CPU families. For example, it is unusual for a PCI expansion card to operate seamlessly with both Intel Pentium and Apple Macintosh systems.

A PCI interface is physically comprised of PCI expansion slots, the actual bus lines and a PCI bridge (see Fig 8.24). The PCI bridge encodes and decodes messages between the system bus and the PCI bus. Its main task is to reorganise messages received from devices on one bus into a form that is

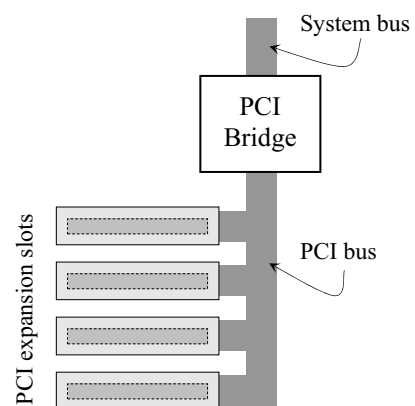


Fig 8.24
The PCI bridge encodes and decodes messages to and from the system and PCI bus.

understood by devices on the other bus. In addition, the bridge allows multiple PCI devices to share some of the same system resources. For example, the PCI bridge allows a PCI modem and a PCI network card to share a single IRQ line to the CPU. On modern motherboards the circuitry within the PCI bridge is integrated within the motherboard's chipset. All PCI expansion slots share the same bus lines and only one PCI device is able to transfer data at any particular time. A single complete transaction often involves addresses and data being transferred in different directions along the same $AD(x)$ wires within the PCI bus.

Signal	Number of pins	Driven by	Description
CLK	1	Initiator	Clock
AD(x)	32(64)	Sender	Address and data lines
C/BE(x)	4(8)	Initiator	Command, byte, enable
FRAME	1	Initiator	Address or data phase
DEVSEL	1	Target	Device select
IRDY	1	Initiator	Initiator ready
REQ	1	Initiator	Request transfer
GNT	1	Chipset	Grant transfer
PAR	1(2)	Sender	Parity bit for AD and C/BE
PERR	1	Receiver	Parity error
RST	1	Initiator	Reset
SERR	1	Any	System error
STOP	1	Target	Stop transfer from target
TRDY	1	Target	Target ready
+5V	8	-	Power
+3.3V	12	-	Power

Fig 8.25
Significant signals on a PCI interface.

In the PCI standard, data is transferred between an initiator and a target, the transfer of actual data can be in either direction. Any device can be an initiator, it could be a PCI device or it could be the CPU or the memory controller. The initiator negotiates control of the bus, including the system bus, by activating the *REQ* wire. The PCI circuitry within the chipset negotiates with all other system devices for control of the bus. Clearly some devices, in particular the CPU, have higher priority during this process. Once the PCI circuitry is able to grant bus control it activates the *GNT* wire. The initiator detects this signal and now has control of the bus.

The command to be executed is specified using the command (*C/BE(x)*) wires. When a 32-bit device is being used there are 4 *C/BE* wires, hence 2^4 or 16 different combinations are possible. Operations involving data transfer include: I/O Read (0010), I/O Write (0011), Memory Read (0110) and Memory Write (0111). Remember the initiator can be the memory controller, the CPU or any PCI device; this does not alter the direction in which the data is transferred. For example, if a PCI modem is the initiator and executes the command 0011 addressed to the CPU, then the CPU will transmit data to the modem. The same occurs if the CPU is the initiator.

Execution of each command commences with an address phase followed by one or more data phases using the following procedure:

- During the address phase the initiator places the command on the *C/BE* wires and the target address on the $AD(x)$ wires. Note that the address is sent using the same $AD(x)$ wires as the data that follows.
- The target acknowledges receipt and acceptance of the command by placing and maintaining a low (0) on the *DEVSEL* wire.
- The sender (which may be the initiator or the target depending on the command) places data on the $AD(x)$ lines and activates its *RDY* line (either *IRDY* or *TRDY* depending on who is the sender).
- The data is clocked into the receiver using the rising edge of the *CLK* signal. Hence data continues to be transferred synchronously.

- Either the initiator or the target can pause the transfer at any time by placing a high (1) on the *IRDY* or *TRDY* signals respectively. In essence both *IRDY* and *TRDY* must be low (0) for data to be transferred.
- As the last data phase is commenced the initiator sets the *FRAME* signal to high (1) indicating the last data phase.
- Finally the transfer concludes and the *IRDY*, *TRDY* and *DEVSEL* wires are returned to high (1).

A complete address phase together with its multiple data phases is known as a frame, whilst a frame is being transferred the *FRAME* signal is held low (0) by the initiator.



Consider the following:

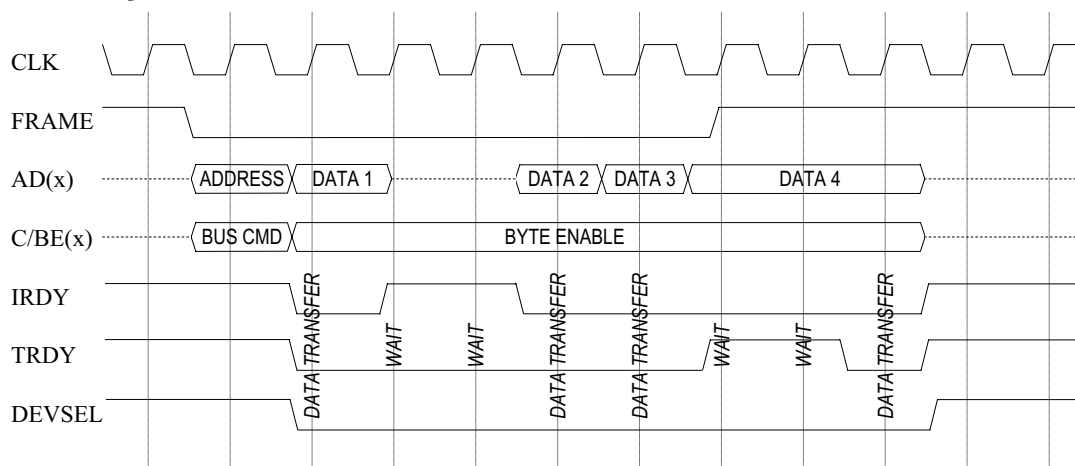


Fig 8.26

PCI write cycle containing 4 data phases.



GROUP TASK Discussion

Describe the sequence of signalling events modelled in the above signal diagram. In particular, identify and describe reasons why data is not transferred on each rising edge of the clock signal.



GROUP TASK Discussion

PCI certainly operates synchronously when transferring data. However, is the transfer of control data really synchronous? Discuss.

USB Ports

Universal Serial Bus (USB) ports are used to connect a wide variety of peripheral devices. Examples of USB devices include, mice, keyboards, scanners, network adapters, printers, telephones, digital cameras and audio systems. Up to 127 individual peripheral devices can be connected to a single universal serial bus (USB).

Clearly computers are not sold with 127 USB ports, consequently the USB standard allows for expansion using USB hubs. A USB hub is like a double adapter, or a power board, it allows a single USB port to be split into multiple ports. For example, the USB hub in Fig 8.27 expands a single USB port into 4 USB ports. Furthermore, even more ports can be added by simply chaining hubs together. It is possible for an individual



Fig 8.27

4-port USB hub.

device to be chained to a computer using a sequence of 4 USB hubs. Computers containing multiple USB ports actually contain an embedded USB hub, called the root hub. All the USB ports connected to a single USB hub transfer data on the same universal serial bus (USB).

Physically, the USB 2.0 standard defines four types of cable connector, an A-type and Mini-A connector for connections heading towards the computer and a B-type and Mini-B connector for connections heading away from the computer. *Fig 8.28* shows the A and B type connectors; the mini connectors are commonly used on small devices such as digital cameras, mobile phones and PDAs. Many peripherals have their cable attached to the device and thus only require an A-type connector. Using A and B-type connectors makes it physically impossible for a user to incorrectly connect a device. This feature assists in achieving one of USB's primary aims; to simplify the installation of peripheral devices for users.



Fig 8.28
USB A and B type connectors.



GROUP TASK Discussion

USB uses A-type and B-type connectors. In reality some devices use their own connectors in place of the B-type connector. Why is this OK in regard to B-type connectors but not for A-type connectors? Discuss.

USB cables contain just 4 wires; two are used for power (+5 volts and ground) and a single twisted pair of data wires (D+ and D-). The power wires are able to supply power to attached devices. However, the available current is only suitable for low power devices such as mice, keyboards, mobile phones and digital still cameras. High power devices, such as scanners and printers still require their own dedicated power supply.

The two data wires combine to form a single communication channel. With just one channel how is it possible for such a varied collection of devices to communicate? Our aim is to answer this question. Furthermore, the current USB 2.0 standard permits the transfer of data at some 480Mbps and the USB 3.0 standard is designed to achieve speeds up to 4Gbps!

Let us base our discussion on a typical setup. The USB ports connect to a USB mouse, USB flatbed scanner and a USB digital video camera (refer to *Fig 8.29*). On the computer's motherboard is the circuitry for controlling the bus; these circuits contain the host controller and an integrated USB root hub. Lines from the root hub connect to each USB port, which in turn are used to connect each of our three peripherals.

Firstly the host controller must detect each attached device. This process is known as enumeration which simply means 'assigning each device a number'. However, USB enumeration is a little more complex; it first assigns each device a unique identification number, queries the device to determine its requirements and assigns it resources based on these requirements. The requirements primarily provide information in regard to the device driver to use and also the type of

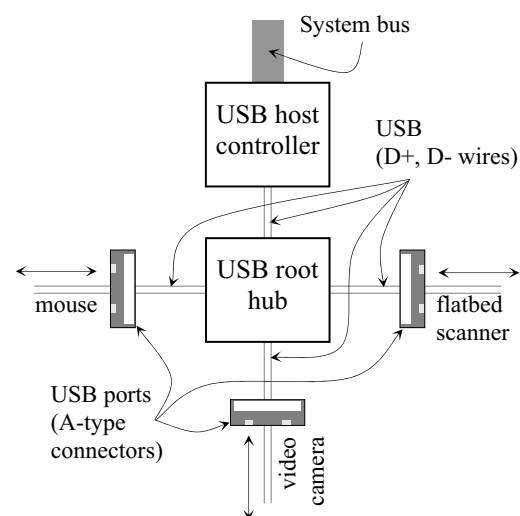


Fig 8.29
Components of a USB interface.

USB connection the device needs. If the appropriate device driver is already installed within the system then the host controller instructs the operating system to load and commence executing the driver. If no device driver is found then the user is prompted to locate and install the driver.

The USB standard defines four types of USB connection, namely control, interrupt, bulk and isochronous. Control connections are used to configure devices, both during enumeration and also during normal operation. All USB devices must be able to receive and also send control messages. However, each USB device uses just one type of connection when actually transferring data; either interrupt, bulk or isochronous. The type of connection used is stored permanently within the device. The host controller accesses this information using a control connection and subsequently assigns the connection type to the device. Let us discuss the differences between each of these connection types.

Interrupt connections are designed for low speed devices that only need to transfer small quantities of data at random times, yet it is vital that the data be transferred as soon as possible. In our example system in *Fig 8.29*, the mouse would use an interrupt connection. Bulk connections are designed for devices that need to transfer large quantities of data but precisely when the data is transferred is not critical. Devices using a bulk connection are assigned transfer space on the fly by the USB host. In our example, the scanner would use a bulk connection. Isochronous connections are used when the transfer is time critical. That is, the nature of the data means that it must be delivered in a constant stream. Usually isochronous connections are used to stream audio and video to and from devices. When an isochronous connection is required the device informs the USB host in regard to its desired transfer speed. The USB host, where possible, allocates such devices a guaranteed and constant transfer speed. Isochronous transfers do not include any error checking. In our example the digital video camera would use an isochronous connection.

Note that USB devices can be ‘hot swapped’, meaning you can plug and unplug cables even when the host computer is on. Therefore enumeration occurs not just when the computer is booted, but every time a USB device is plugged in or unplugged. Hot swapping and the enumeration process further simplify the installation of USB devices for users. In our example it would be perfectly fine to disconnect the video camera and connect a printer. The USB host would simply detect the change and perform the enumeration process.



GROUP TASK Discussion

Brainstorm a list of USB peripheral devices. Discuss whether each of these devices would use interrupt, bulk or isochronous USB connections.

In terms of our example USB system from *Fig 8.29*, let us assume enumeration has completed successfully. This means we have a mouse assigned to say, device ID 1 using an interrupt USB connection. We have a scanner, say device ID 2, using a bulk USB connection. And we have a video camera, say device ID 3, using an isochronous USB connection. All three devices are connected to the single USB communication channel controlled by a single USB host controller.

Before we commence discussing USB data transfer, it is worth noting that many of the tasks performed by the host controller are actually performed using software. Our current discussion focuses on hardware, however in regard to USB the line between hardware and software is somewhat blurred. Our discussion is based on the original USB 1.1 standard; this standard provides for a maximum USB transfer rate of

12Mbps. The newer USB 2.0 and 3.0 standards operate on similar principles, but offer higher speeds and hence the timing and transfer rates are different.

All data is transferred within frames. A new frame commences every millisecond, so one thousand frames are transferred every second. To achieve a transfer rate of 12Mbps each frame needs to contain 1500 bytes of data ($12,000,000 \text{ bits per second} \div 1000 \text{ frames per second} \div 8 \text{ bits per byte}$). In fact, all USB frames must be of identical size, even when no actual data is being transferred.

Every frame commences with a start of frame packet, this packet is used to synchronise each device to the host controller's clock in a similar manner to the preamble used within Ethernet transmissions (refer p285-6). Data packets follow the start of frame packet. Each data packet is preceded by the device ID of the intended recipient.

Isochronous devices have a predetermined length of data packet that is included within all frames transmitted. If the host is sending data to such a device then the signal representing the data is placed on the bus by the host. If the device is sending to the host then the device places the signals on the bus at the appropriate times. At all times it is the host that creates each frame, including the organisation of packets within each frame. For example, if our video camera needs to transfer at 3Mbps then it requires one quarter of each frame for its data packets, namely 375 bytes. Hence every frame will contain a data packet addressed to device ID 3, that contains 375 bytes. This is true regardless of whether the video camera is actually sending or receiving data. When the video camera is transmitting it detects its device ID within each frame and then commences placing 375 bytes of data on the wire, one bit for each tick of the clock.

Data packets for interrupt devices have data packets created within frames at regular intervals, perhaps every 10th frame. The packets are generally far smaller than those used for isochronous devices. For example in Chapter 3 (p87), we mentioned that a typical mouse sends data 40 times per second, and each transmission includes X and Y distance data, direction data, button data and scroll wheel data; around 4 bytes in total. As 1000 frames occur per second then only every 25th ($1000 \div 40$) frame need contain a 4-byte data packet for use by the mouse.

What about bulk transfers? These devices are allocated whatever bytes remain within each frame. This implies that some frames may well be full, and therefore bulk devices will never get a chance. In fact, the USB standard ensures that only 90% of the available frame size is used for isochronous and interrupt connections, the remaining 10% being reserved for both bulk and control packets. Nevertheless, bulk transfers will certainly suffer as extra devices are added to the bus. Consider our scanner, even a moderately sized image, say 1500 by 2000 contains 3 million pixels. If each pixel is represented using 3 bytes (24-bit colour depth), then the total file is 9MB or 72Mb. If say half of each frame is available for bulk scanner data then the total file will take 12 seconds to transfer; $72 \text{ megabytes} \div 6 \text{ megabits per second}$ or $9,000,000 \text{ bytes} \div 750 \text{ bytes per frame} \div 1000 \text{ frames per second}$. 12 seconds is not an unreasonable amount of time, as it is roughly the time it would take the scanner to physically scan the image. If the USB is at full capacity and no control packets are transferred then only 10% of each frame is available, under these circumstances the image will take 60 seconds to transfer, an unreasonable amount of time.

Finally let us briefly examine how bits are physically represented on a USB. 0's are represented by transitions and 1's by no transition, using a system similar to Manchester encoding (see p286) called NRZI (Non-return to Zero Inverted). NRZI,

like Manchester encoding, is a self-clocking system. Clearly, maintaining the synchronisation of both host and device clocks is essential for efficient USB transfer.



Consider the following:

According to its formal specifications, the Universal Serial Bus was developed with the following purpose:

1. To achieve ubiquitous and cheap connectivity to accommodate the convergence of computing and communications, in particular telephony.
2. To make the connection of peripheral devices less confrontational and easier to configure.
3. To make more low-cost bi-directional ports available to the growing number of peripheral devices, including telephone, fax and modem adaptors, answer machines, personal digital assistants, keyboards and mice.



GROUP TASK Discussion

Has USB achieved its purpose in relation to each of the above points?
Discuss using evidence from our USB discussion.



Consider the following:

You may or may not have noticed that during our discussions in regard to PCI and USB we made no mention of the terms simplex, half-duplex or full-duplex. Both PCI and USB utilise the same communication channels for both addresses and data. Furthermore, these same channels can operate in both directions. In many instances a single message contains data travelling in opposing directions.



GROUP TASK Discussion

How is it possible for a single message to contain data travelling in opposing directions? Discuss.



GROUP TASK Discussion

Are the terms simplex, half-duplex and full-duplex suitable for describing PCI and USB? Perhaps none of these terms are appropriate at all, or perhaps the terms are useful for describing some aspects of the communication. Discuss.

SET 8B

1. The system bus is composed of:
 - (A) the internal and external bus.
 - (B) a data, address and control bus.
 - (C) the CPU, main memory and I/O systems.
 - (D) all the elements of von Neumann's stored program concept.
2. Which collection of terms best describes the operation of the data bus?
 - (A) Asynchronous, parallel, full duplex.
 - (B) Synchronous, parallel, full duplex.
 - (C) Synchronous, parallel, half duplex.
 - (D) Synchronous, serial, half duplex.
3. Which is the best description of the rising edge of a transition?
 - (A) A rising edge occurs as the signal changes from a high voltage to a low voltage.
 - (B) A rising edge occurs as the signal changes from a low voltage to a high voltage.
 - (C) A rising edge occurs when a strobe signal changes state.
 - (D) A rising edge occurs each time the clock signal changes state.
4. Which of the following is used by motherboard components to determine whether they should act upon instructions?
 - (A) The IRQ line that has been activated.
 - (B) The instruction on the control bus.
 - (C) The address on the address bus.
 - (D) The device ID present on the data bus.
5. In terms of USB cables, which of the following is true?
 - (A) A-type connectors plug into ports on the computer.
 - (B) B-type connectors plug into ports on the computer.
 - (C) The connectors on either end of a USB cable are identical.
 - (D) Devices with attached cables only require B-type connectors.
6. What is the primary purpose of all external buses?
 - (A) To transfer data between the system bus and other hardware devices.
 - (B) To enable efficient communication between the CPU and main memory.
 - (C) To provide direct access to main memory for attached devices.
 - (D) To provide ports for connecting external devices.
7. I/O devices request the attention of the CPU by:
 - (A) activating their DMA lines.
 - (B) placing their address on the address bus.
 - (C) activating their IRQ line.
 - (D) setting their own internal registers.
8. What is the purpose of the four wires within a USB interface?
 - (A) 2 wires are used to power external devices and the other 2 form two communication channels.
 - (B) 1 wire provides power to external devices and the other 3 are used for data transfers.
 - (C) 2 wires are used to power external devices and the other 2 form a single communication channel.
 - (D) Each pair of wires forms a communication channel - one channel in each direction.
9. Which of the following is true for a PCI bus?
 - (A) It allows many devices to share a single IRQ line.
 - (B) Only 1 device can use the bus at a time.
 - (C) It is a parallel synchronous bus.
 - (D) All of the above.
10. USB frames always include data packets for:
 - (A) interrupt transfer devices.
 - (B) bulk transfer devices.
 - (C) isochronous transfer devices.
 - (D) All of the above.
11. Identify and describe the purpose of the three major components within the system bus?
12. What is the purpose of each of the following?
 - (a) IRQ lines.
 - (b) DMA lines.
 - (c) The system clock.
13.
 - (a) Identify all the ports and slots present on your home or school computer.
 - (b) Identify the port or slot that connects each peripheral device to the motherboard on your home or school computer.
14. Describe how data is transmitted to a printer connected to a USB port?
15. PCI and USB are able to support multiple devices using quite different techniques. Compare and contrast the techniques used by PCI and USB to support multiple devices.

THE ROLE OF MODEMS

The term modem is a shortened form of the terms modulation and demodulation, these are the primary processes performed by all modems. Today most modems are used to connect a computer to a local Internet Service Provider (ISP), as shown in *Fig 8.30*, the ISP supplying a high-speed connection to the Internet. However, dial-up modems are also used to connect computers directly to other computers, to the Internet and also to facsimile machines. In fact virtually all dial-up modems are able to both send and receive fax transmissions.

So what is modulation and demodulation? Modems modulate digital signals by altering the phase, amplitude and/or frequency of electro-magnetic waves. That is, modulation is the process of encoding digital data into an analog waveform. Similarly, modems reverse the modulation process by demodulating received analog waves into a digital form suitable for use by the computer. That is, demodulation decodes analog signals back into their original digital form. Clearly both sender and receiver must agree on the method of modulation used if communication is to be successful.

Modems are commonly connected to a computer's system bus via a PCI slot, USB port or a network interface. All these interfaces are considered digital links, they do use electromagnetic waves however bits are represented distinctly using high and low voltages. These voltage changes are suitable for use by the electronic circuits within the computer. In contrast, analog waves, such as those transmitted down telephone lines or coaxial cables, are not suitable for direct use by the circuits within the computer. Hence the primary role of modems is to provide an interface between analog waves used for long distance transfer and digital signals suitable for use by computers.

The concepts of baud rate and bandwidth are central to understanding the nature of analog signals. In this section we examine three common types of modem, namely traditional dial-up modems, ADSL modems and cable modems. All these modems transmit and receive at a particular baud rate within a particular bandwidth.



Modulation

The process of encoding digital data onto an analog wave by changing its amplitude, frequency or phase.



Demodulation

The process of decoding a modulated analog wave back into its original digital signal. The opposite of modulation.

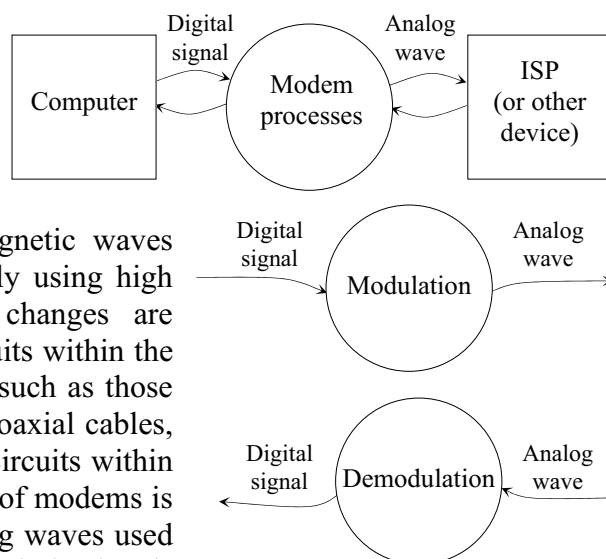


Fig 8.30

Context and dataflow diagrams describing the operation of a modem.



GROUP TASK Research

Survey members of your class in regard to their Internet connection. Record the type of modem and the interface used to connect the modem to their computer. Discuss how these results would have differed 5 years ago and are likely to differ in 5 years time.

Dial-up modems

Dial-up modems transfer data over standard telephone lines. These lines were designed for voice communication, primarily speech, and therefore many of the switching devices present within the telephone network filter out frequencies below about 200Hz and above about 3400Hz. As a consequence traditional dial-up modems must operate within a restricted bandwidth of around 3200Hz.

Before we commence describing how these modems modulate and demodulate we need to clear up one common misconception, standard telephone lines do NOT use two separate wires for transmitting and receiving. They do use two physical wires, however these two wires form a single circuit composed of an active wire and a ground. When talking to someone both voices are present on this single line. Circuitry between the microphone and the speaker within your own phone filters out your own voice, consequently you only hear the other person's voice through your phone's speaker.

Transmitting overlapping frequencies in both directions is fine for voice communication. The filtering process does not need to be that precise; even if some detail within the received signal is lost the overall voice remains intact. Up until the late 1980s filters were not sufficiently sensitive for use when transferring data. Hence older modems achieved full-duplex data communication by essentially splitting the bandwidth into two distinct channels. The modem dialling a number, called the originate modem, was assigned one frequency range and the modem answering a call, called the answer modem, was assigned the remaining frequency range. Such a system effectively halved the total bandwidth available for data transfers. Some modems operated symmetrically, meaning sending and receiving operated at the same speed, and others operated asymmetrically, where a larger portion of the bandwidth was allocated for downloads. For example, during the mid 1980s 1200/75bps was a common modem speed, data was received at 1200bps and transmitted at 75bps. This was acceptable at the time as 75bps equated to roughly 7 characters per second, a fairly reasonable typing speed. As most modems were used in conjunction with dumb terminals the only data sent had to be typed via the keyboard. Data was being received by these dumb terminals at around 120 characters per second, a reasonable reading speed. Clearly these speeds were unacceptably slow when personal computers became common and users wished to transfer large files.



Consider the following:

From the 1960s up to the early 1980s modems operated at 300bps and 300baud. These 300bps modems transferred data in full-duplex using 4 different tones (frequencies). Such modems were connected to terminals and computers via a standard, but slow, RS232 serial port. During the transmission of data the originate modem sends a 1070Hz tone to transmit a 0 and a 1270Hz tone for each 1. The answer modem, when transmitting, uses a 2025Hz tone for each 0 and a 2225Hz tone for each 1. Therefore when receiving it is a simple matter of listening for the two tones transmitted by the other modem and converting them into negative voltage for 1s and positive voltage for 0s.



GROUP TASK Discussion

Describe what occurs during modulation and demodulation of a typical 300bps full-duplex transmission.

Since the early 1990s dial-up modems have used the same frequency range for both sending and receiving. The filtering system is able to detect the difference between the weaker received signal and the stronger signal being sent, essentially the filter removes the stronger signal during the receiving process.

During the 1990s dial-up modem speeds increased from 9600bps, through 28.8kbps, 33.6kps and then finally the current standard 56kbps. All these speeds, apart from 56kbps, were achieved by modulating the signal within the 200 to 3400Hz range. Due to the inherent noise present on standard phone lines the maximum baud rate possible is 3200 baud, commonly achieving 3200 baud is not possible and modems revert to 3000 baud or even 2400 baud. These speeds are achieved using quadrature amplitude modulation (QAM); we discussed QAM on page 288. Essentially the amplitude and phase of the wave is altered to represent multiple bits within each signal event (baud).

Theoretically 33.6kbps is the absolute maximum speed possible using the 3200Hz bandwidth available over standard analog telephone connections. So how do 56kbps modems achieve higher speeds? They capitalise on the fact that the only analog part of a complete transmission is between the modem and the local telephone exchange, all other intermediate links being high-speed digital. These digital links, although still designed for voice, provide a much cleaner signal. Thus if the appropriate digital hardware is present at your local telephone exchange and your ISP has a dedicated digital link into this digital phone network then the only analog link is between the telephone exchange and your home. A different coding system, called pulse code modulation (PCM) is used. PCM is a system similar to audio sampling where the amplitude of the signal is sampled at precise intervals. To achieve such precision a clock signal is transmitted along with the data, the modem uses this clock signal to synchronise the sampling process. Approximately 8000 samples are detected each second and each sample represents 7 bits, hence 8000 times 7 results in 56000bps. Although 56kbps is theoretically possible, in reality most connections do well to achieve speeds of around 48kbps.



Fig 8.31
A Netcomm 56kbps
external modem



Consider the following:

When a dial-up modem wishes to connect to another dial-up modem it first listens for a dial tone and then dials the telephone number of the answering modem. The phone company transmits the dial tone, in fact the phone company supplies all power on the line. Dialling involves transmitting a sequence of different frequencies corresponding to the digits within the phone number, for example to dial the digit 1, a frequency of 1209Hz and a frequency of 696Hz is transmitted. The telephone network interprets these frequencies, sets up a circuit between the two lines and transmits a ring tone down the line. In Australia, the ring tone is a 25Hz alternating wave, however different countries use slightly different frequencies. The answering modem detects the ring tone on the line and commences communication with the originating modem. This function is known as 'automatic answer' and is a feature of all dial-up modems. The two modems then begin negotiating in regard to baud rate, and the number and nature of the symbols used per baud. Both the number of symbols and the baud rate are progressively reduced until accurate data transfers are achieved. Once this 'handshaking' process is complete data transfer can commence.

**GROUP TASK Discussion**

Should it be possible for two 56kbps modems to connect at 56kbps? Discuss.

**GROUP TASK Discussion**

Should it be possible to connect two modems directly using a standard phone cable? Discuss.



Consider the following:

Virtually all current dial-up modems are able to transmit and receive facsimiles. In fact, a modern fax machine is essentially a scanner, modem and printer combined into a single unit. Also multi-function devices are available that connect to computers and integrate the functions of a scanner, modem and printer.

So does the transmission of faxes differ from other types of data transmission? In general terms, no, the only significant difference being the details in regard to how the digital data is encoded. The International Telegraph and Telephone Consultative Committee (CCITT) is responsible for maintaining the rules for encoding fax transmissions. Currently most faxes communicate using the CCITT Group 3 standard, although a Group 4 standard has been released. The Group 3 standard includes speeds of up to 14400bps where QAM is used to encode the data. Lower speeds are used when noise is present on the line or the receiving fax complies with a lower CCITT standard. The CCITT standards also specify precise details in regard to the compression of data prior to encoding. Clearly modems capable of transmitting and receiving faxes must be aware of these CCITT standards.

**GROUP TASK Discussion**

Discuss how a computer and a modem can be used to transmit and receive faxes.

**GROUP TASK Discussion**

Should it be possible to connect a modem directly to a dedicated fax machine and use the fax machine like a printer? Discuss.

ADSL modems

Asymmetrical digital subscriber lines (ADSL) use existing copper telephone lines to transfer broadband signals. Although these copper wires were originally designed to support frequencies from 200 to 3400Hz, they are physically capable of supporting a much wider range of frequencies. It is the various switching and filtering hardware devices within the standard telephone network that prevent the transfer of frequencies above about 3400Hz. To solve this problem requires dedicated hardware to be installed where each copper line enters the local telephone exchange.

If the telephone line from your house does not connect directly to the local exchange then ADSL will not be available. For example, the signals on multiple copper wires are often combined onto a single cable. When this occurs between a building and the local telephone exchange then ADSL will not be available.



*Fig 8.32
2-pair copper
telephone cable.
Each pair
supports a single
telephone line.*

ADSL signal strength deteriorates as distances increase, the signal cannot be maintained at all for distances greater than about 5400 metres. Voice lines much greater than 5400 metres are possible using amplifiers. Unfortunately these amplifiers boost only the lower frequencies required for voice, hence ADSL is not currently available in many remote areas. Even when distances are short and the copper runs directly into the exchange problems can occur as a consequence of interference. In general, phone lines within a building and out to the street are not shielded against interference. This interference is rarely significant enough that a connection cannot be established, however it often reduces the speed of such connections.

So how does ADSL transfer data between an ADSL modem and the local telephone exchange? Using a modulation standard known as Discrete MultiTone (DMT). DMT operates using frequencies from about 20kHz to around 1.1MHz for the ADSL and up to 2.2MHz for ADSL2+.

This bandwidth is split into hundreds of individual 4kHz wide channels as shown in Fig 8.33. Each channel is modulated using QAM. DMT's task is to specify the channels that are used for actual data transfer. If interference is present on a particular 4kHz channel then DMT will shut down that channel and assign a new channel. This channel switching occurs in real time and is completely transparent to the user. In a sense ADSL is like having hundreds of dial-up modems all working together, each modem using QAM and DMT ensuring they all work together efficiently. The ADSL modem and the DSL hardware at the telephone exchange communicate to agree on the channels currently being used.

At the local telephone exchange all the copper wires from the neighbourhood are connected to a splitter (see Fig 8.34). This splitter directs the 0-4kHz frequencies to the normal telephone network and the higher ADSL frequencies to a DSL Access Multiplexor (DSLAM). The DSLAM (see Fig 8.34) performs all the DMT negotiations with individual ADSL modems and directs data to and from ISPs, where it heads onto the Internet. The term *multiplexor* simply refers to the DSLAM's task of combining multiple signals from customers onto a single line and extracting individual customer signals from this single line.

In most ADSL systems the lower bandwidth ADSL channels are used for upstream data (from modem to exchange) and higher frequency channels are used for downstream data (exchange to modem).

Some channels are able to transfer data in both directions. ADSL is one example of a DSL technology, the A stands for asymmetrical, meaning communication in each direction occurs or can occur, at a different speed. At the current time typical ADSL connections in Australia achieve speeds of 512kbps upstream and 1.5Mbps downstream and ADSL2+ achieves 1Mbps upstream and up to 20Mbps downstream.

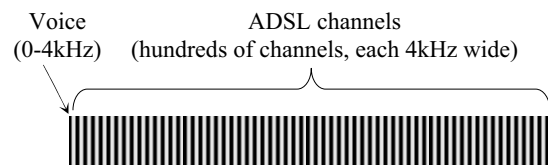


Fig 8.33
ADSL splits higher frequencies into hundreds of channels, each 4kHz wide.



Fig 8.34
A splitter (left) and DSLAM (right).



GROUP TASK Activity

Construct a diagram showing the major hardware components present in a typical ADSL connection. Describe the function of each component.



Consider the following:

When first installing an ADSL connection it is necessary to install one or more low-pass (LP) filters. Sometimes a single filter is installed where the phone line enters the premises. In this case a qualified technician is required to install a dedicated ADSL line from the LP filter to the location of the ADSL modem. In other cases, the user installs a separate LP filter, like the one shown in Fig 8.35, between each telephone and wall socket.



Fig 8.35
Inline LP filter.



GROUP TASK Discussion

What is the function of an LP filter? Describe how the two LP filter installation methods described above achieve the same outcome?

Cable modems

Earlier in this chapter (p289) we examined the specifications for a particular cable modem, the Motorola Surfboard SB4200. During this discussion we noted that cable modems use a single 6MHz bandwidth channel for downstream data; 6MHz being the width of a single cable TV station. This 6MHz wide channel is assigned within the range 88 to 860 megahertz. A narrower bandwidth channel is used for upstream, commonly 1.6MHz wide however various other bandwidths are supported ranging from 200kHz to 3.2MHz. The upstream channel is assigned within the range 5 to 42 megahertz. The particular frequencies used for both channels are determined by the cable Internet provider and cannot be altered by individual users.

The bandwidth used in a cable system is significantly larger than that used for ADSL. Hence, one would imagine the rate of data transfer would be much larger. In reality cable connections achieve speeds similar to ADSL connections. Why is this? Cable connections are shared amongst multiple users. A single 6MHz downstream channel is likely to be shared by hundreds of users. In a sense, all the cable modems sharing a particular channel form a local area network. Every cable modem within the network receives all messages, they just ignore messages addressed to other modems. Consequently when only a few users are downloading then higher speeds are possible than when many users are downloading. Clearly the same situation occurs when uploading. This is why cable Internet companies include statements within their conditions stating that speeds quoted are not guaranteed.

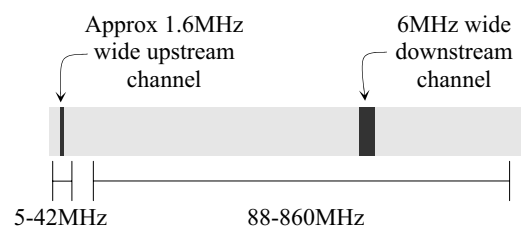


Fig 8.36

Cable modems share a bandwidth of 6MHz downstream and a lower bandwidth upstream.

Most current cable modems comply with standards based on the *Data over cable service interface specifications* (DOCSIS). These DOCSIS standards allow cable modems manufactured by a variety of different companies to be used on different cable networks. Each individual DOCSIS modem is only able to communicate directly with the *Cable Modem Termination System* (CMTS) located at the cable Internet provider's premises. The CMTS provides the connection between the cable network and ISPs, the ISPs transmit and receive data to and from the Internet. The CMTS performs a similar task as the DSLAM in an ADSL system.



Consider the following:

Cable modems connect using coaxial cable whilst ADSL systems use standard copper telephone wires. Coaxial cable is shielded to exclude outside interference and also to ensure the integrity of the signal.



GROUP TASK Discussion

ADSL uses DMT and many small bandwidth channels, whilst cable uses QAM and two relatively large bandwidth channels. Discuss reasons for these differences in terms of the transmission media used by each system.



Consider the following:

At the time of writing many cable Internet providers offered plans whereby the upstream speed is reduced when a specified download limit has been exceeded. Many business plans offered higher upload speeds than most home plans.



GROUP TASK Discussion

How can cable Internet providers alter upload speeds? And can you suggest techniques they could use to increase download speeds? Discuss.



Consider the following:

Back in Chapter 3 (p96) we described the operation of a simple analog to digital converter (ADC). The principles underlying our simple ADC are similar to those used when demodulating analog signals received by modems. But what about the modulation process that converts digital signals to analog? Fig 8.37 describes a simple digital to analog converter (DAC). This example DAC converts each 4-bits into an analog signal with varying amplitudes (voltages). Hence there are 16 different symbols represented by 16 different voltages.

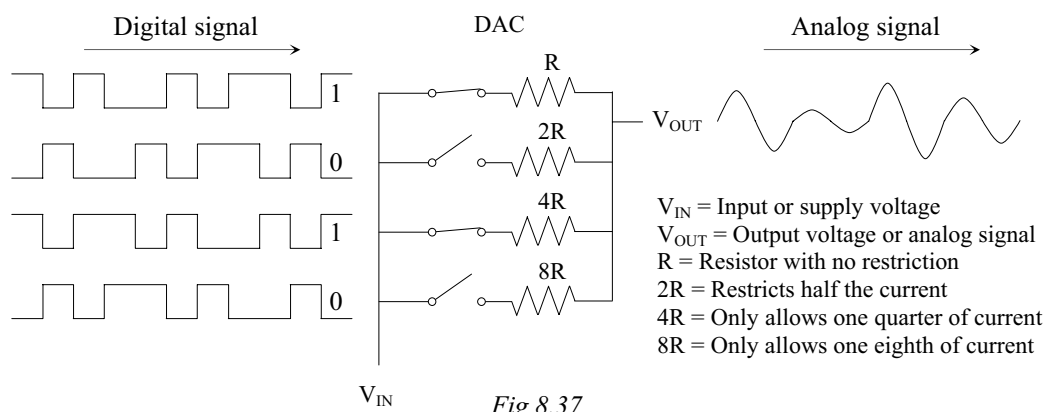


Fig 8.37

A simple binary weighted DAC uses weighted resistors to alter the signal's output voltage or amplitude.



GROUP TASK Discussion

Assume the above DAC has an input voltage of 2.5V. Make up a table listing the voltages output for all possible 4-bit combinations. Identify specifications necessary to accurately receive this analog signal.

LOCAL AREA NETWORKS (LANs)

Local area networks are used to connect computers and other devices that are physically close to each other, usually within the same building or within adjoining buildings. Each individual computer or other device connected to a LAN is called a node. For communication between nodes to occur, all nodes must agree on a precise method of transmission. A specific set of such communication rules is called a protocol. In most cases a single LAN operates using the same protocol, hence nodes connected to a LAN are usually able to transmit and receive messages to and from all other nodes without the need to alter the organisation of the messages.

For a LAN to operate each node must be able to encode and also decode messages using the LAN's protocol. The hardware device that performs this function is called a network interface card or NIC. The messages are transferred along a transmission medium, such as twisted pair, optical fibre, coaxial cable or wirelessly through the air. These two hardware components, NICs and the transmission media are sufficient for a simple LAN connecting two nodes. However when many nodes are present further devices, such as hubs, switches, bridges, wireless access points and routers are used to assist in monitoring and moving messages across the transmission medium.

In the HSC course, we examine the detailed operation of various protocols and LAN hardware devices. In the Preliminary course, we restrict our discussion to a brief description of a selection of device's within a LAN. Let us consider the purpose of NICs, hubs, switches, bridges and routers.

Network Interface Cards (NICs)

NICs are used to convert messages between the computer and the LAN. They also ensure messages are only transmitted when the communication channel is free. Each NIC has a unique number, called a Media Access Controller (MAC) address or physical address. This address is hardwired into the NIC at its time of manufacture and is ultimately what is used to identify each individual node on a network.

Most NICs are connected to computers via the PCI bus using an expansion slot (see Fig 8.38) or embedded on the motherboard itself.



Fig 8.38

A PCI Ethernet network interface card.

Hubs

Hubs are used to connect more than two nodes together. When a hub receives a message on one wire it merely sends it down all other connected wires. Hubs are dumb devices, they make no attempt to understand the messages. Hubs are also called multi-port repeaters; this alternative name is really a better description of their function. Hubs operate at the physical layer, they don't examine any of the data within messages at all.

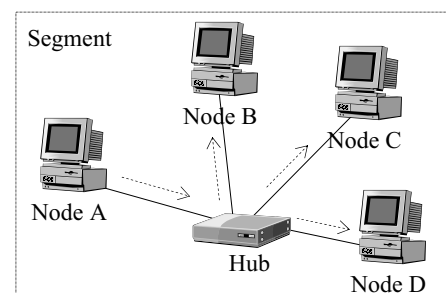


Fig 8.39

Hubs repeat all messages to all nodes on a single LAN segment.

When a hub is used to connect nodes only one node is able to transmit at a time. Each hub (or switch), together with its attached nodes forms a single segment. Using LAN terminology, a segment is a collection of devices that all share a single line, all nodes receive all messages transmitted on that segment; they just ignore messages that are addressed to other nodes.

Switches

A switch can be thought of as an intelligent hub. Switches determine the MAC address of the sender and intended receiver that precedes each message. The receiver's address is used to identify the destination node and forward the message to that node only. In essence, a switch sets up a direct connection between the sender and the receiver. Hence each node exists on its own segment, the switch being the only other device on the segment. As no other nodes exist on each segment each node is free to transmit messages at any time.

Most switches can simultaneously receive and forward messages from and to multiple pairs of nodes. As long as both the sender and the receiver of each message do not conflict with other simultaneous messages then the switch will direct the message correctly. Many switches allow nodes to communicate in full duplex. In *Fig 8.40*, Node A is sending a message to Node B whilst it simultaneously receives a message from Node D, neither message is ever present on Node C's segment. Switches significantly reduce the amount of traffic flowing over each wire resulting in vastly improved transfer speeds compared to speeds achieved using hubs.

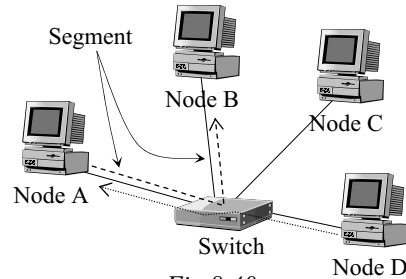


Fig 8.40

Switches forward messages to the destination node only. Each switch – node connection forms a segment.

Routers

Routers are even more intelligent than switches, they specialise in directing messages over the most efficient path. Routers include the functionality of a gateway. They are able to communicate with networks that use different protocols and even completely different methods of communication. A router operates at a higher level than a switch, they do not use the MAC address of NICs, rather they use the protocol address of each machine to determine its location. For example, many routers identify the destination of messages using IP (Internet Protocol) addresses.

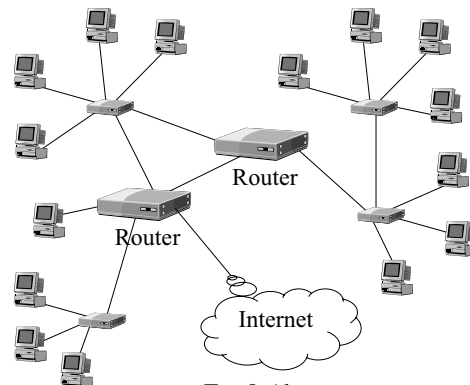


Fig 8.41

Routers forward messages over the most efficient path and can alter this path as needed.

Many routers include a variety of different security features. They are able to block messages based on the sender's address, block access to specific web sites and even restrict communication to certain protocols.

Routers learn the layout of networks surrounding them by communicating with other routers. Based on this information the router determines the most efficient path for each message. However, should any connections within the most efficient path fail then routers automatically direct the message over an alternate path. On larger wide area networks, and in particular the Internet, thousands of routers work together to pass messages to their final destination.



GROUP TASK Research

Many hardware devices integrate the functionality of switches, routers, wireless access points and modems within a single device. Such devices are often just referred to as modems or as routers. Research examples of such devices and determine their built-in functionality.

WIDE AREAS NETWORKS (WANS)

Wide area networks connect computers over large physical distances, the Internet being the largest example. Because of the large distances and massive amounts of data involved most WANs utilise fast broadband links. Governments together with large communication and telecommunication companies provide such links. Each individual link is of high quality and has a wide bandwidth, many able to achieve transfer rates of many hundreds of gigabits per second. However, due to the enormous amount of data being transferred between an ever increasing number of users, most WAN links are generally slower for end-users than their local LAN links.

In this section we briefly introduce some of the common types of communication links used by WANs, in the HSC course we examine such links in more detail. To assist our discussion let us consider the communication links and substantial entities present in a typical connection between a computer communicating with a web server over the largest WAN in the world, the Internet.

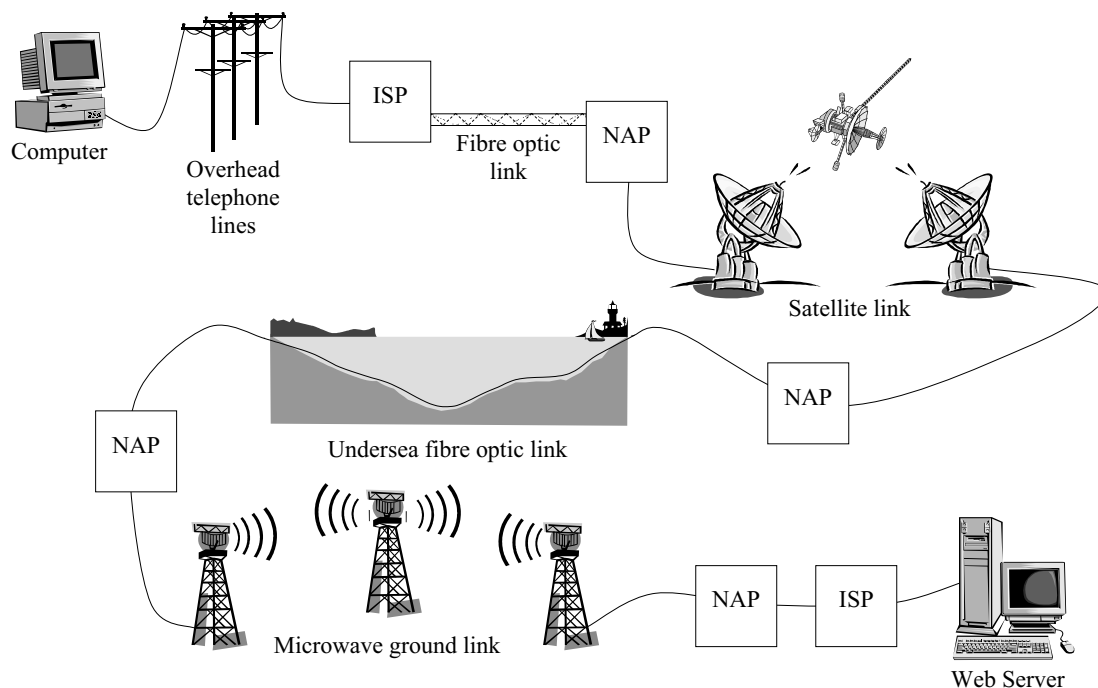


Fig 8.42

One possible communication link between a computer and a web server.

In Fig 8.42 above, both the computer and the web server communicate with their local ISP (Internet Service Provider). Each ISP has at least one Point of Presence (PoP), a PoP being composed of all the communication equipment that allows individual users (or companies) to connect to their ISP. Remember that between each PoP and user connection there may also be various devices, such as DSLAMs for DSL connections, banks of modems for dial-up connections and CMTSs for cable connections. PoPs need to be located relatively close to the ISP's customers.

Each ISP connects to a Network Access Point (NAP); sometimes called an Internet Exchange (IX). Typically a NAP provides connections between many different ISPs and also provides high-speed connections to other adjoining NAPs. Each ISP and NAP contains powerful routers to ensure each message takes the most efficient path to its destination. Connections between NAPs span every continent in the world. For example, undersea fibre optic links exist between all continents except Antarctica. Governments or very large communication companies provide and maintain the physical communication links between NAPs.



GROUP TASK Practical Activity

The Ping command sends a single message to a server and requests a response. For example, typing the command ping se.yahoo.com at a Windows command prompt will ping the Yahoo server in Sweden. The time taken to respond is then reported. Try Pinging a server located on every continent on earth and record the average time for the return trip.

In Fig 8.42 there are a number of different types of link that we are yet to discuss, namely fibre optic, microwave ground links and satellite links.

Fibre optic links

Most optical fibres transmit ultra-violet light. UV light is a form of electromagnetic radiation or wave that occupies a frequency range from 7.5×10^{14} Hz to 3×10^{17} Hz, or 750,000GHz to 300,000,000GHz, which is just above visible light. Messages are transmitted as pulses of light down extremely pure glass fibres thinner than a human hair. These fibres are covered in a reflective coating, hence the light pulses reflect internally within the fibre. Modulation of digital data down optical fibres involves altering the characteristics of each light pulse to represent different bit patterns. This is much the same process as that used to modulate signals for transmission down copper wires.

Microwave ground and satellite links

Satellite links use microwaves, hence their operation is similar to microwave ground links. Microwaves are high frequency electromagnetic waves occupying frequencies from 3×10^9 Hz to 3×10^{12} Hz, or between 3GHz and 3000GHz. Microwaves can only travel in straight lines, thus receivers and transmitters must be located in direct line of site. Microwaves are used for both land and satellite based communication. The devices that receive and transmit microwaves are called transponders, which is a shortened form of the words transmitter and responder. Transponders receive messages, amplify them and then transmit them on to the next station.

For ground-based applications a series of transponders is able to cover large distances. Each transponder is located at the highest point on the terrain, often the top of a mountain or hill in rural areas or the top of a tall building in urban areas.

Almost all communication satellites are geostationary. This means they orbit at precisely the same speed as the Earth and are normally located directly above the equator. As a consequence, geostationary satellites appear stationary when viewed from Earth. Satellite dishes are essentially microwave transponders that are directed at a particular satellite. As most satellites are above the equator then in the southern hemisphere all such dishes point in a northerly direction, and in the northern hemisphere they point in a southerly direction.



GROUP TASK Research

Using the Internet or otherwise determine the maximum distances between microwave transponders. Now observe actual microwave ground transponders located around your local area. Does the distance between these transponders appear to comply with your research?



GROUP TASK Practical Activity

Observe satellite dishes installed around your local area. Do all these dishes point towards the equator? If you have a satellite TV or Internet connection, determine the precise direction of your dish?

SET 8C

1. What is demodulation?
 - (A) The process of converting an analog wave into a digital signal.
 - (B) The opposite of the modulation process.
 - (C) A process that determines the original digital signal within an analog signal.
 - (D) All of the above.
2. Why do dial-up modems operate within a bandwidth of just 3.2kHz?
 - (A) Copper wire only supports frequencies from about 200Hz to 3400Hz..
 - (B) Switching hardware within standard telephone networks filters out frequencies above about 3400Hz.
 - (C) Filters within dial-up modems are not sensitive enough to detect frequencies above 3400Hz.
 - (D) 3.2kHz is the bandwidth needed to support transfer speeds of 56kbps.
3. A modem transmitting at 1200baud is able to transfer data at a speed of 9600bps. How is this achieved?
 - (A) Each signal event represents 256 bits.
 - (B) There are 8 different wave patterns that can be transmitted.
 - (C) Each signal event represents 8 bits.
 - (D) The bandwidth is 8kHz.
4. When a dial-up modem dials a telephone it becomes the:
 - (A) answer modem.
 - (B) receiving modem.
 - (C) originate modem.
 - (D) transmitting modem.
5. Which term means transmitting and receiving occur at different speeds?
 - (A) asymmetrical
 - (B) symmetrical
 - (C) multiplexing
 - (D) broadband
6. What is the function of a splitter within an ADSL system?
 - (A) To combine the signals from many lines into a single high-speed signal.
 - (B) To connect many individual lines to a single DSLAM.
 - (C) To direct voice frequencies to the standard telephone network and higher frequencies to the DSLAM.
 - (D) To assign a particular bandwidth to each customer's ADSL modem.
7. In regard to connections between modems and service providers, which of the following is true?
 - (A) Many customers share cable bandwidth.
 - (B) Many customers share ADSL bandwidth.
 - (C) Many customers share dial-up bandwidth.
 - (D) All of the above.
8. Which device directs messages based on MAC addresses?
 - (A) NIC
 - (B) hub
 - (C) switch
 - (D) router
9. Which of the following are used to transfer electromagnetic waves?
 - (A) copper wires.
 - (B) satellites.
 - (C) optical fibres.
 - (D) All of the above.
10. The main purpose of a NIC is to:
 - (A) connect two computers together.
 - (B) convert messages between a computer and a LAN.
 - (C) direct messages on a LAN to their destination.
 - (D) ensure that each node has a unique MAC address.
11. In regard to dial-up modems:
 - (a) Describe how they are able to transmit and receive at speeds greater than their baud rate.
 - (b) Explain how a dial-up modem is able to send faxes to a standard fax machine.
12. Identify and describe the function of the major hardware devices present between a computer and an ISP when using:
 - (a) a dial-up connection (b) an ADSL connection (c) a cable connection.
13.
 - (a) Explain how a binary weighted DAC operates.
 - (b) Explain how a DAC can be used within an ADC (analog to digital converter).
14. ADSL and cable systems allocate bandwidth to individual customers differently. Compare and contrast these two systems in terms of their allocation of bandwidth.
15. Explain the differences between LANs and WANs. Use examples of different types of connection to assist your explanation.

SOFTWARE FOR TRANSMITTING AND RECEIVING

Software is used to control and direct the operation of hardware. In regard to transmitting and receiving information processes there are at least two different hardware devices involved, namely the sender and the receiver. These devices, whether they are peripheral devices, network devices or even interfaces within a computer, must agree on how the hardware will be used to transfer messages. This is not a simple matter, a wide variety of applications transfer data using a wide variety of operating systems, protocols, devices and transmission media. In 1978 a set of standards was first developed by the International Standards Organisation (ISO) to address such issues. These standards are known as the *Seven-Layer Model for Open Systems Interconnection* or more simply as the *OSI Model*. This seven-layer model has been accepted and used by network engineers when creating all types of transmission hardware and software.

In the IPT HSC course we split the seven OSI layers into three levels (refer Fig 8.43). The IPT Application Level includes OSI Layers 6 and 7, the IPT Communication and Control Level includes OSI Layers 3, 4 and 5, and the IPT Transmission Level includes OSI Layer 1 and 2. In this section we briefly introduce the seven layers of the OSI model. We then examine particular examples of transmitting and receiving software, namely:

- Software that interfaces with hardware.
- Software applications for transferring text, numeric, image, audio and video, including electronic mail.

THE SEVEN-LAYER MODEL FOR OPEN SYSTEMS INTERCONNECTION (THE OSI MODEL)

Most of our work so far in this chapter has involved transmission hardware. The hardware actually used for transmission resides within Layer 1, the physical layer. The physical layer includes NICs, hubs and the various types of wired and wireless transmission media. These components actually physically move the data. How they do this is determined by the higher software layers. Each layer interfaces with the layer above it and the layer below it.

The seven layers are often referred to as the OSI stack. Each packet of data must descend the stack, be transmitted and then ascend the stack on the receiving computer. To explain the general tasks performed by each layer consider the transmission of a message:

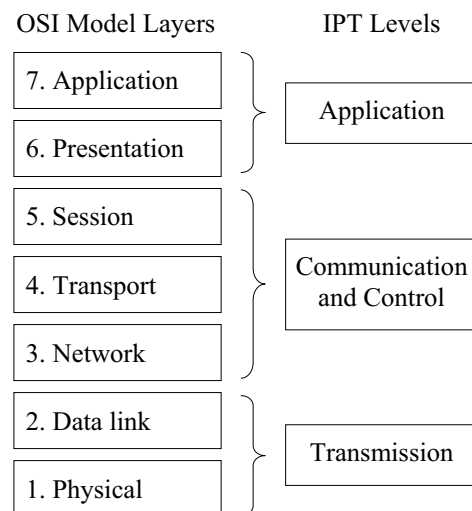


Fig 8.43

Comparison of the seven layers of the OSI model with the three levels in IPT HSC.

7. Application – The actual data to be transmitted is created by a user within a software application, this data is organised in a format understood by the application that will receive the data.
6. Presentation – The data is reorganised into a form suitable for subsequent transmission. For example, compressing an image and then representing it as a sequence of ASCII characters. The presentation layer is commonly part of the application or is executed directly by the application. Protocols operating at this level include HTTP, DNS, FTP, SMTP, POP and IMAP.

5. Session – This is where communication with the network commences and is maintained. It determines when a communication session is started and also when it ends. For example, when performing an Internet banking transaction it is the session layer that ensures communication continues until the entire transaction is completed. Layer 5 also includes security to ensure a user has the appropriate access rights.
4. Transport – The transport layer manages the correct transmission of each packet of data. This layer ensures that packets failing to reach their destination are retransmitted. For example, TCP (Transport Control Protocol) operates within layer 4. TCP is used on TCP/IP networks, such as the Internet, to ensure the correct delivery of each data packet.
3. Network – This is where packets are directed to their destination. IP (Internet Protocol) operates here, its job is to address and forward packets to their destination. Routers also operate at this layer by directing packets along the best path based on their IP address. Routers often have their software stored in flash memory and can be configured remotely from an attached computer.
2. Data link – This layer defines how the transmission media is actually shared. Device drivers that control the physical transmission hardware operate at this layer. They determine the final size of transmitted packets, the speed of transfer, and various other physical characteristics of the transfer. Switches operate at this level, directing messages based on their destination MAC address.
1. Physical – This layer performs the actual physical transfer, hence it is composed solely of hardware. It converts the bits in each message into the signals that are transmitted down the transmission media. Most of the previous section on hardware involved processes occurring within the physical layer.



GROUP TASK Discussion

The OSI model aims to standardise the design of communication hardware and software across the IT industry. What are the advantages of such standardisation? Discuss.

SOFTWARE THAT INTERFACES WITH HARDWARE

Between all hardware devices and software applications are various levels of other software tools. The OSI model describes these levels in terms of network communication, however similar levels exist for communication between devices within a single computer and between its attached peripherals. For example, a printer is controlled by a device driver that performs functions similar to software in the lower layers of the OSI model. The Basic Input Output System (BIOS) is software that controls the lowest level of communication between hardware located on the motherboard, in this respect the BIOS performs functions similar to the OSI data link layer. Software at these lower levels commonly includes a user interface provided to configure various hardware settings. Although these configuration screens are merely a window into the software, they do provide insight into the operation of the actual interface software itself.

Basic Input Output System (BIOS)

The BIOS is contained on a ROM or flash memory chip and is loaded prior to the operating system as the computer boots. The various configuration settings used by the BIOS are stored on a CMOS (Complementary Metal Oxide Semiconductor) chip, which is powered by a small battery located on the motherboard. Hence on many

machines the software that configures various BIOS settings is often called a CMOS setup utility, an example of such a screen is shown in *Fig 8.44*.

The BIOS provides a standard method for the operating system to communicate with various hardware interfaces and devices. For example, the settings on the screen in *Fig 8.44* indicate that the hard disk will be accessed using logical block addressing (LBA) and has a capacity of approximately 40 gigabytes. Operating systems and device drivers are written to interact with particular BIOSs. Fortunately BIOSs for particular CPU families are produced by a small number of companies, and in most cases are compatible with all mainstream operating systems designed for these CPUs. Motherboard manufacturers alter the BIOS to suit the different hardware combinations installed on their motherboards. However these BIOSs must retain a consistent interface for operating systems and for device drivers. Clearly it is in the best interests of both BIOS and operating system developers to ensure compatibility between their products.

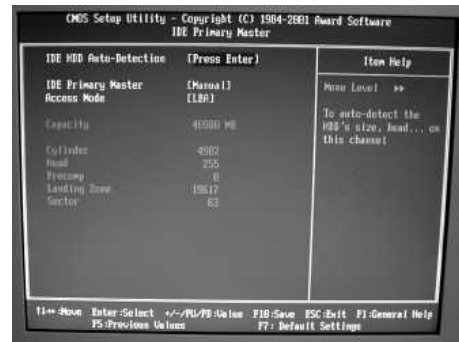


Fig 8.44
Screen from a BIOS setup utility.



Consider the following:

When new technologies first become available it is common for them to be included on motherboards and be supported by the BIOS before operating systems supporting the new technology are available. This problem occurred when USB ports were first introduced. Many computers were sold that physically contained USB ports, together with appropriate BIOS support, yet no operating system was able to access these ports.



GROUP TASK Discussion

Why do you think BIOS support for USB ports was available prior to operating system support for USB ports? Discuss.

Device Drivers

We first discussed device drivers back in Chapter 3 (p103) and then again in Chapter 6 (p226). We stated that a device driver is a program that provides the interface between the operating system and a peripheral device. This is true, however in terms of transmitting and receiving a device driver must send and receive their data over an interface that includes the BIOS. Let us examine some of the common configuration settings that determine how device drivers communicate with hardware via the BIOS.



Consider the following:

Within Microsoft Windows XP device driver details can be viewed and altered via *Device Manager*, which is a software utility included within the operating system. Device drivers, on Windows systems, have a .sys file extension. *Fig 8.45* shows the various driver files used to communicate with a mouse. In this example, specific Logitech drivers are being used in conjunction with the generic HID mouse drivers.

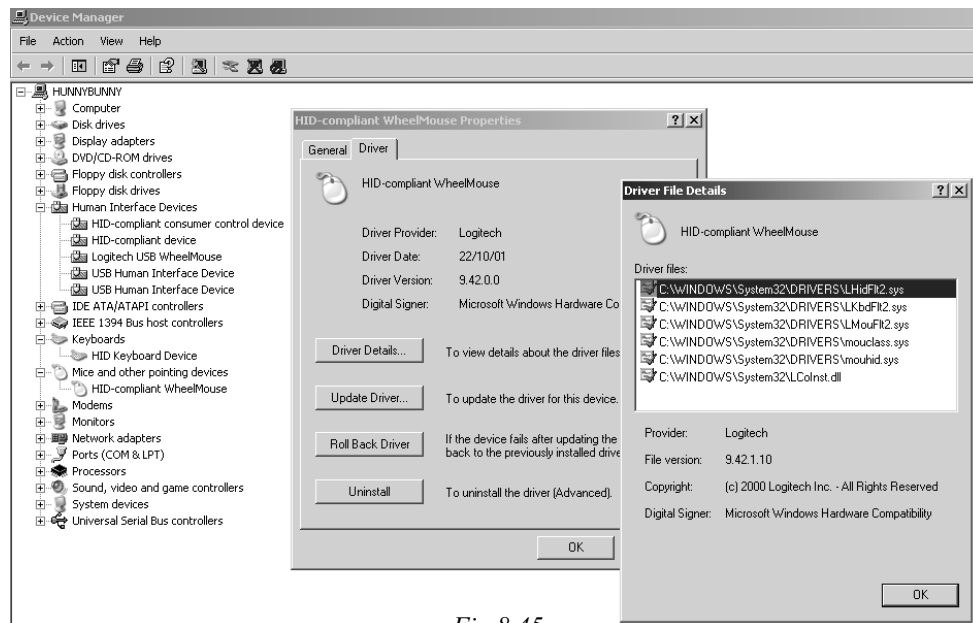


Fig 8.45

Drivers used by an HID compliant mouse in Windows XP.



GROUP TASK Research

Device drivers are not just used for connecting peripheral devices, they are also used to communicate with the PCI bus, USB, firewire, network and HDD interfaces, such as SATA. Examine the configuration screens for such device drivers. Identify settings that specifically relate to transmitting and receiving processes.



Consider the following:

The screens in Fig 8.46 and 8.47 are used to configure a single local area network connection. The network interface card is plugged into a PCI expansion slot on the motherboard. The network cable connects this computer's NIC to a network hub.

All the computers connected to this LAN are able to share files and printers with all other computers. Furthermore, each computer has Internet access via a cable modem that is attached to the USB port on one of the computers attached to the LAN.

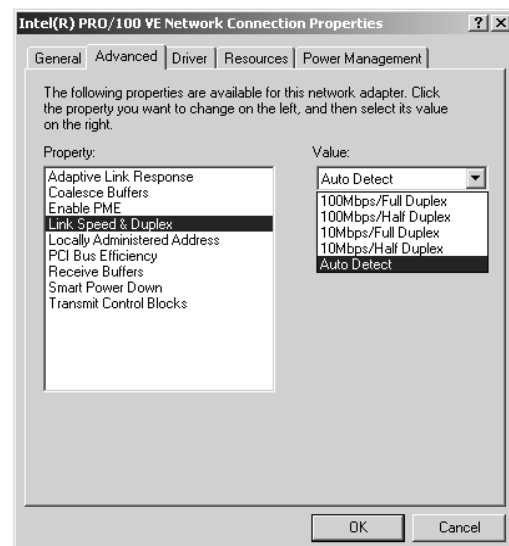


Fig 8.46

NIC configuration settings in Windows XP.



GROUP TASK Discussion

Identify and describe the interface connections being configured in each of the three screens within Fig 8.46 and 8.47.



GROUP TASK Research

What is a gateway and what is a DNS server? What is the purpose of these two settings on the screen in Fig 8.47?

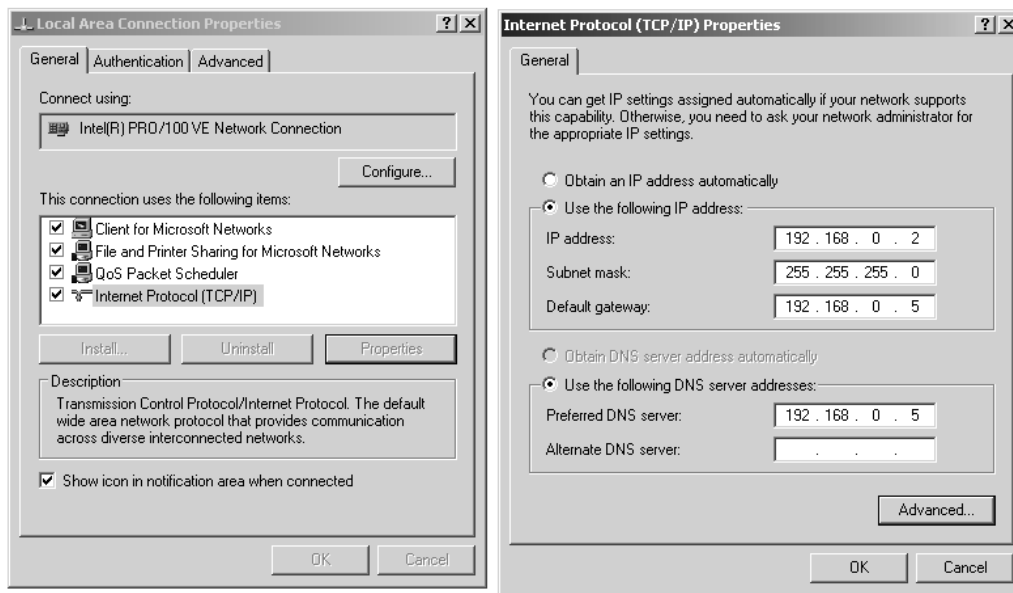


Fig 8.47

Local area network configuration settings within Windows XP.



Consider the following:

The series of screens shown in Fig 8.48 are for a 56kbps dial-up modem installed on a machine running Windows XP. This modem is plugged into the PCI bus via an expansion slot on the motherboard.

Based on our earlier discussions we know that both the PCI bus and also 56kbps dial-up modems do not transfer data asynchronously. However the data bits and stop bits settings on the screen in Fig 8.48 indicate data will be transferred asynchronously. Furthermore, it is possible to choose 1.5 stop bits, how is this possible? Surely you must have either 1 stop bit or 2 stop bits.

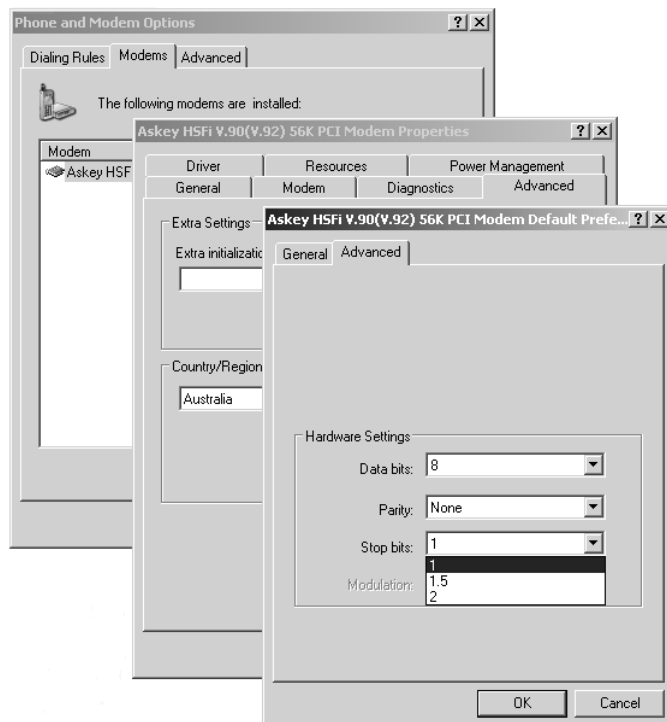


Fig 8.48

Windows XP Driver settings for a dial-up modem.



GROUP TASK Discussion

Describe the various interfaces present between the actual modem and the phone line. Identify the communication link whose settings are being specified on the screen in Fig 8.48.



GROUP TASK Discussion

What is a stop bit? And how is it possible to have 1, 2 or even 1.5 stop bits? Discuss.

SOFTWARE FOR TRANSFERRING TEXT, NUMERIC, IMAGE, AUDIO AND VIDEO

Ultimately transmitting and receiving information processes are about efficiently moving data from one location to another. If only small quantities of data are to be moved then the speed and efficiency of the connection is less significant than for large quantities of data, this applies to data transfers within a computer, across a LAN and more significantly across a WAN such as the Internet.

Most data is stored in files. During a file's transmission the sender splits it into a sequence of chunks or packets, each packet being sent individually. For example a serial port operating asynchronously sends each character as a separate packet, whilst a USB port sends a frame that may contain up to 1500 bytes of a file. The receiver must combine all the received packets back into a complete file. In general, the complete file cannot be displayed until the receiving process is complete. Such a system is fine for smaller files or files containing data that is not dependant on time. However, audio and video files are often large and furthermore they need to be displayed progressively over time. Waiting for an entire audio or video file to be received takes an unacceptable amount of time, thus when real time displaying is required such data is transferred using a system called streaming. This means the file is transmitted and received at a constant rate. Streaming allows the displaying process to commence whilst further packets are simultaneously being received.

In this section we restrict our discussion to three examples of software applications specialising in the transfer of data, namely FTP client software, e-mail applications, and streaming media players and streaming servers. These applications operate at the application and presentation layers of the OSI model.

FTP Client Software

FTP stands for File Transfer Protocol, predictably FTP is a set of rules used to transfer files between computers. Whenever you upload or download files over the Internet your computer is most likely using FTP to negotiate the transfer. Your computer is running FTP client software and the file is downloaded from an FTP server on the Internet, hence FTP is an example of a client-server protocol operating at the presentation layer of the OSI model. Clearly, an active Internet connection is needed before an FTP client-server transfer can take place.

FTP is so often used that support is included within most operating systems in a variety of different forms. For example, current versions of Microsoft Windows operating systems contain a simple FTP program (ftp.exe), this program operates from a command prompt. FTP client functionality is also included within Windows Explorer and Internet Explorer. There are also many dedicated FTP clients available that include a more user-friendly interface, such products are commonly used for uploading website files to web servers.



Consider the following:

The screen shot in *Fig 8.49* is from Ipswitch software's WS_FTP LE ftp client software. It includes views of files on both the local machine and also on the remote machine. In this screen an anonymous connection has been made to the FTP server called ftp.microsoft.com. Many FTP servers allow anonymous users to download files, in fact an anonymous FTP connection is commonly used when files are downloaded over the Internet. Anonymous FTP connections generally do not permit files to be uploaded to or deleted from the server.

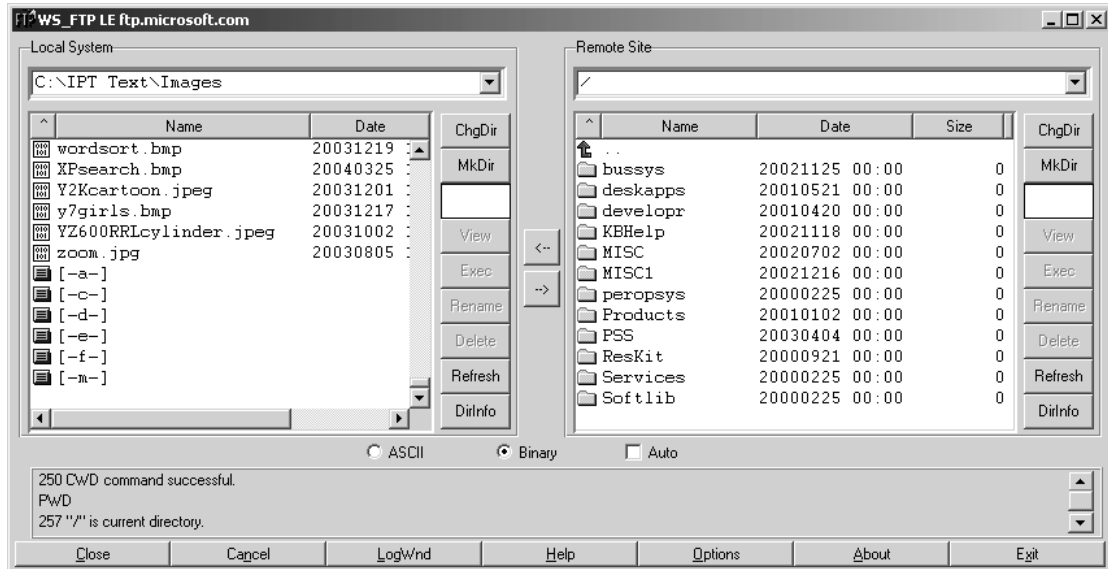


Fig 8.49

Screen from WS_FTP LE, an FTP client application from Ipswitch software



GROUP TASK Practical Activity

Use an FTP client application to connect anonymously to an FTP server. Examine various directories on the server and then copy one or more files from the server to your local machine.



Consider the following:

The program ftp.exe can be executed on a machine running Windows by typing ftp at a command prompt. The screen in Fig 8.50 shows an example of such an FTP session where an image file is being uploaded. The command binary causes the transfer to include all 8 bits within each byte of the file. If the binary command is not used then 7-bit ASCII is assumed.

```

C:\WINDOWS\System32\ftp.exe
ftp> open users.bigpond.net.au
Connected to users.bigpond.net.au.
220 Serv-U FTP-Server v2.4a for WinSock ready...
User (users.bigpond.net.au:(none)): pedc
331 User name okay, need password.
Password:
230 Welcome to the Big Pond Cable Web Hosting service.
230 User logged in, proceed.
ftp> mkdir test
257 "/test" directory created.
ftp> cd
Remote directory test
250 Directory changed to /test
ftp> binary
200 Type set to I.
ftp> put alfa.jpg
200 PORT Command successful.
150 Opening BINARY mode data connection for alfa.JPG.
226 Maximum disk quota limited to 10240 Kbytes
Used disk quota 8051 Kbytes, available 2188 Kbytes
226 Transfer complete.
ftp> 29086 bytes sent in 0.00Seconds 29086000.00Kbytes/sec.
ftp> ls
200 PORT Command successful.
150 Opening ASCII mode data connection for /bin/ls.
alfa.JPG
226 Maximum disk quota limited to 10240 Kbytes
Used disk quota 8051 Kbytes, available 2188 Kbytes
226 Transfer complete.
ftp> 10 bytes received in 0.00Seconds 10000.00Kbytes/sec.
ftp> disconnect
221 Goodbye!
ftp> quit

```

Fig 8.50

FTP session using ftp.exe, a command line FTP client supplied with Microsoft Windows.

A similar transfer process can be accomplished from within Windows Explorer (see Fig 8.51). The login procedure involves entering ftp:// to specify the ftp protocol, the user name, an @ symbol and the server address into the address bar. Explorer then requests the password and makes the ftp connection. Files on the server can then be manipulated in the same manner as files on a local drive.



GROUP TASK Discussion

Describe the sequence of events occurring during the FTP session in Fig 8.50.



GROUP TASK Activity

Use ftp.exe to perform an FTP upload or FTP download. Perform exactly the same processes using Windows Explorer. Describe how Explorer implements each of the ftp.exe commands you used.

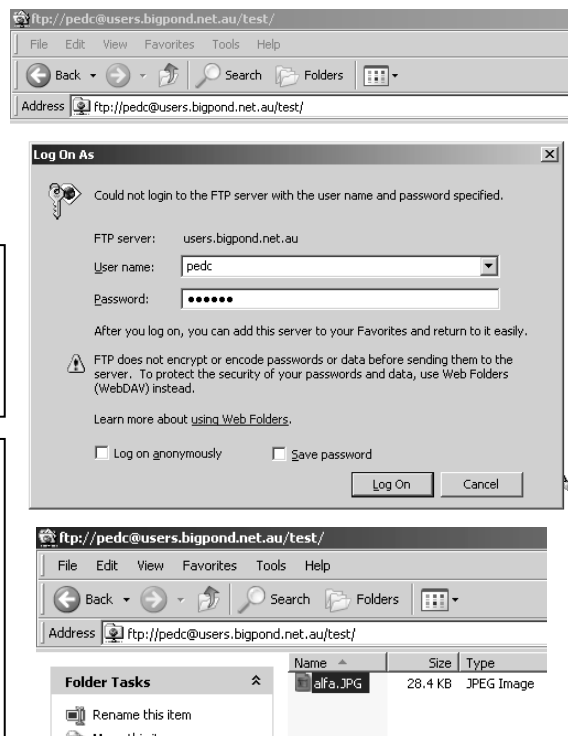


Fig 8.51

FTP session using Windows Explorer.

Electronic Mail Applications

No doubt you have all sent email messages. You simply enter the recipients email address, a subject line and your actual message. Click the send button and off it goes. Receiving is even simpler, just click the send and receive button and your emails arrive in your inbox. How does all this work? In this section we aim to answer this question.

Electronic mail, or email, uses two different protocols SMTP and either POP or IMAP. Email client applications, such as Microsoft Outlook, must be able to communicate using these protocols. SMTP (Simple Mail Transfer Protocol) is used to send email messages from an email client application to an SMTP server. Emails are received by an email client application from a POP (Post Office Protocol) server or IMAP (Internet Message Access Protocol) server. Fig 8.52 shows these server settings for a particular email account within Microsoft Outlook.

Sending an email using the account in Fig 8.52 involves the email client, in this case Microsoft Outlook, establishing an SMTP connection to the SMTP server called smtp.mydomain.com.au. The email is then transferred to this server. If the user wishes to download their email then Microsoft Outlook establishes a POP connection with pop.mydomain.com.au, logs into the server using the account name and



Fig 8.52

Emails are received from a POP server and transmitted to an SMTP server.

password, and finally receives all messages stored in the mailbox for that account. Note that the account name is the first part of the user's email address. If the email address is sam.davis@mydomain.com.au, then sam.davis is the account name and it is also the mailbox name on the POP server, mydomain.com.au is the domain name of the email server.

So how does email arrive into the mailbox on the POP, or IMAP, server of the recipient? The sender's SMTP server establishes an SMTP connection with the recipient's SMTP server. To do this it first needs to determine the IP address of the recipient's SMTP server. It does this by performing a DNS lookup. DNS stands for domain name server, these are servers that map domain names to IP addresses. For example, the email address fred@nerk.com.au includes the username fred and the domain name nerker.com.au. A DNS lookup determines the IP address of the email server that stores all mail for the domain nerker.com.au. The email message is sent over the Internet to the machine with this IP address. Once the message has been sent to the recipient's SMTP server it is passed to the corresponding POP, or IMAP server. This server places the message into the mailbox of the recipient ready for collection.

SMTP, POP, IMAP and DNS are protocols operating at the application and presentation layers of the OSI model. SMTP, POP and IMAP servers are part of software applications running on both email clients and email servers. It is possible, and highly likely, that a single machine is an SMTP, POP and IMAP server. In fact many email server applications include all three of these protocols within a single application. DNS servers are usually separate entities to email servers, they provide DNS lookup services to many other Internet applications, not just to email servers.



Consider the following:

The systems flowchart in Fig 8.53 describes the sequence of events occurring as email messages are transmitted and received. Notice that each email server includes an SMTP server and also a POP (or IMAP) server. Messages destined for a particular protocol are sent to a unique TCP/IP port. SMTP servers communicate on port 25, POP servers on port 110 and IMAP servers on port 143.

Most SMTP servers do not require a user name and password, hence anybody in the world can transmit email messages using almost any SMTP server on the Internet, however some SMTP servers will only deliver mail to or from their own customers. It is possible to perform such a transmission using a simple Telnet program (Telnet is yet another protocol used on the Internet). Microsoft Windows includes a program called telnet.exe. A typical SMTP session using telnet.exe is reproduced below in Fig 8.54. This session was initiated by typing the command telnet mail-hub.bigpond.net.au 25 at the Windows XP run dialog, this command executes the telnet program and establishes a connection with the remote SMTP server on port 25. All lines

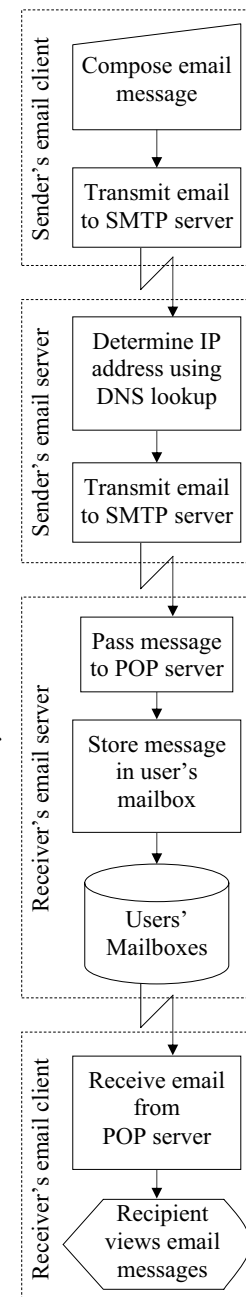


Fig 8.53
Systems flowchart
describing email
transmission.

preceded with a number are responses from the server, all other lines were entered at the keyboard. The interactions detailed in this SMTP session are identical to those performed automatically by email clients when sending email.

```
Telnet mail-hub.bigpond.net.au
220 bigpond.net.au service ready <identifier 8/46123778>
helo it's me
250 bigpond.net.au
mail from:me@here.com.au
250 ok
rcpt to:info@pedc.com.au
250 ok
data
354 start mail input; end with <CRLF>.<CRLF>
from: me@here.com.au
to: info@pedc.com.au
subject: text message
This is just a test message sent using SMTP and Telnet.
.
250 ok
quit
```

Fig 8.54

Sending an email directly using Telnet.



GROUP TASK Discussion

Work through the systems flowchart in Fig 8.53 describing the processing occurring at each step. Identify areas where the procedure can halt for a period of time.



GROUP TASK Practical Activity

Send an email message to yourself using a telnet program such as telnet.exe. Confirm the email has been sent correctly by receiving the email using your usual email client application.

Streaming Media Players and Streaming Servers

Audio and video playback, or display, requires a continuous and steady supply of data. If the data stops or its speed of delivery changes then the sound or video will not be displayed as intended. The purpose of streaming media players is to solve such problems without the need to download the entire file prior to playback commencing. Currently Adobe's Flash Player is the dominant streaming media player. Other popular players include Real Networks Real Player, Microsoft's Windows Media Player (see Fig 8.55), and Apple Computer's QuickTime Player. All these players can be embedded within web pages.

Streaming media players' use a buffer into which received data is placed. Data is received into one end of the buffer and removed from the other end for decompression and subsequent display. The aim of the buffer is to even out any inconsistencies in the rate of transmission. This buffering system works well as long as the video playback rate approximates the data transfer speed of the communication link. That is, data must enter the buffer at a higher or at least similar rate to the file's playback rate. If the



Fig 8.55

Windows Media Player is an example of a streaming media player.

playback rate is greater than the receive rate then playback must halt whilst sufficient data is received into the buffer. Current video formats and media players allow users to jump to any part of the video footage. This is accomplished by including the location of key frames within the initially transmitted data. The player can then request playback to commence at specific locations within a particular file.

The playback rate of audio and video files is determined when the file is first created. Playback rates for audio are determined by the compression method used, the size of each sound sample, and the number of samples per second. Similarly video playback rates are determined by the compression technique, the size of each frame and the number of frames per second. Files should be created so that their playback rate is less than the transmission rate of the intended communication link. For example, *Fig 8.55* shows Windows Media Player displaying a video with a playback rate of 81kbps. In this case the communication link must be able to support data transfer speeds in excess of 81kbps.

When the communication link is over the Internet then the intended users are often unknown, it is therefore impossible to predict the final speed of transfer when the file is being created. Furthermore different users will have different types of Internet connections and even similar connection types will operate at different speeds at different times. There are two techniques used to help overcome such issues, creating different versions of the file that play at different rates and using a streaming server to adjust the playback rate based on the real connection speed. The user's streaming media player doesn't care or even know which technique is being used, it merely accepts data into its buffer and plays it at the designated rate.

The first technique is the simplest as it does not require any special software at the server end. When this option is used websites commonly request users to indicate the type of Internet connection they have. Based on the user's selection a version of the file is transmitted that has a playback rate less than the normal expected speed of their connection. Should the connection speed deteriorate during the transmission then playback problems are likely to occur.

The second technique requires streaming server software to be installed and running on the remote machine. It also requires the file to be coded as a single multi-rate file, essentially various different playback rates are encoded into this single audio or video file. At playback time the streaming server detects the transmission rate and sends the stream that most closely matches this transfer rate. Should the transmission rate change during playback then the streaming server will alter the data sent to match the new rate. Currently Adobe's Flash Media Streaming Server software is the dominant server application. RealNetworks uses the term 'SureStream' to describe the process within their RealServer software and Microsoft uses the term 'Adaptive Rate Streaming' to describe the similar process performed by their Windows Media Server software.



GROUP TASK Practical Activity

Copy a video file that requires a fast data transfer rate onto one of the hard disks in your classroom. Using a streaming media player, have one person play the file and then have the whole class play this single file simultaneously. Clearly playback problems are likely to occur. Would installing a streaming server solve this problem? Discuss.



GROUP TASK Discussion

Why does audio and video data need to be delivered at a constant rate whereas text, numeric and image data can be transmitted at variable rates?

NON-COMPUTER TOOLS FOR TRANSMITTING AND RECEIVING

Tools, or more accurately systems, used to transmit and receive that do not directly use computers include:

- mail
- phone
- fax
- radio
- television

Each of the above systems includes various non-computer and computer based tools. Furthermore, each can be described using the communication concepts introduced at the start of this chapter. For example, radio could be described as a synchronous simplex transmission method. Synchronous because the signal is received in time with its transmission rate, and simplex because the signal travels in one direction only, namely from the radio transmitter to each radio receiver.



GROUP TASK Discussion

Classify each of the above systems in terms of the communication concepts described at the start of this chapter. Justify your responses.



GROUP TASK Discussion

Computer systems are able to perform functions similar to each of the above systems. Compare and contrast each of the above systems with a similar computer-based system.



HSC style question:

Outline the technology and processes occurring from when an email message is sent until it has been received.

Suggested Solution

The email client software connects to its SMTP mail server and then begins transmitting the message. During transmission the email client and the SMTP server communicate with each other to ensure error free transmission occurs. A variety of other lower level protocols would also be operating and ensuring correct delivery. When the complete message has been received by the mail server it examines the recipient's email address and determines the address of the recipient's mail server. The server then establishes a connection with this mail server (or perhaps some intermediate mail server). The two mail servers then use SMTP to transfer the message. Eventually the email arrives at the recipient's mail server. This server stores the message in the recipient's mail box. At some later time the recipient opens their email client software. The email client establishes a connection with the mail server using POP. Once established the emails in the users mailbox are transferred from the server to the email client. The email client stores the message in the user's local inbox where it can then be displayed for the user to read.

SET 8D

1. What is the relationship between ISO and OSI?
 (A) ISO developed the OSI model.
 (B) OSI developed the ISO model.
 (C) ISO develops standards, whilst OSI is a standard.
 (D) Both A and C.
2. FTP is used to:
 (A) view files on remote computers.
 (B) transfer files between computers.
 (C) stream audio and video files.
 (D) secure data during transmission.
3. The FTP command put this.txt would:
 (A) download the file this.txt
 (B) upload the file this.txt
 (C) delete the file this.txt
 (D) create the file this.txt
4. A video is designed for playback at 100kbps, is being transmitted. The receiving modem is connected at 50kbps. What will occur?
 (A) The video will display correctly.
 (B) The video will not play at all.
 (C) The video will play at half speed.
 (D) The video will start and stop during playback.
5. In relation to the transmission of email messages, which of the following is true?
 (A) SMTP is always used to transmit and POP is always used to receive.
 (B) SMTP is used for all transfers except from the email server to the email client.
 (C) POP or IMAP is always used to transmit and SMTP is always used to receive.
 (D) POP or IMAP is used for all transfers except from the email server to the email client.
6. A DNS server:
 (A) is usually part of email server software.
 (B) is used to determine MAC addresses.
 (C) exists solely to determine IP addresses for requesting email servers.
 (D) returns the IP address for a domain.
7. The function of the BIOS is to:
 (A) specify configuration settings for hardware devices on the motherboard.
 (B) provide a user interface to control the operation of hardware.
 (C) provide an interface between hardware devices and device drivers.
 (D) ensure operating systems work with all possible hardware configurations.
8. A normal telephone conversation can be best described as:
 (A) half-duplex and serial.
 (B) full-duplex and serial.
 (C) half-duplex and parallel.
 (D) full-duplex and parallel.
9. What is the purpose of the buffer within a streaming media player?
 (A) To even out data transfer speed inconsistencies.
 (B) To alter the playback speed to suit the real rate of data transfer.
 (C) To determine the real speed of data transfer.
 (D) To ensure a copy of all data is kept should the user wish to rewind the clip.
10. Email client software operates at which layers of the OSI model?
 (A) Layer 6 and 7
 (B) Layer 1 and 2
 (C) Layer 3 to 5
 (D) All layers.
11. (a) Identify the significant software used to transmit and receive a single email message.
 (b) All email attachments are converted to text prior to transmission. Describe why this conversion is needed.
12. Explain the purpose of the following protocols:
 (a) SMTP (b) POP (c) DNS (d) FTP
13. Briefly describe the function of each layer of the OSI model.
14. Discuss the relationship between the BIOS, device drivers and the operating system.
15. Video files can and are transferred over the Internet via FTP, as email attachments and using streaming media players. Compare and contrast these techniques and provide example scenarios where each technique would be suitable.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH TRANSMITTING AND RECEIVING

The widespread use of digital data and its ease of transmission, particularly over the Internet, has created a whole new set of social and ethical issues. Today it is simple for anybody to publish information. This certainly increases the ability of individuals to communicate their thoughts and ideas, however determining the accuracy of such information or its original source can be difficult.

Today many of us use email, newsgroups and instant messenger systems to communicate with people known to us and also to complete strangers. These systems are primarily text based, hence emotions, gestures and other human communication signals are difficult to communicate. A series of unwritten rules has evolved, they determine reasonable and acceptable communication and include techniques for humanising communication.

The Internet connects the world, this opens up incredible opportunities for even small businesses to market their products globally. However the world includes individuals from all walks of life. Most, we hope, are honest but clearly some are not. We therefore need to secure data during transfer to ensure messages sent arrive at their destination without having been read, copied or altered.

In this section we concentrate on social and ethical issues arising as a consequence of the widespread transmission of digital data.

QUALITY OF INFORMATION RECEIVED FROM THE INTERNET

When performing research it is vital to evaluate the quality of the information. Traditionally print media, such as books and journals, have been evaluated using five criteria, namely accuracy, authority, objectivity, currency and coverage. The use of such evaluation criteria is even more critical when using information received from the Internet.

A possible checklist that could be used to address each of these five traditional criteria follows:

1. Accuracy

Is the information well written and edited?

Have sources upon which the information is based been acknowledged?

2. Authority

Who wrote or is responsible for the information?

Are the author's qualifications clearly stated?

Is a phone number and address for the author or their company included?

3. Objectivity

Is the information free of advertising?

Is the information trying to alter or sway your opinion?

On commercial sites, is the information biased towards the company's products?

4. Currency

Is the information up-to-date?

Is it clear when the information was published?

5. Coverage

Is the information complete or is it still under construction?

What topics are covered and are they explored in depth?

Is this the entire work or is there a more detailed version?

**GROUP TASK Research**

Use a search engine to research a topic of personal interest. Choose three websites on this topic and evaluate the quality of the information presented using the above checklist.

SECURITY OF DATA BEING TRANSFERRED

Data transferred over the Internet travels through a generally unknown series of servers. In most cases data is not encrypted and consequently can be viewed, and even altered, by anyone with access to a machine within the communication link. Such issues are of particular concern in regard to email messages. Emails are always text messages and there is no security during their transmission. Even attachments are converted to text prior to transmission. Encryption can be added by users or by ISPs, however this is rarely the case.

Some possible security problems during transmission of email messages include:

1. **Eavesdropping:** It is a fairly simple matter for someone with access to an email server to read and copy messages without your knowledge. This is just like somebody listening into a conversation from an adjoining table.
2. **Identity theft:** Anybody can send an email message using any email address; recall our example SMTP session on page 333. This means it is possible to pretend to be somebody else, essentially stealing their identity. Many viruses use this facility, they send you emails that appear to be from a trusted contact, in reality they contain a copy of the virus. It is even possible to intercept passwords as they are transmitted to POP servers. This means the thief can not only send emails using your identity but they can also read your received messages.
3. **Message modification:** Those who administer SMTP servers are able to view and also modify email messages. The person receiving the message has no idea that this has occurred. Encrypting messages solves this problem however it doesn't stop the message from being deleted completely.
4. **Backups:** Most email servers are backed up regularly. This means a copy of your emails could be kept for many years. These copies can be read even after you think all copies have been deleted.
5. **Proof of delivery:** In general, email messages do reach their intended destination, however there is no full proof method for ensuring delivery. This means recipients can deny receiving a message. This has significant implications when the message contains important information, such as legal contracts or financial transactions.

**GROUP TASK Discussion**

We discussed encryption and decryption techniques in Chapter 6 (p226). Could encryption and decryption solve each of the above security problems? Justify your responses.

**GROUP TASK Research**

Many websites that transfer sensitive information use the Secure Sockets Layer (SSL) protocol. SSL uses digital certificates to ensure the authenticity of sender and receiver. Research and briefly describe the operation of the SSL protocol and how it uses digital certificates.

NET-ETIQUETTE

Net-etiquette is a term used to describe a code of behaviour that has evolved for polite communication on the Internet. It includes various symbols and techniques used to express emotion together with abbreviations used to reduce the number of keystrokes.

Following is a list of some items that are considered good net-etiquette:

- Never reply to rude or threatening messages, just ignore or delete them.
- The use of upper case is considered to be shouting. Don't use upper case, except to emphasise a specific word.
- Always use the subject field when composing emails. This helps the recipient to determine the nature of your message.
- Personal emails should not be forwarded to others unless the sender has given their permission.
- When using newsgroups, ensure your messages are relevant to the group. It's considered good practice to observe the dialogue for a period of time prior to posting.

Most Internet communication is text based, hence the emotions and gestures present during face-to-face communication are not possible. Emoticons and various abbreviations and acronyms can be used to express emotions and gestures. For example, :-) means smiling or agreeing, in fact many applications will automatically convert :-) to ☺. IMO means 'in my opinion' and LOL means 'laughs out loud'. Such acronyms save keystrokes and they also lend a more casual or conversational air to the exchange. Often asterisks or even brackets are used to surround actions, for example, ***Leaves the room***.



GROUP TASK Discussion

Create a list of emoticons and acronyms known and used by members of your IPT class. Discuss how the use of such items affects the nature of Internet conversations.

GLOBAL ISSUES – TIME ZONES, DATE FIELDS, EXCHANGE RATES, FOREIGN LANGUAGES

Although the Internet has brought the concept of the 'global village' closer, there are still many differences between countries. They exist in different time zones, use different date formats, different currencies and speak in different languages. These differences affect communication via the Internet in similar ways to traditional communication.

Theoretically, every 15° difference in longitude equates to a one-hour difference in time zone. In reality, this is not the case. For example, all of China uses a single time zone and many countries, including Australia, have implemented daylight saving. Communication between countries in real time must take account of these differences, in some cases workers may need to be present 24 hours a day.

The organisation of date fields can also cause problems when communicating globally. For example, does 02/03/04 mean 2nd of March 2004, 3rd of February 2004 or 4th of March 2002? The answer depends on the country in which you live. Back in Chapter 2 (p58) we described how dates are commonly represented using double precision floating-point and then displayed using settings specific to each computer. This solution solves many problems, however dates are often displayed on web pages and transmitted as text. In these cases it is safer to avoid any confusion by writing dates in words.

Different countries use different currencies, therefore financial transactions between countries involve exchanging one currency for another. This is not something that can be performed by individuals, rather they must engage the services of large financial institutions. Furthermore, financial institutions are free to set their own exchange rates. In effect both businesses and their foreign customers are at the mercy of the financial institution performing the currency exchange.

Language differences are more difficult to accommodate. Some websites use images of different countries flags to link to versions of their site written in different languages. Clearly it is not possible for web designers to create a version for every possible foreign language, however the use of icons or simple well recognisable words can greatly assist foreign language users.



GROUP TASK Practical Activity

Visit a number of foreign commercial websites that sell goods over the Internet. Identify how these sites deal with time zone, date format, currency exchange and language issues.

CHANGING NATURE OF WORK FOR PARTICIPANTS – WORKING FROM HOME OR TELECOMMUTING

Back in Chapter 1 (p22-25) we discussed various trends in regard to the changing nature of work. Many of these trends have occurred either directly or indirectly as a consequence of the widespread use of computer-based communication. However we did not discuss participants working from home or telecommuting. Part of Adam Turner's article 'Home Alone' (published in 'The Age' newspaper on 30th September 2003) is reproduced in *Fig 8.56*. The article discusses the hopes and realities of telecommuting.



GROUP TASK Discussion

Read the article in *Fig 8.56*. Identify within the article perceived and real advantages and disadvantages of telecommuting for employees.

THE IMPACT OF THE INTERNET ON BUSINESS

Today it is almost obligatory for all businesses to have a website and each employee to be assigned a business email address. The Internet has provided many opportunities for business to better communicate, reduce costs and improve their efficiency. Some specific areas where the Internet is able to improve business performance include:

- Sending and receiving business correspondence
- Keeping tabs on competitors
- Generating and processing sales
- Reducing advertising costs
- Reducing customer support calls
- Enhancing customer communications
- Providing product information to customers



GROUP TASK Discussion

For each of the above dot points, identify and describe real examples where businesses have improved their performance in each area as a consequence of their use of the Internet.

Home alone

By Adam Turner

30th September 2003

It is 30 years since the smog and chaos of a Los Angeles traffic jam inspired author-consultant Jack Nilles to coin the phrase "telecommuting". Stuck in peak-hour traffic, it is easy to dream about turning your back on the office and working from the comfort of home. Many of us spend our days in front of a computer screen anyway, so why not do it in our pyjamas and save the time, money and stress of the commute?

Spurred by the oil crisis of the 1970s and Los Angeles's deteriorating air quality, early telecommuting trials were undertaken in California by the Smart Valley Consortium, which included Pacific Bell, Deloitte & Touche, 3Com Corporation and Hewlett-Packard.

Telecommuting went on to become one of the great promises of the IT revolution, offering a daily commute from the bedroom to the study, with a quick detour to boil the kettle.

Wild predictions foresaw half the workforce telecommuting within 20 years but, as technology once again failed to deliver on the hype, the workerless office ended up in the "too hard" basket along with the paperless office.

Thirty years later, work is becoming more flexible, computers faster and network connections cheaper. Conditions for telecommuting would finally seem to be right - so why are most of us still fighting our way into the office each day?

Australia had a workforce of 9 million in June 2000, according to the Australian Bureau of Statistics' Labour Force Survey. Of the 8.6 million people at work during the week of the study, almost 7.5 million were employees, yet only 224,000 of those employees "mainly" worked at home.

Separate ABS research found that about 430,000 employees spent at least some of their working life at home in 2000 through "an agreement with their employer". A spike in 2001 saw this figure hit 555,000 but more than half of these gains were lost last year as the figure slumped to 480,000.

How many of these people could be classified as telecommuters is hard to say but less than half of them used a portable computer at home or accessed their employer's computer system.

So in this land of the laid-back lifestyle, why have fewer than 250,000 of Australia's 9 million workers embraced the dream of telecommuting?

Despite the stereotype of pyjama-clad slackers lazing around the house, working at home is no holiday, says Griffith University management lecturer Dr George Lafferty, who uses the term "teleworkers" to reach beyond those who work remotely to reduce commuting.

Teleworkers are in danger of becoming workaholics as they blur the line between work and home, says Lafferty, who was part of a three-year research project, commencing in 1999, on the adoption of telework in Australian organisations.

Lafferty and fellow researchers defined "regular teleworkers" as employees who consistently spent more than 40 per cent of their working hours away from the office, using telecommunications technology to access it.

Telework can include "flexi-place arrangements" with employees working on the road, at remote sites or on-site with clients.

Managers comprised the largest single group of teleworkers identified by the study, followed by IT professionals and administrative and clerical workers.

But while teleworkers often work harder, they find it harder to climb the corporate ladder because they are out of the office loop.

"If you're not visible, in many organisations you tend to be overlooked," Lafferty says.

"We have generally recommended that there should be a limit in how much teleworking people do because people really need to be in the organisational culture and politics. It's probably not the greatest thing to be away from the office for a long period of time."

The push for telework comes primarily from senior management looking for greater productivity and to give employees more flexible hours. Allowing employees to balance work and home life falls further down the list.

Lafferty says that teleworking arrangements should be preceded by a pilot study and require systematic rather than ad hoc arrangements.

More than half of the organisations surveyed employing teleworkers had no formal agreement on terms and conditions.

Lafferty's research found there was a lack of systematic analysis of industrial

relations, labour processes and regulatory frameworks for teleworking.

The possibility of organisations hiring teleworkers from low-wage economies and creating electronic sweatshops has left the union movement wary of telecommuting.

The Community and Public Sector Union negotiated a federal award for home-based working in 1994 but it has rarely been invoked, says CPSU spokesman Dermot Browne.

The lack of interest means teleworking is not a key issue for the CPSU, says Browne, but the union is "keeping an eye on" issues such as lifestyle balance and the career impact of working outside the office.

"When it was first introduced, everybody's assumption was that teleworking was going to be a lark but the reality is in fact people work harder," he says.

"More and more people are working from home but the sad reality is they're probably doing it at 10 o'clock at night or on Saturdays and Sundays and gently tipping over into unpaid overtime. There needs to be a proper balance for the employee."

The Australian Chamber of Commerce and Industry supports the establishment of more flexible working arrangements, provided they are agreed to by employers and employees.

"While there is some demand for telecommuting, most employers would prefer employees to have contact with the primary workplace and the business," says ACCI workplace policy director Peter Anderson.

Considering managers were the primary category of teleworkers in Lafferty's research, it is ironic he found the greatest resistance to teleworking came from managers who were not prepared to trust employees to work at home and whose own positions may become threatened if there is no apparent need for direct supervision.

This is compounded by the fact that office-based workers tend to take on urgent tasks that would otherwise be done by a teleworking colleague.

"There's suspicion from people who aren't teleworking that they're basically just people having a holiday," Lafferty says.

"In a lot of cases letting people work away from the office isn't treated as an entitlement, it's treated as doing people a favour."

Fig 8.56

Extract of an article by Adam Turner published in "The Age" newspaper on 30th September 2003

CURRENT AND FUTURE TRENDS IN DIGITAL COMMUNICATION

It is likely that in a few years time you will look back at the communication technology available today and comment on how simple and unintelligent it all was. Let us consider some current trends and try to predict what the future may hold.

Wireless networks – At the time of writing (2009) wireless LANs were popular within homes and educational institutions, and wireless high-speed Internet connections that use the 3G mobile network were common. These wireless Internet connections use high frequency waves but they do not require a direct line of site to the transceivers. Currently wireless solutions do not deliver the quality of service that wired technologies deliver, however one would expect this to improve in the future.

Integration of TV, radio and Internet – Currently TV stations broadcast analog and more recently digital transmissions. Digital transmissions are broadcast using radio waves and also coaxial cable. Similar digital radio networks also exist. A logical step is to combine the interactive nature of Internet communication with TV and radio. Already digital cable services are able to offer on demand movies, and limited interactivity for some broadcasts. In the future, it is likely that we will be able watch and navigate through TV shows in the same we do with DVDs and video cassettes.

Home automation – Currently home automation is in its infancy. Many homes include integrated sound and video systems, and a separate LAN enabling the sharing of peripherals and Internet connections. Some household appliances are available that can connect to the home's LAN. Automation systems are available for controlling lights, security systems and various appliances. In the future all these systems are likely to be standard integrated inclusions in all new homes. This could mean lights turn on as you enter a room, the volume on the stereo turns down when the phone rings, or a message appears on your TV when an email arrives.

Proactive devices – Currently computers react to input from users, they cannot sense the user's presence or predict their needs. In the future most devices will contain sensors of all types. Many sensors within various devices will provide input to powerful processors. These processors will use artificial intelligence to determine people's need. Hence computers will be proactive, that is they will respond appropriately without being asked or directed.

Seamless connectivity – Currently there is a wide variety of different network technologies that all communicate using different protocols. The Internet has resulted in the standardisation of many of these rules and technologies and mobile phone technologies have experienced similar levels of standardisation. Perhaps in the future such standardisation will continue to the point where devices of all types are able to communicate seamlessly. Each device could provide connections to other devices, hence a large network is formed without the need for separate and expensive infrastructures.



GROUP TASK Discussion

From a technological viewpoint all the above predictions sound exciting, however, on a human level, do you think they will really improve our lives? Discuss.



GROUP TASK Discussion

One thing is definite, should each of the above predictions come true there will be social and ethical consequences. Discuss possible social and ethical consequences for each of the above future predictions.



HSC style question:

- (a) Traditionally processing speed was improved by increasing CPU clock speeds and/or increasing the bus capacity. Recently speed increases have been achieved by packaging multiple CPUs within a single chip to implement parallel processing.
 - (i) Explain how different clock speeds affect processing speed.
 - (ii) Explain how different bus capacities affect processing speed.
 - (iii) Outline TWO situations where multiple CPUs on a chip would NOT increase processing speed.
- (b) In general data transmitted between nodes on a LAN is not modulated whilst data transmitted to and from the Internet is modulated – commonly using an ADSL or cable modem.
 - (i) Outline the modulation process.
 - (ii) Explain why modulated signals are used to transfer data over long distance such as to and from the Internet but are not used when transferring data over shorter distances such as over a LAN.

Suggested Solutions

- (a) (i) The clock speed of a CPU determines the speed at which instructions are executed. Slower clock speeds mean fewer instructions are performed per second, whilst faster clock speeds result in more instructions being executed per second.
- (ii) The bus capacity is the width or number of parallel connections between the CPU and other components on the motherboard. The number of connections determines the number of bits that can be moved into and out of the CPU in parallel. A larger bus capacity means more data is moved and processed simultaneously, whilst narrower bus sizes process less data at a time.
- (iii) Situations where multiple CPUs on a chip would not increase processing speed include:
 - Processes where instructions must execute in sequential order – for example, creating a running total within a spreadsheet.
 - Multiple processes where one process uses data altered by another process – for instance, changing a value in a field whilst summing values that include that same field value.
- (b) (i) Modulation is the process of encoding digital data onto analog waves. Different bit patterns are represented by altering the amplitude, frequency and/or phase of the analog wave.
- (ii) Modulation is used over long distances but not over short distances because:
 - Digital voltage changes used over shorter distances (such as LANs) would degrade over longer distances; hence modulated electromagnetic waves must be used for long distances such as Internet connections.
 - The binary high/low voltages used by LANs can be processed directly by digital computers. It is therefore simpler and also cheaper to use such signals over shorter distances.
 - The number of signal events per second that can be accurately detected is lower as distances increase. The effect of this is reduced by representing multiple bits within each signal event within modulated signals.

CHAPTER 8 REVIEW

1. Examples involving parallel transmission include:
 - (A) The system and PCI bus.
 - (B) SATA and USB interfaces.
 - (C) ADSL and cable connections.
 - (D) LAN connections.
2. Short data packets are used for asynchronous communication because:
 - (A) they can be used to represent individual characters.
 - (B) it means less time for the receiver to lose synch with the transmitter.
 - (C) asynchronous communication is used over short distances.
 - (D) only small amounts of data are ever transferred asynchronously.
3. Devices attached to the system bus achieve synchronisation using:
 - (A) Manchester or NRZI encoding.
 - (B) a short preamble.
 - (C) 'self-clocking' code.
 - (D) a common, but separate clock signal.
4. QAM is an example of a:
 - (A) modulation scheme that uses amplitude and phase changes.
 - (B) modulation scheme that uses frequency, amplitude and phase changes.
 - (C) system for encrypting and decrypting data.
 - (D) communication protocol.
5. Baud rate is equivalent to:
 - (A) bps
 - (B) bandwidth
 - (C) symbols/second
 - (D) All of the above.
6. Approximately how long would it take to transfer a 1MB file over a connection operating at 10Mbps?
 - (A) 1/10 second.
 - (B) 1/2 second.
 - (C) 4/5 second.
 - (D) 10 seconds.
7. The system bus connects:
 - (A) the CPU, main memory and I/O systems.
 - (B) the CPU to main memory.
 - (C) the components within the CPU.
 - (D) all other interfaces to their attached devices.
8. Which of the following is true for isochronous USB connections?
 - (A) They provide different rates of data transfer as the need arises.
 - (B) They communicate control messages to and from all USB devices.
 - (C) They are suitable for devices that require a constant rate of data transfer.
 - (D) Data packets are created at regular intervals, but not within every frame.
9. When a modem is transmitting it is:
 - (A) demodulating.
 - (B) modulating.
 - (C) encrypting.
 - (D) decrypting.
10. Discrete MultiTone (DMT) is used to:
 - (A) Swap channels when using ADSL.
 - (B) Modulate and demodulate ADSL signals.
 - (C) Simulate hundreds of dial-up modems operating in parallel.
 - (D) Remove interference from individual ADSL channels.
11. Define the following terms:

(a) serial	(d) half-duplex	(g) asynchronous
(b) parallel	(e) full-duplex	(h) bps
(c) simplex	(f) synchronous	(i) bandwidth
12. Explain the nature of the signals used by the following types of modem:
 - (a) dial-up modems
 - (b) ADSL modems
 - (c) cable modems
13. Identify and briefly describe the various hardware interfaces commonly existing between the CPU and an ISP when using an ADSL modem.
14. Various software tools work together during the transfer of data. Identify and briefly describe all the software involved during an FTP session.
15. In regard to email:
 - (a) Describe how email messages are transferred.
 - (b) Discuss issues in regard to the security of email messages.

In this chapter you will learn to:

- choose and justify the most appropriate method for displaying information given a particular set of circumstances
- describe the operation of display hardware
- use a range of hardware and software combinations to display different types of information
- format a text document with appropriate use of fonts, spacing and layout for printed and screen displays
- design and develop a simple web page
- generate reports for display within a database
- mail-merge information from a database into another application for display
- create audio, image and video displays with presentation software
- compare and contrast displays created without a computer to those created with a computer
- identify, discuss and appreciate the widespread use of non-computer methods of displaying information
- design a display for a wide variety of users

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams.

In this chapter you will learn about:

Displaying – the method by which information is output from the system to meet a purpose

Hardware for displaying

- screens (LCD, CRT and plasma screens) for displaying text, numbers, images and video
- printers and plotters for displaying text, numbers and images
- speakers for audio output
- digital projectors and interactive whiteboards for displaying text, numbers, images and video

Software for display

- interfaces for hardware display devices
- display features in applications packages, including:
 - reporting
 - formatting
 - spacing
 - merging
 - tables
 - charts

Non-computer tools:

- traditional methods for displaying the different types of data

Social and ethical issues associated with displaying

- communication skills of those presenting displays
- past, present and emerging trends in displays
- appropriate displays for a wide range of audiences, including:
 - standards for display for the visually impaired
 - displays suitable for young children

TOOLS FOR INFORMATION PROCESSES: DISPLAYING

The displaying information process outputs information from an information system to an external entity within the system's environment (see *Fig 9.1*). The external entity viewing the displayed information is generally one or more persons. The purpose of the information system is achieved via the information displayed to these people. For example, a search engine displays a list of websites matching the user's entered criteria. The purpose of the search engine is to suggest possible matches to the user. This purpose is achieved via the displayed list of websites.

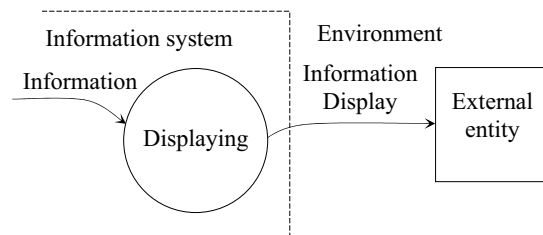


Fig 9.1

Displaying outputs information from the system to the environment.

Displaying is often a sub-process within other information process. For example, when collecting data online a monitor is used to display the data entry form, similarly progress bars are often displayed whilst intensive processing takes place. These are examples of information that is being displayed to inform and direct users. In fact, whenever a person receives information from a system a displaying information process has occurred.

We briefly discussed the displaying information process back in Chapter 2 (p50). We stated that the general meaning of displaying is to show, exhibit or put on view. This meaning encompasses sound and video as well as text, numeric and image data. Hence display hardware includes speakers as well as various types of screens and printers. In fact any device that performs actions based on information from a system is a display device. For example, the switches within traffic lights are display devices, they cause the lights to turn on or off based on information from the traffic light control system.

Software is used to interface with display hardware and it is also used to prepare information into a suitable form for display. For example, generating a sales report based on information from a database, or formatting a document within a word processor. Choosing the most effective method of display enhances the information, which means the system's purpose is more effectively achieved.

In this chapter we first discuss the operation of common display hardware. We then consider software in terms of its general features rather than the detail of specific applications. Finally we consider some non-computer display tools and some of the social and ethical issues associated with displays.



GROUP TASK Activity

Brainstorm a list of output devices. Categorise the list according to the different types of media each device is designed to display.



Consider the following:

The idea of displaying being a sub-process within other information processes seems to conflict with the notion that displaying outputs information from the system to an external entity within the system's environment. This conflict is resolved by considering each subsystem as a complete information system. For example, consider entering data via a keyboard as an information system. In this case a collecting information process is the primary information process occurring, yet a monitor is most likely being used for display. Indeed numerous other information processes are occurring between typing a character and it appearing on the monitor. The user entering the data is the sole external entity to this system. This person both enters data and also views the characters displayed as they type. The entered characters are data and the displayed characters are information. Displaying the characters typed confirms to the user that the collection process has occurred.



GROUP TASK Discussion

Identify and briefly describe the information processes taking place as a user presses a single key on the keyboard until the corresponding character is displayed on the monitor.



GROUP TASK Activity

Construct a dataflow diagram to model the flow of data and information through the processes identified in the above discussion.

HARDWARE FOR DISPLAYING

We cannot hope to describe the operation of all the different types of display devices currently available, therefore we restrict our discussion to the following devices:

Screens (or monitors) including:

- Video cards (display adapters)
- LCD (liquid crystal display) based monitors
- CRT (cathode ray tube) based monitors
- Plasma screens
- Projectors
- Interactive Whiteboards (IWBs)

Printers including:

- Laser printers
- Inkjet printers

Audio display including:

- Sound cards
- Speakers



GROUP TASK Discussion

The technology underpinning each of the above display devices is used within various different dedicated hardware devices. For example, photocopiers include similar technology to laser printers. Identify and describe examples of dedicated devices that use technologies present within each of the above display devices.

SCREENS

Information destined for the screen is received by the video system via the system bus. In most applications the video system retrieves this data directly from main memory without direct processing by the CPU. The video system is primarily composed of a video card (or display adapter) and the screen itself. The video card translates the data into a form that can be understood and displayed on the screen.

Video cards (display adapters)

A typical video card contains a powerful processor chip known as a GPU (Graphics Processing Unit), random access memory chips (often called Video RAM or VRAM) and various interfaces. Currently (2009) most video cards use at least 128MB of VRAM and some contain up to 4GB. When the video card is embedded as part of the motherboard it is common for some of the systems RAM to be used as VRAM. On most computers the functionality of a standard video card is embedded on the motherboard, whilst more powerful video cards, such as the one in *Fig 9.2*, are installed for intensive graphics applications such as video editing and high resolution gaming.

The video card in *Fig 9.2* communicates with the motherboard via a PCIe (PCI Express) port and transmits digital video data via its DVI (Digital Visual Interface) and HDMI (High Definition Multimedia Interconnect) interfaces. This particular video card also includes a TV tuner so it can be used to both collect and display video data. The PCIe interface has recently (2007) replaced the older AGP (Advanced Graphics Port); PCIe supports the high data transfer speeds required to move and process high definition and high frame rate video data.



Fig 9.2
ATI All-in-Wonder includes plugs into a PCIe slot and includes DVI and HDMI interfaces.

Digital computer monitors have largely replaced older analog screens. Currently most digital computer monitors use a DVI interface and most widescreen televisions include HDMI connections. HDMI interfaces can send and receive video and audio and also include the ability to control connected devices. For example, turning devices on and off, and altering contrast, brightness and volume settings. Older analog monitors were connected using VGA cables which included separate analog channels for red, green and blue, together with connections for vertical and horizontal synchronisation.



GROUP TASK Discussion

Many video cards contain large amounts of VRAM, whilst others utilise part of main memory (RAM). Discuss advantages and disadvantages of each of these approaches.



GROUP TASK Discussion

Many users of intensive graphics applications install more powerful video cards containing large amounts of VRAM. Identify applications where the purchase of such high performance video cards is justified.

LCD (liquid crystal display) based monitors

Flat panel displays, such as LCD based monitors, have largely replaced CRT based monitors. This has occurred for both computer monitors and television monitors. At the time of writing the most common flat panel technology for computers and television applications is based on liquid crystals. Gas plasma technologies are still used for larger televisions but their popularity is declining. In this section we consider the operation of LCD based monitors.

Liquid crystals have been used within display devices since the early 1970s. We see them used within digital watches, microwave ovens, telephones, printers, CD players and many other devices. Clearly the technology used to create the LCD panels within these devices is relatively simple compared to that contained within a full colour LCD monitor, however the basic principles are the same. Hence we first consider the operation of a simple single colour LCD panel and then extrapolate these principles to a full colour computer monitor.

So what are liquid crystals? They are substances in a state between liquid and solid, as a consequence they possess some of the properties of a liquid and some of the properties of a solid (or crystal). Each molecule within a liquid crystal is free to move like a liquid, however they remain in alignment to one another just like a solid (see *Fig 9.3*). In fact the liquid crystals used within liquid crystal displays (LCDs) arrange themselves in a regular and predictable manner in response to electrical currents.

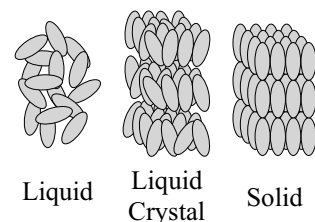


Fig 9.3

The molecules within liquid crystals are in a state between liquids and solids.

LCD based panels and monitors make use of the properties of liquid crystals to alter the polarity of light as it passes through the molecules. The liquid crystal substance is sandwiched between two polarizing panels. A polarizing panel only allows light to enter at a particular angle (or polarity). The two polarizing panels are positioned so their polarities are at right angles to each other. For light to pass through the entire sandwich requires the liquid crystals to alter the polarity of the light 90 degrees so it matches the polarity of the second polarizing panel. Each layer of liquid crystal molecules alters the polarizing angle slightly and uniformly, hence if the correct number of liquid crystal molecule layers are present then the light will pass through unheeded. This is the resting state of LCDs.

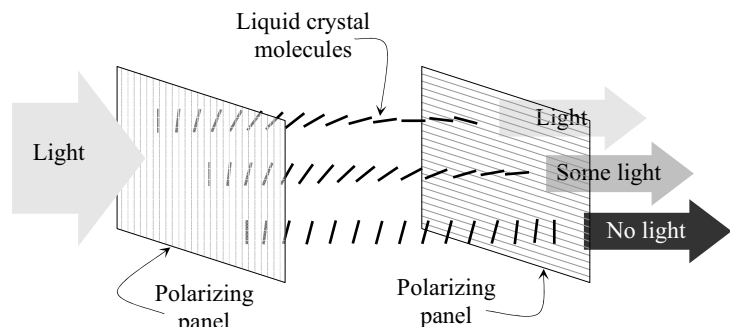


Fig 9.4

The primary components within a LCD.

To display an image requires that light be blocked at certain points. This is achieved by applying an electrical current that causes the liquid crystal molecules to adjust the polarity of the light so it does not match that of the second polarizing panel. Furthermore different electrical currents result in different alignments of the molecules and thus varying intensities of light pass through. In *Fig 9.4* the first sequence of molecules has no electrical current applied and hence most of the light passes through. A medium electrical current has been applied to the second sequence of molecules hence some light passes through. A larger current has been applied to the third molecule sequence and hence virtually no light passes through to the final display causing that pixel to appear dark.

In a CRT monitor, light is produced by glowing phosphors, therefore no separate light source is required. Within an LCD no light is produced, thus LCD based panels and monitors require a separate light source. For small LCD panels, such as those within microwave ovens and watches, the light within the environment is used. A mirror is installed behind the second polarizing panel, this mirror reflects light from the room back through the panel to your eye. LCD based computer monitors include fluorescent lights or a series of LEDs (Light Emitting Diodes) mounted behind the LCD, the light passes through the LCD to your eye. Such monitors are often called 'backlit LCDs'.

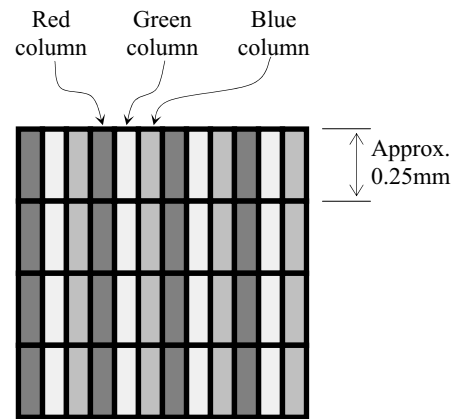


Fig 9.5
Section of the filter within a colour LCD based monitor.

So how are liquid crystals used to create full colour monitors? Each pixel is composed of a red, green and blue part. A filter containing columns of red, green and blue is contained between the polarizing panels (see Fig 9.5). A separate transistor controls the light allowed to pass through each of the three component colours in every pixel.

In current LCD screens transistors known as 'Thin Film Transistors' or TFTs are used, so for that reason LCD monitors were often known as TFT monitors. A two dimensional grid of connections supplies electrical current to the transistor located at the intersection of a particular column and row. The transistor activates a transparent electrode, which in turn causes electrical current to pass through the liquid crystals (see Fig 9.6). However, as each transistor is sent electrical current in turn, usually rows then columns, there is a delay between each transistor receiving current. To counteract this delay storage capacitors are used; each capacitor ensures the electrical current to its transparent electrode is maintained between each pixel refresh.

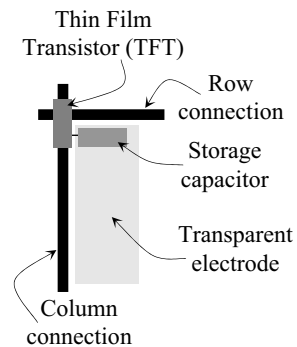


Fig 9.6
Components within each colour of each pixel in a TFT display.



GROUP TASK Discussion

LCD based computer monitors have almost completely replaced CRT based monitors. Why do you think this occurred? Discuss.



GROUP TASK Investigation

Resolutions less than the physical resolution of an LCD monitor mean part of the screen is not used. Is this true? Investigate and explain.

CRT (cathode ray tube) based monitors

Let us consider the components and operation of a typical cathode ray tube based monitor. The cathode is a device within the CRT that emits rays of electrons. Cathode is really just another name for a negative terminal. The cathode in a CRT is a heated filament that is similar to the filament in a light globe. The anode is a positive terminal; as a result electrons rush from the negative cathode to the positive anode. In reality, a series of anodes are used to focus the electron beam accurately and to accelerate the beam towards the screen at the opposite end of the glass vacuum tube. The flat screen at the end of the tube is coated with phosphor. When electrons hit the

phosphors they glow for a small amount of time. The glowing phosphors are what we see as the screen image.

To accurately draw an image on the screen requires very precise control of the electron beams. Most CRTs use magnetic steering coils wrapped around the outside of the vacuum tube. By varying the current to these coils the electron beams can be accurately aimed at specific phosphors on the screen. To further increase accuracy a shadow mask is used. This mask has a series of holes through which the electron beam penetrates and strikes the phosphors. There are various types of phosphors that give off different coloured light for different durations. In colour monitors there are groups of phosphors. Each group contains red, green and blue phosphors. When a red dot is required on the screen the red electron gun fires electrons at the red phosphors. To create a white dot all three guns fire. Firing the electrons at different intensities allows most monitors to display some 16.8 million different colours.

The entire screen is drawn at least 60 times each second; this is known as the refresh rate or frequency and is expressed in Hertz. Each refresh of the screen involves firing the red, green and blue electron beams at each picture element (pixel) on the screen. A screen with a resolution of 1280 by 1024 has approximately 1.3 million pixels to redraw 60 or more times every second. The electron guns fire in a raster pattern commencing with the top row of pixels and moving down one row at a time.

Most CRT monitors are *multisync*, meaning that they can automatically detect and respond to signals with various refresh, resolution and colour-depth settings. The software driver for the video card allows changes to be made to the refresh rate, resolution and colour-depth. Faster refresh rates, increases in resolution or increases in colour-depth require more memory and processing power. Often compromises need to be made between refresh rate, resolution and colour depth to maintain performance at a satisfactory level.

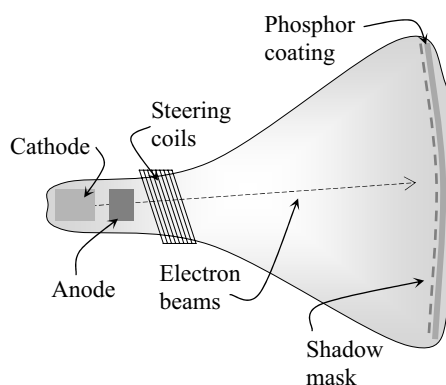


Fig 9.7

Detail of a Cathode Ray Tube (CRT).

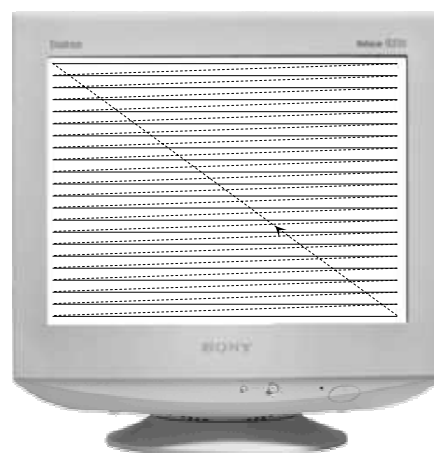


Fig 9.8

The screen is refreshed at least 60 times per second using a raster scan.

Colour Depth (Bits per pixel)	Number of colours
1	2 (monochrome)
2	4 (CGA)
4	16 (EGA)
8	256 (VGA)
16	65,536 (High colour)
24	16,777,216 (True colour)

Fig 9.9

Colour depth table showing number of bits required per pixel.



GROUP TASK Investigation

Examine the different settings available for the video card and monitor on either your school or home computer. Observe the effect of altering these settings. Which settings were the most satisfactory?



Consider the following:

The controllers within most monitors (including both LCD and CRT based monitors) are able to generate 256 different levels of electrical current corresponding to each 8-bit binary colour value received from the video card. Consequently 256 levels of light intensity are possible for each colour within each pixel. As there are three colours within each pixel there are $256 \times 256 \times 256$ or 16777216 different possible colours. Furthermore, current TFT based LCD monitors have a physical resolution of at least $1024 \times 768 = 786432$ pixels, as there are 3 transistors per pixel then these screens contain some $786432 \times 3 \approx 2.3$ million transistors. Each of these transistors is refreshed approximately 70 times per second, this means $2.3 \text{ million} \times 70$ or approximately 161 million transistors are being refreshed each and every second!



GROUP TASK Discussion

TFT based monitors include capacitors that maintain the electrical current in each pixel between screen refreshes. How is the screen image maintained between refreshes within CRT based monitors? Discuss.



GROUP TASK Activity

Dots per inch (dpi) and also dot pitch (width of each pixel in mm) are common measures of screen definition or crispness. If a screen is 12 inches (305mm) wide and has a resolution of 1024×768 pixels, calculate its dpi and dot pitch.

Plasma Screens

Plasma screens are common within large televisions. Plasma screens, like LCD screens can also be used as computer monitors and also for large advertising displays. In general, LCD screens dominate the computer monitor market, whilst LCD and plasma screens compete in the large wide-screen television market.

A plasma is a state of matter known as an ionised gas. It possesses many of the characteristics of a gas, however technically plasma is a separate state of matter. When a solid is heated sufficiently it turns to a liquid, similarly liquids when heated turn into a gas. Now, when gases are heated sufficiently they form plasma; a fourth state of matter. Plasma is formed as atoms within the gas become excited by the extra heat energy and start to lose electrons. In gases, liquids and solids each atom has a neutral charge, but in a plasma some atoms have lost negatively charged electrons, hence these atoms are positively charged. Therefore plasma contains free-floating electrons, positively charged atoms (ions) and also neutral atoms that haven't lost any electrons. The sun is essentially an enormous ball of plasma and lightning is an enormous electrical discharge that creates a jagged line of plasma – in both cases light (photons) is released. Photons are released as all the negative electrons and positive ions charge around bumping into the neutral atoms – each collision causes a photon to be released. In summary, when an electrical charge is applied to a plasma substance it gives off light. Within a plasma screen the gas is a mix of neon and xenon. When an electrical charge is applied this gas forms plasma that gives off ultraviolet (UV) light. We can't see ultraviolet light, however phosphors (much like the ones in CRT screens) glow when excited by UV light. This is the underlying science, but how is this science implemented within plasma screens?

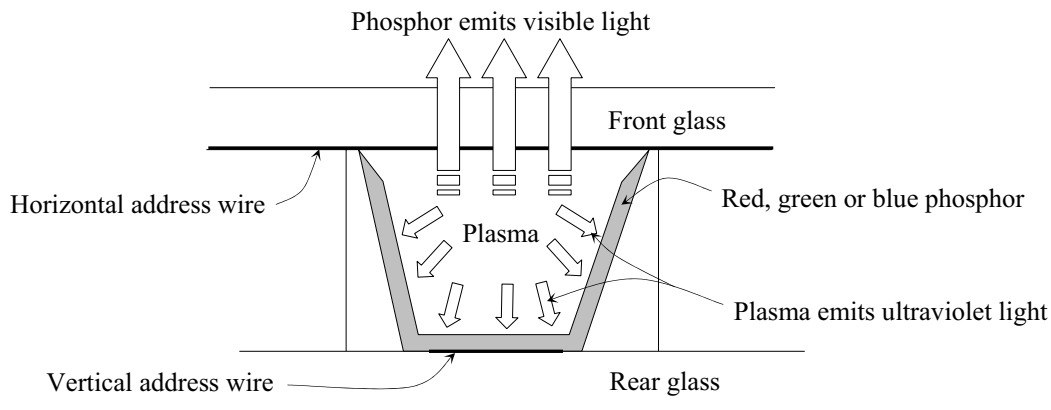


Fig 9.10
Detail of a cell within a plasma screen.

A plasma screen is composed of a two dimensional grid of cells sandwiched between sheets of glass. The grid includes alternating rows of red, green and blue cells – much like a colour LCD screen. Each set of red, green and blue cells forms a pixel. Each cell contains a small amount of neon/xenon gas and is coated in red, green or blue phosphors (refer Fig 9.10). Fine address wires run horizontally across the front of the grid of cells and vertically behind the grid. When a circuit is created between a cell's horizontal and vertical address wires electricity flows through the neon/xenon gas and plasma forms within the cell. The plasma emits ultraviolet light, which in turn causes the phosphors to glow and emit visible light. By altering the current passing through the cell the amount of visible light emitted can be altered to create different intensities of light. As with other technologies, the different intensities of red, green and blue light are merged by the human eye to create different colours.

Projectors

Projectors use a strong light source, usually a high power halogen globe, to project images onto a screen. In this section we consider the operation and technology used within such projectors. There are two basic projection systems; those that use transmissive projection and those that use reflective transmission. Transmissive projectors direct light through a smaller transparent image, whereas reflective projectors reflect light off a smaller image (see Fig 9.11). In both cases the final light is then directed through a focusing lens and then onto a large screen.

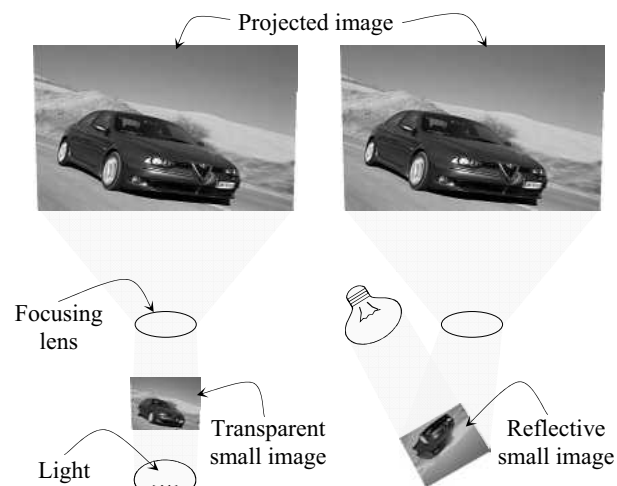


Fig 9.11
Transmissive (left) and reflective (right)
projector systems.

Older projector designs are primarily transmissive, the oldest operate similarly to CRTs. CRT based projectors have been largely phased out, and transmissive LCD projectors are marketed to low-end applications such as home theatre and other personal use systems. For high-end applications, such as conference rooms, board rooms and even cinemas, reflective technologies are predominant. Let us briefly consider three technologies used to generate the small reflective images within reflective projectors, namely liquid crystal on silicon (LCOS), digital micromirror devices (DMDs) and grating light valves (GLVs).

- **LCOS (Liquid Crystal on Silicon)**

Liquid crystal on silicon is essentially a traditional LCD where the transistors controlling each pixel are embedded within a silicon chip underneath the LCD. A mirror is included between the silicon chip and the LCD, hence light travels through the LCD and is reflected off the mirror and back through the LCD to the focusing lens. LCOS chips, such as the one shown in *Fig 9.12*, are also used in devices such as mobile phones and other devices where a small screen is required. For these applications the two polarizing panels are included as an integrated part of the LCOS chip. When used within projectors the polarizing panels are usually independent of the LCOS chips (see *Fig 9.13*). This means the light must only pass through each polarizing panel once on its journey to the screen. At the time of writing LCOS is a new technology and it appears likely to gain a large part of the projector market. Projectors for high quality digital cinema applications are under development that use a separate LCOS chip to generate each of the component colours.



Fig 9.12
LCOS chip suitable for use in a mobile phone or PDA.

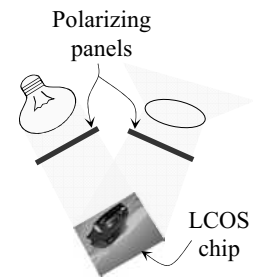


Fig 9.13
Most LCOS based projectors use two independent polarizing panels.



GROUP TASK Discussion

Brainstorm a list of possible applications where LCOS chips would be suitable.

- **DMD (Digital Micromirror Device)**

DMDs are examples of micro-electromechanical (MEM) devices. As the name suggests, DMDs are composed of minute mirrors where each mirror measures just 4 micrometres by 4 micrometres and are spaced approximately 1 micrometre apart. Each mirror physically tilts to either reflect light towards the focusing lens or away from the focusing lens. *Fig 9.14* shows just 16 mirrors of a DMD, in reality millions of individual mirrors are present on a single DMD chip (one mirror for each pixel). Each mirror is mounted on its own hinge and is controlled by its own pair of electrodes. DMD chips were developed by Dr. Larry Hornbeck at Texas Instruments and they are produced and marketed by their DLP™ Products Division. DLP is an abbreviation of “digital light processing”, hence DMD based projectors are often known as DLP projectors.

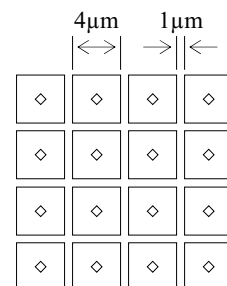


Fig 9.14
DMDs are composed of tilting mirrors.

To produce a full colour image current DMD projectors include a colour filter wheel between the light source and the DMD. This wheel alternates between red, green and blue filters in time with the tilting of the mirrors. To produce different intensities of light each mirror is held in its on position for varying amounts of time. The human eye is unable to detect such fast changes and hence a consistent image is seen. DMD based projectors currently produce better quality images due to their much larger percentage of reflective surface area compared to competing LCD based technologies. DMD manufacturers currently claim the reflective surface is approximately 89% of the chips surface area compared to LCD devices where the figure is less than 50% of the total surface area.



GROUP TASK Discussion

DMDs are an example of a MEM device. What do you think the term ‘Micro-electromechanical’ means? Discuss with reference to DMDs.

• GLV (Grating Light Valve)

GLVs were first developed at Stanford University and are currently produced by Silicon Light Machines, a company founded specifically to produce GLV technologies. GLVs are another example of a MEMs device. A single GLV element consists of six parallel ribbons coated with a reflective top layer (see Fig 9.15). Every second ribbon is an electrical conductor and the surface below the ribbon acts as the common electrode. Applying varying electrical voltages to a ribbon causes the ribbon to deflect towards the common electrode. Consequently, the light is altered such that it corresponds to the level of voltage applied.



Fig 9.15
A single GLV element.

The major advantage of GLVs is their superior response speed compared to other current technologies. Some GLV chips apparently have response times 1 million times faster than LCDs. This superior response speed allows GLV based projectors to use a single linear array or row of GLVs rather than a 2-dimensional array. For example, high definition TV has a resolution of 1920×1088 pixels, this resolution can be achieved using a single linear array of 1088 GLV elements, compare this to other technologies that require in excess of 2 million pixel elements. In reality current GLV projectors utilise a separate linear array of GLVs for the red, green and blue components of the image (see Fig 9.16). The light source for each GLV linear array being a similar linear array of lasers generating red, green and blue light respectively. The red, green and blue strips of light are combined using a light multiplexer. Finally a rotating mirror directs each strip of light to its precise location on the screen.

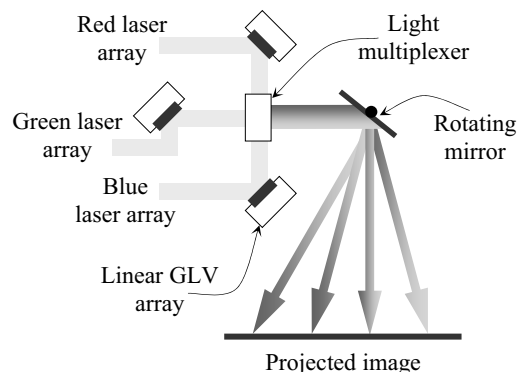


Fig 9.16
Major components of a GLV projector.



GROUP TASK Discussion

Discuss similarities and differences between computer monitors and projectors. Consider the signal received from the computer together with the operation of the device as part of your discussion.

Interactive Whiteboards (IWBs)

Interactive whiteboards (IWBs) are now common in many classrooms. An IWB system includes both collection and display devices. Commonly a projector is used as the display device; however IWBs are available for use with large LCD and plasma monitors. Users enter data, including location and control data, using touch. Many IWBs include pens designed specifically for the system, however in most cases a finger (or almost anything else) can be used for input. The pen or finger simulates the movement and clicks of a mouse. The IWB transmits inputs to the computer in the same way as a mouse; hence user inputs are processed by software applications and reflected on the screen as normal.

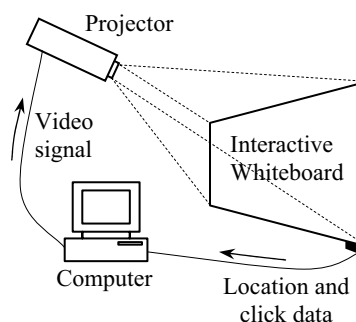


Fig 9.17
Typical IWB system.

Most IWB systems include a notebook style software application that allows an image of the screen to be captured and saved for later use. Using this software the user is able to draw or write on the surface much like drawing within a paint software application. The drawing is superimposed over the current screen image and both the screen image and drawing can be saved as an image for later use. To enter text most of these software applications include handwriting recognition so that hand written notes can be converted from image to text data. They also include an onscreen virtual keyboard, however for entering large amounts of text many users prefer to use a wireless keyboard.

There are a variety of technologies currently used to determine the location of each touch. Technologies include:

- Resistive membrane IWBs use two resistive membranes which cover the entire surface of the whiteboard. The outside membrane is separated from the inside membrane by a small air gap. Touching the surface causes the outer membrane to touch the inside membrane, which reduces the electrical resistance at that point and the coordinates are sent to the computer. As resistive IWBs are soft to the touch they are known as “soft boards”. A finger or any other object can be used as a pen. Most models allow use of standard dry-erase whiteboard pens.
- Electro-magnetic IWBs are made of hard material with a two-dimensional grid of wires behind. These boards require a special pen that includes a wire coil. The coil within the pen alters the electromagnetic waves within the board’s grid to determine the point of contact. Clearly these “hard boards” cannot be used with a finger or other object.
- Optical technologies are often used to convert a standard whiteboard or large LCD or plasma screen into an IWB. Some use infrared (IR) sensors together with pens that include an IR light in their tip. Others use optical sensors, much like those used within an optical mouse. The optical sensors are setup to cover the surface of the whiteboard or screen. Any object, such as a finger or pen, is detected and the location calculated and sent to the computer.



Consider the following

An inexpensive and remarkably effective IWB can be created using the remote (Wiimote) from a Wii gaming console combined with a pen containing a single infrared LED. The Wiimote includes Bluetooth connectivity and an infrared sensor. The infrared sensor detects the location of the infrared light emitted by the pen’s infrared LED. The location data is transmitted to the computer via the Bluetooth connection. Open source software is available for the Wiimote IWB system which includes functions similar to other commercial IWB software applications.



Fig 9.18

Pen with infrared LED and switch.
(Source: IR Great Innovations)



Fig 9.19

Wiimote mounted above a projector.
(Source: IR Great Innovations)



GROUP TASK Research

A variety of different applications have been devised that utilise the various sensors within the Wiimote. Research different applications of the Wiimote including its use within inexpensive IWB systems.

PRINTERS

Currently most printers receive their data via USB connections, however network printers often use Ethernet or wireless to connect directly to a LAN. Most current printers on the market are classified as either laser printers or inkjet printers. Specialised printers that use thermal technologies and impact dot matrix technologies are available. For example, most small receipt printers use thermal technology and many businesses use impact dot matrix printers to print documents in triplicate onto carbonised paper (examples of each are reproduced in Fig 9.20). In this section, we restrict our discussion to the operation of laser and inkjet printers.



Fig 9.20
Epson's TM-T88 thermal receipt printer
and FX-880 Impact dot matrix printer.



GROUP TASK Research

Use the Internet to research different types of printer technologies (not including laser and inkjet technologies). Print specific examples of printers that use each technology you find and describe where they are used.

Laser printers

Laser printers use static electricity to form images on paper. Static electricity is a charge built up on insulated materials in such a way that materials with opposing charges attract one another. Laser printers use static electricity to temporarily attract toner and then transfer it to paper. As no physical contact is used to form images laser printers are an example of non-impact printers.

Software applications send their output to the printer's software driver. The printer driver translates this data into a form that can be sent to the printer. The data is usually sent to the printer via a USB cable and is received by the printer controller within the laser printer. The printer controller is itself a dedicated computer containing significant amounts of RAM. Its job is to communicate with the host computer, format and prepare each page ready for printing and finally to create a rasterised image and send it progressively to the print engine.

So how does the print engine transform the information from the printer controller into hardcopy? The main component of the print engine is the photoreceptor. This is normally a rotating drum coated in a photo-sensitive material that is able to hold a static electrical charge. First the drum is given a positive charge by the charge corona wire. The drum then rotates past the laser-scanning unit. This unit traces out the image using a laser which discharges the static electricity on portions of the drum. The drum now holds the image as discharged areas (areas to be black) and positive charged areas (areas to be white).

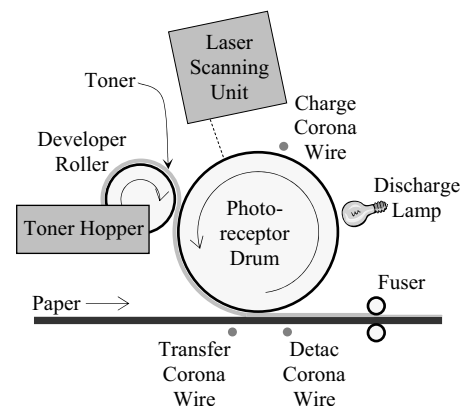


Fig 9.21
The main components of a laser printer.

The drum now rotates past the developer roller. The developer roller is coated in fine positively charged magnetic particles. As the developer roller passes through the toner hopper, these particles act like a brush, collecting a coating of positively charged toner.

The toner is attracted to the discharged areas of the drum and repelled by the positively charged areas. As a consequence the image areas on the drum are coated with toner.

The paper now approaches the drum, travelling at precisely the same speed as the drum. The transfer corona wire first negatively charges the paper, as a result the paper attracts the toner off the drum and onto the paper. The detach corona wire then discharges the negative charge held in the paper. This is necessary to stop it sticking to the photoreceptor or other sheets of paper. The fuser then fixes the toner to the paper. The fuser is essentially a pair of hot rollers, which melt the fine plastic toner particles into the fibres of the paper. The drum finally revolves past the discharge lamp, which removes all traces of the previous image.



GROUP TASK Investigation

Most laser printers contain replaceable toner and drum cartridges. Examine these cartridges and identify components from Fig 9.21.

Inkjet printers

Inkjet printers form images by depositing minute drops of ink onto the page. Within most current inkjet printers the diameter of each dot is approximately 20 to 60 micrometres. Full colour images are formed using the CMYK or four colour process system (we discussed CMYK back in Chapter 4, p147). This system requires dots of cyan, magenta, yellow and black to be deposited on the paper, hence most inkjet printers include cartridges containing ink in each of these colours. The Epson printer shown in Fig 9.22 includes a black ink cartridge and a cartridge containing cyan, magenta and yellow inks. The dots produced are too small for the human eye to detect, thus adjoining dots merge and we perceive a full colour image.

Inkjet technology is used within small point of sale printers right up to large commercial printers, Fig 9.23 shows a large commercial inkjet printer capable of printing on a variety of different materials up to 6 metres wide. Wide format inkjet printers have totally replaced the older plotters that were previously used for CAD and architectural applications.

So how do inkjet printers operate? There are two stepper motors, one advances the paper through the printer and the other moves the print head assembly left and right across the page. Most inkjet printers deliver a separate colour during each pass across the page. Once all colours have been printed the page is advanced slightly ready for the next strip of the image to be printed. The stepper motor and toothed belt that drives the print head (see Fig 9.24) actually moves a small precise

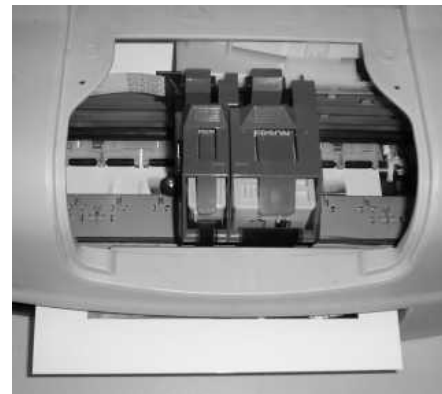


Fig 9.22

An inkjet printer showing the black ink cartridge alongside the cyan, magenta and yellow ink cartridge.



Fig 9.23

An inkjet printer capable of printing on various materials up to 6m wide.

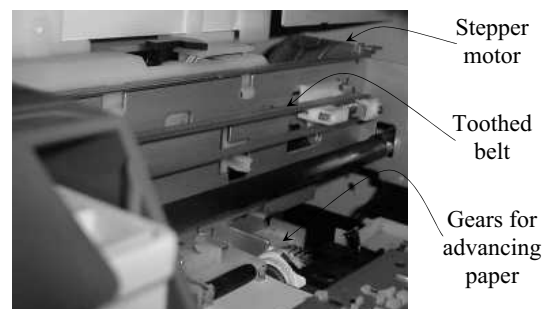


Fig 9.24

Detail of the inside of an inkjet printer.

amount and then stops for an instant whilst ink is deposited. This start-stop operation occurs so fast that it appears that the print head moves across the page at a continuous rate.



GROUP TASK Research

Within the text above, we noted that wide inkjet printers have totally replaced plotters. Research how plotters worked and why wide inkjet printers have completely replaced them.



GROUP TASK Activity

Create a list of steps that describes the processes occurring during the operation of an inkjet printer.

The print head within an inkjet printer contains the inkjet nozzles that form the individual droplets of ink together with the electronics required to operate the nozzles. Current printers contain more than 300 nozzles for each colour. There are two common technologies used to form the droplets, one based on heat and one based on the expansion of piezo crystals. Let us consider the operation of an individual nozzle based on each of these technologies.

Heat or thermal inkjet printers include a heating element within each nozzle (refer to Fig 9.25). When voltage is applied to the heating element the ink close to the element is heated to the point where it begins to vaporize. This vaporized ink forms a bubble within the nozzle chamber – this is why Canon uses the term ‘bubblejet’ to describe their thermal inkjet printers. The vaporised ink takes up more space and hence pressure increases and a droplet begins to form at the nozzle opening. A drop of ink is released once the pressure within the nozzle chamber is sufficient to overcome the surface tension at the nozzle opening. As the drop is released the heating element is switched off, this causes a pressure drop as the vaporized ink returns to its liquid state. The pressure drop causes ink from the adjoining reservoir to refill the nozzle chamber. This process is occurring thousands of times per second at each nozzle.

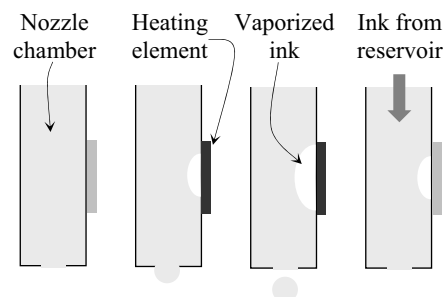


Fig 9.25
Operation of a thermal inkjet nozzle.

Piezo crystals expand and contract predictably as electrical current is increased or decreased. Essentially piezo crystals are able to transform electrical energy into mechanical energy due to vibration within the crystals. In the case of inkjet printers the mechanical energy is used to push ink out the nozzle chamber as microscopic droplets. When the electrical current is reduced or removed the piezo crystals contract. This contraction lowers the pressure within the nozzle chamber and causes ink from the adjoining reservoir to refill the nozzle chamber. Piezo based inkjet printers are able to produce a wide range of different sized droplets in response to different levels of electrical current. This is much more difficult to achieve with thermal systems. Also thermal systems must heat ink to high temperatures (thousands of degrees) and then quickly cool it down, for this

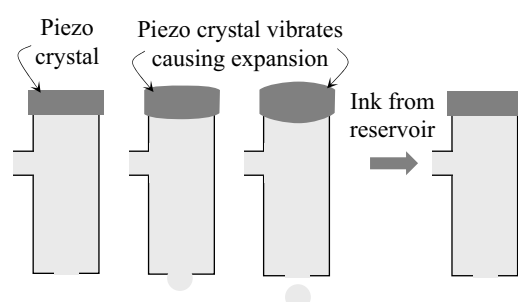


Fig 9.26
Operation of a piezoelectric inkjet nozzle.

reason, special inks are required that can withstand such extreme conditions. Piezo systems do not have such limitations and are therefore suited to a wider range of inks. Currently Epson's inkjet printers are based on piezo technology.



GROUP TASK Investigation

Take note of the inkjet printers around your home, school and local area. Research whether each of these printers uses thermal or piezo nozzles.



GROUP TASK Discussion

Some inkjet cartridges include the entire print head as part of the cartridge whilst others merely contain the ink reservoir. Compare and contrast these two approaches.

AUDIO DISPLAY

In Chapter 3 (p94-96) we discussed the operation of microphones and sound cards as collection devices. The components within speakers are similar to those found within microphones. In fact the processes occurring to display audio are essentially the reverse of the processes occurring during audio collection. Many older sound cards used many of their components for both sound collection and display. This meant that sound could either be collected or displayed but not at the same time; in essence these old sound cards operated in half duplex. Modern sound cards operate in full duplex, that is, they can collect and display audio simultaneously.



GROUP TASK Discussion

Identify applications where it is useful for sound to be both collected and displayed simultaneously.

Sound card

Most computers today include the functionality of a sound card embedded on the motherboard, however it is common to add more powerful capabilities through the addition of a separate sound card that attaches to the PCI bus via a PCI expansion slot. In either case similar components are used to perform the actual processing.

In regard to displaying the purpose of a sound card is to convert binary digital audio samples from the CPU into signals suitable for use by speakers and various other audio devices. Although many of today's audio devices include digital inputs ultimately an analog signal is required to generate sound through the system's speakers. Hence we restrict our discussion to the generation of analog audio signals. Analog audio signals are electromagnetic waves composed of alternating electrical currents of varying frequency and amplitude. The frequency determines the pitch and the amplitude determines the volume (we discussed this representation back in Chapter 2, p60). An alternating current is needed to drive the speakers, as we shall see later.

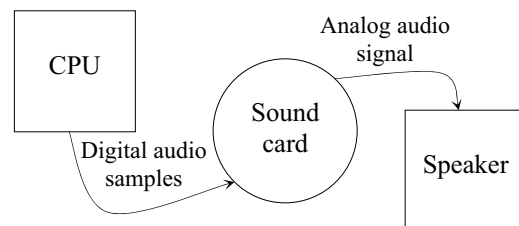


Fig 9.27
Context diagram for a sound card.

The sound card receives binary digital audio samples from the CPU via the PCI bus and transforms them into an analog audio signal suitable for driving a speaker. The context diagram in Fig 9.27 models this process. On the surface it would seem a simple digital to analog converter (DAC) could perform this conversion. In reality

audio data is time sensitive, meaning it must be displayed in real time, the DFD in Fig 9.28 describes this process. To achieve real time display sound cards contain their own RAM which is essentially a buffer between the received data and the card's digital signal processor (DSP). The DSP performs a variety of tasks including decompressing and smoothing the sound samples. The DSP then feeds the final individual samples in real time to a DAC. The DAC performs the final conversion of each sample into a continuous analog signal.

The analog signal produced by the sound card's DAC has insufficient power (both voltage and current) to drive speakers directly. This low power signal is usually output directly through a line out connector and a higher-powered or amplified signal is output via a speaker connector. Obviously the line out connector is used to connect display devices that include their own amplifiers, such as stereo and surround sound systems.

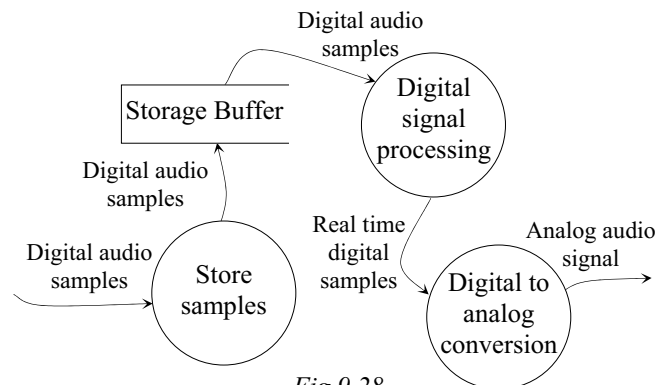


Fig 9.28
A sound card's display processes modelled dataflow diagram.



GROUP TASK Research

Many sound cards also contain a MIDI port, that often doubles as a joystick port. Research different types of audio display devices that connect to MIDI ports.

Speakers

Most speakers include similar components as dynamic microphones (refer p94). This includes an electromagnet, which is essentially a coil of wire surrounded by a magnet. As current is applied to the coil it moves in and out in response to the changing magnetic fields. As an alternating current is used to drive the speaker the coil vibrates in time with the fluctuations present within the alternating current. The coil is attached to a paper diaphragm, it is the diaphragm that compresses and decompresses the air forming the final sound waves. The coil and diaphragm are held in the correct position within the magnet using a paper support known as a 'suspension spider'.

The size of the diaphragm in combination with the coils range of movement determines the accuracy with which different frequencies can be reproduced. Large diameter diaphragms coupled with coils that are able to move in and out over a larger range are suited to low frequencies (0Hz to about 500Hz). Such speakers are commonly used within woofers. Smaller diameter diaphragms are tighter and hence respond more accurately to higher frequencies. Speakers with very small diameter diaphragms respond to just the higher frequencies and are known as tweeters. Commonly speaker systems include a separate low frequency woofer or sub-woofer, combined with a number of speakers capable of producing all but the lowest frequencies. Just a single large woofer is sufficient as low

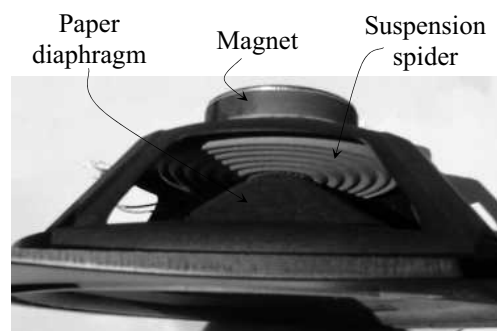


Fig 9.29
Underside of a typical speaker.

frequency sound waves are omnidirectional, that is they can be heard in all directions. Conversely, high frequency sounds from say 6000Hz up to 20000Hz are very directional, hence tweeters need to be arranged to produce sound in the direction of the listener.



GROUP TASK Practical Activity

Listen to the various sounds around you to determine their source. Is it easier to determine the direction of the source of higher or lower frequency sounds? How can you explain your results?



GROUP TASK Practical Activity

Most speakers can indeed be used as microphones, try connecting a speaker into the line in or microphone jack and see if it works as a microphone. The reverse experiment is not advisable; if a microphone is plugged into a speaker jack you're likely to destroy the microphone.



HSC style question:

In almost all cases the resolution of printed output is far greater than the resolution of screen output.

- (a) Define the term resolution.
- (b) A photograph when displayed on a screen appears clear, however when printed it appears jagged and of generally poor quality. Explain how this is possible when the resolution of printed output is apparently far greater than the resolution of screen output.

Suggested Solutions

- (a) Resolution is a measure of how detailed an image appears when displayed. Higher resolution images have more pixels within a given area than low resolution images. This means each pixel is smaller in a high resolution image than in a low resolution image. In terms of storage, images containing more rows and columns of pixels are said to be of a higher resolution than those with fewer total pixels. As a consequence the resolution of an image file is expressed in terms of the width in pixels by the height in pixels.
- (b) Possible explanations include:
 - Screens are low resolution display devices whilst printers are high resolution display devices. Screen pixels are larger and their edges are less well defined than printed pixels. The blurring of adjacent screen pixels makes the image appear clearer (or at least less jagged), whilst the definite edges to the printed pixels are more obvious.
 - Perhaps the printed version is physically much larger than the screen version. If the screen version shows all pixels within the image file then the printed version must contain larger versions of each pixel and these larger edges will appear jagged.

SET 9A

1. Which of the following is NOT a display device?
 - (A) printer
 - (B) monitor
 - (C) microphone
 - (D) speaker
2. In regard to HDMI, which of the following is true?
 - (A) HDMI is the name of the plugs used to connect audio visual equipment.
 - (B) Each colour is transmitted down a separate wire.
 - (C) HDMI transfers digital video, audio and control data.
 - (D) HDMI includes more than 128MB of VRAM and a GPU.
3. How is light produced by CRT monitors?
 - (A) Phosphors on the inside surface of the screen glow when struck a beam of electrons.
 - (B) Small florescent tubes behind the screen emit light which passes through the phosphor coating on the inside surface of the screen.
 - (C) Light is reflected off a mirror and through the front of the CRT.
 - (D) Each phosphor is controlled by its own transistor. The phosphors glow when their transistor is on.
4. 'Refresh rate' is best described as:
 - (A) the total number of colours that a screen can reproduce.
 - (B) the number of times per second that a screen image changes.
 - (C) the time taken to redraw the screen.
 - (D) the number of times the screen is redrawn each second.
5. Approximately how much memory is needed to hold a single 1024 by 768 pixel screen using a colour depth of 24-bits?
 - (A) 0.75MB
 - (B) 2.25MB
 - (C) 6MB
 - (D) 18MB
6. What are liquid crystals?
 - (A) Substances in a state between liquid and solid.
 - (B) Substances where the molecules move freely and randomly.
 - (C) Substances where the molecules are locked into alignment with each other.
 - (D) Substances that do not behave in a predictable manner.
7. What is the main purpose of the polarizing panels within an LCD?
 - (A) To form a sandwich between which the liquid crystal substance is contained.
 - (B) To ensure light enters and exits the sandwich at precisely the same angle.
 - (C) To alter the intensity of light passing through the display.
 - (D) To ensure light can only exit the LCD at right angles to the light entering the LCD.
8. What is the essential feature of all transmissive projectors?
 - (A) Light passes through a transparent image.
 - (B) Light is reflected off a small image.
 - (C) They are based on MEM devices.
 - (D) They require a high powered light source.
9. What is the function of lasers within laser printers?
 - (A) To fuse the toner particles into the paper.
 - (B) To apply an electrical charge to the toner.
 - (C) To charge the drum by tracing out the image.
 - (D) To discharge the drum by tracing out the image.
10. Droplets of ink are formed within thermal inkjet printer nozzles using:
 - (A) Piezo crystals which expand to increase the pressure within each nozzle.
 - (B) a heating element which thins the ink causing it to pass through the nozzle.
 - (C) a heating element which vaporises the ink. The vaporised ink passes through the nozzle.
 - (D) a heating element which vaporises some of the ink. The vaporised ink increases the pressure in the nozzle.

11. Each of the following is a component of one or more display devices. Identify the display device (or devices) that includes each component.

(a) VRAM	(g) neon and xenon gas
(b) cathode	(h) polarizing panel
(c) fuser	(i) laser
(d) piezo crystals	(j) electromagnet
(e) shadow mask	(k) suspension spider
(f) focusing lens	(l) storage capacitors
12. Explain how colour is produced by each of the following display devices.
 - (a) CRT based monitors.
 - (b) LCD based monitors.
 - (c) Plasma screens.
 - (d) Inkjet printers.
13. Compare and contrast the operation and physical characteristics of:
 - (a) CRT based monitors with LCD based monitors.
 - (b) transmissive projectors with reflective projectors.
 - (c) inkjet printers with laser printers.
14. Both DMDs and GLVs are MEM devices. What is a MEM device? Include a description of the operation of DMDs and GLVs as examples to justify your response.
15. Identify and describe the devices and processes occurring to:
 - (a) display a file containing sound samples through a speaker.
 - (b) display a bitmap file on a laser printer.
 - (c) display a bitmap file using a transmissive LCD projector.
 - (d) collect data using a resistive membrane IWB.

SOFTWARE FOR DISPLAYING

Software used by the displaying information process includes all the various display features present within virtually all software applications. For example, changing fonts, increasing line spacing, merging information from a database with a form letter, placing text in columns or tables, graphing or charting information, and so on. All these display features improve the presentation of information in preparation for display. The aim is to present the information in a form that best achieves the purpose of the information system.

In this section we first consider software that interfaces with display hardware and then examine common display features present in many software applications.

SOFTWARE THAT INTERFACES WITH DISPLAY HARDWARE

The software interface between software applications and display devices is essentially the same as the interface between collection devices and software applications (compare *Fig 9.30* to the similar diagram in Chapter 3 p104). The only significant difference being that data or information is moving from the software application to the display device.



Consider the following:

Device drivers for screens operate slightly differently from other drivers. Most display adaptors in use today include their own processors and RAM. Software applications, in consultation with the operating system, send instructions rather than final bitmaps to the video device driver. For example, when the text within a window is scrolled the application in charge of the window does not send a fresh bitmap of the entire window, rather it sends sufficient instructions to the driver to allow the video adaptor's processor to create the final screen bitmap. A typical screen display is likely to include windows and icons from a number of different software applications, hence the video device driver is receiving instructions from many software applications. The driver must pass these instructions to the display adapter, who in turn creates and transmits the final bitmap frames to the actual screen.



GROUP TASK Discussion

Would changes made to settings on the screen in *Fig 9.31* alter the data sent from software applications? Discuss.

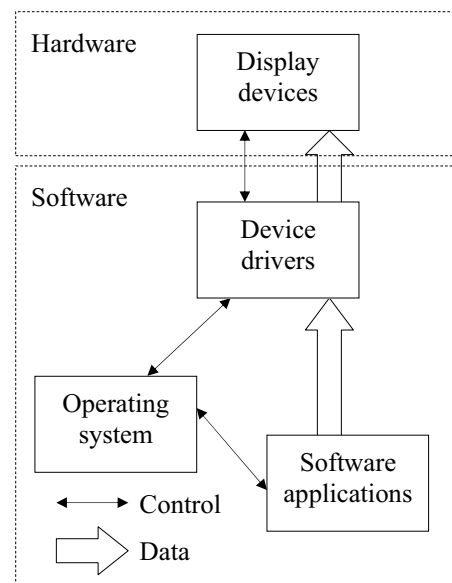


Fig 9.30

The software interface between display devices and software applications.

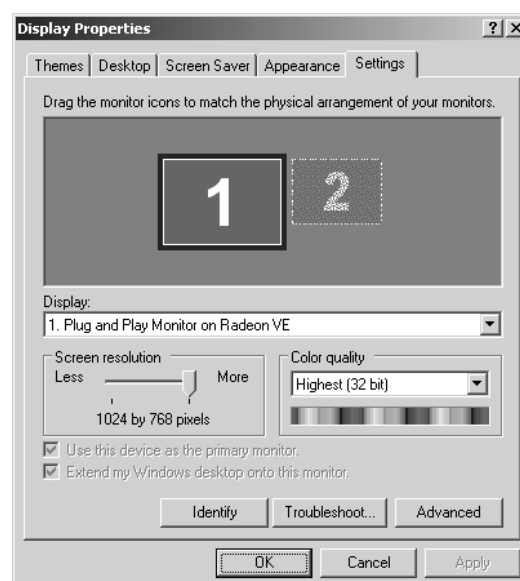


Fig 9.31

Display properties window within Microsoft Windows XP.



Consider the following:

Most inkjet printer drivers include various functions similar to those shown in *Fig 9.32*. Similar functions are not included as part of the drivers for laser printers.



GROUP TASK Discussion

Why are functions like those shown in *Fig 9.32* included for inkjet printers but similar functions are not included for laser printers. Discuss.



GROUP TASK Research

Research how inkjet nozzles are cleaned by an inkjet head cleaning utility.

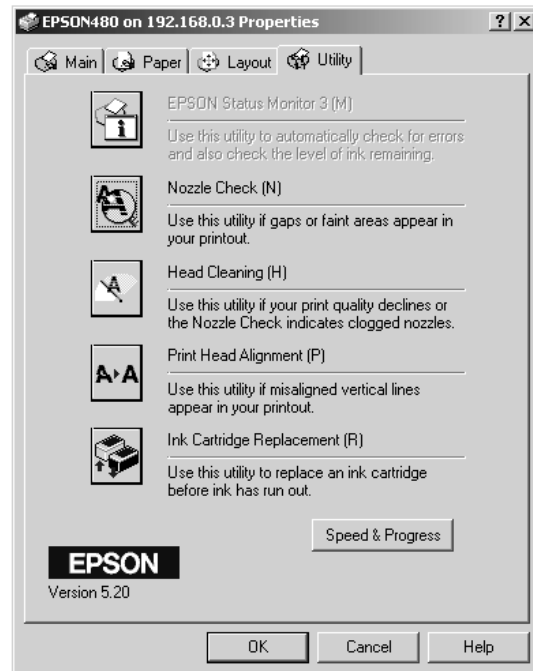


Fig 9.32

Utilities for an Epson Inkjet printer.



Consider the following:

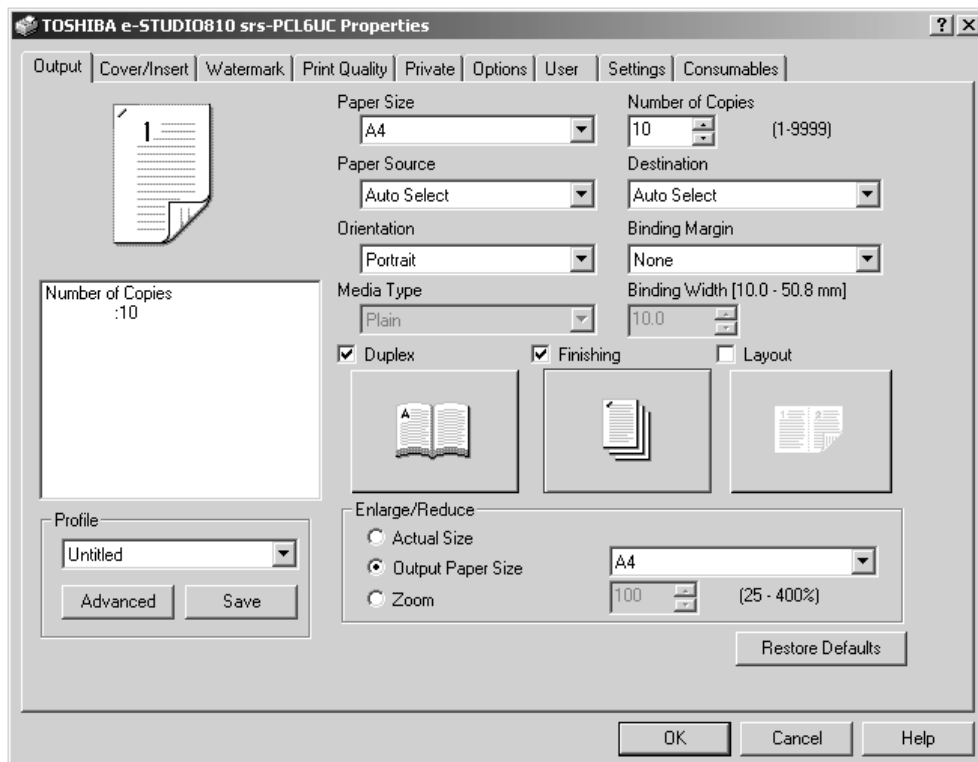


Fig 9.33

Properties dialogue for a Toshiba e-Studio 810 multifunction printer.



GROUP TASK Activity

List and describe printing features of the Toshiba e-Studio 810 implied by the various settings on the above screen.

DISPLAY FEATURES WITHIN APPLICATION SOFTWARE

Clearly we cannot examine all possible display features within all possible applications, so we restrict our discussion to some broad areas, namely:

- reporting within database applications.
- formatting within word processing applications.
- designing and developing simple web pages.
- combining information from different sources.

Each of these areas prepares information for final display. They do not alter the actual information, rather their purpose is to enhance information by presenting it in a form suited to the intended audience and the display devices on which it will be viewed.

Reporting within database applications

Reporting, in terms of database applications, refers to the processes used to present the output from a database. Reports can be designed for printing or they can be designed for screens. For example, a search engine retrieves a list of websites from its database, the information retrieved is then formatted using a report and the final result is displayed on the user's screen. Similarly a report is used to generate invoices for a business, the report formats the information in a manner suitable for printing.

In most cases one or more queries are used to retrieve the information from the database, the result of these queries becomes the data source for the report. The report specifies how the data will be displayed, therefore a report can be thought of as a template describing how the information will be presented.



Consider the following:

Microsoft Access is an example of a database management system that includes its own reports module. The design of a simple report is reproduced in *Fig 9.34* below.

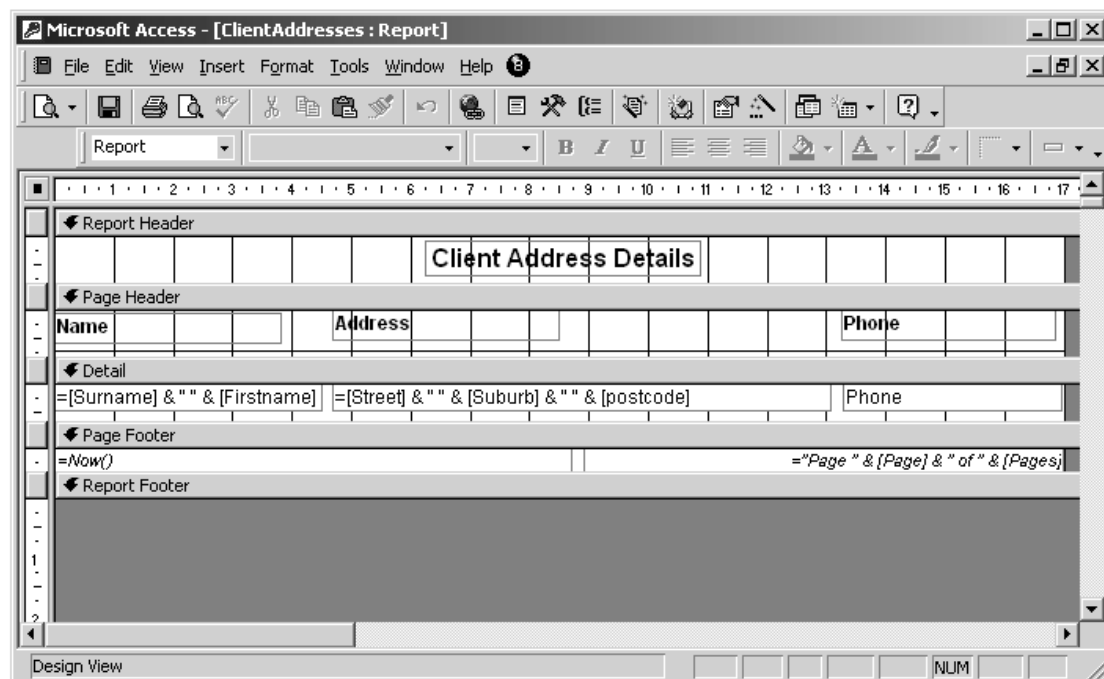


Fig 9.34
Report design window within Microsoft Access.

The report shown in *Fig 9.34* includes various sections, namely a report header, page header, detail and page footer section. On this particular report every record within the retrieved data causes a separate detail section to be produced during display. Similarly a separate page header and footer are generated for each new page.

Static data that is displayed on all reports is entered into controls called labels, whilst data that changes during generation of a report is specified using textboxes. Notice that a single textbox can combine multiple fields, for example one textbox within the detail section contains =[Surname] & " " & [FirstName]. Textboxes can also be used to generate information as the report is created, for example =Now() causes the current date and time to be displayed and [Page] displays the page number within the report.



GROUP TASK Discussion

Discuss the differences between each of the sections within the report shown in *Fig 9.34*.



GROUP TASK Activity

Create a report, similar to the one shown in *Fig 9.34*, using Microsoft Access or a similar database reporting application.

Formatting within word processing applications

Formatting is the process of specifying how the information will be presented. In relation to word processing it includes features for altering fonts, margins, paragraph indenting, line spacing, columns, borders and shading. In fact any feature designed to enhance the presentation of the information is an example of formatting.

In this section we briefly consider some of the more common formatting functions present in most word processors. Many of these features are included within a variety of software applications used to display text.

• Fonts

A font is a specific example of a particular typeface. For example, Times New Roman is a typeface, and Times New Roman Italic 12 point is an example of a font. The term 'font' is often used incorrectly when specifying a typeface. For example, all the settings within the dialogue shown in *Fig 9.35* combine to specify a particular font, yet within this dialogue the term font is used incorrectly as the label displayed above the list of available typefaces. Similar errors are so widespread that the term font is now used interchangeably with the term typeface.

Each font is composed of a particular typeface, style, weight and size. The style, weight and size alter the characteristics of the typeface. Commonly the size of a font is measured in points from the top of the highest ascender to the bottom of the lowest descender (see *Fig 9.36*). There are 72 points per inch, therefore a 12-point font is 12/72 of an inch high

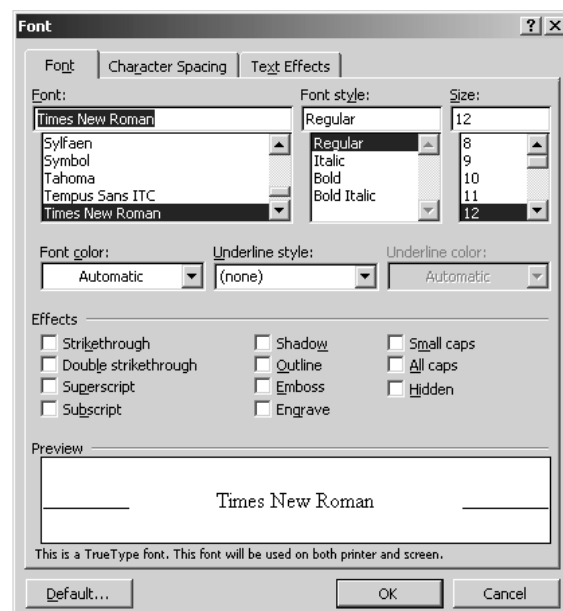


Fig 9.35
Font dialogue from Microsoft Word.

which is equivalent to approximately 4.2mm. Point size determines the height of a font when printed, however problems occur when point sizes are used for screen displays, in particular web pages. The resolution and the physical dimensions of the monitor will alter the actual size of the font when displayed.

Fonts are broadly classified as either serif or sans serif fonts. Serifs are the small strokes present at the extremities of each character. A sans serif typeface, such as Arial, does not have any serifs. Serif typefaces are generally used for the main body of printed

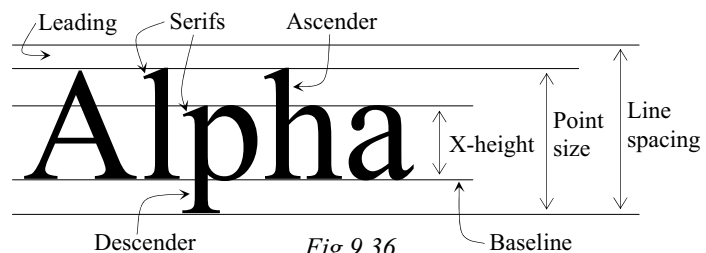


Fig 9.36
Attributes of fonts.

documents whereas sans serif typefaces are used for titles and most screen displays. It is said that the serifs assist the reader's eye to combine characters and consequently determine words rather than individual characters. On most screens the resolution is insufficient for serif fonts to be displayed accurately and hence their use is generally avoided. Titles normally use larger fonts and a small number of words, therefore the use of serif fonts is unnecessary.

- Spacing

Spacing refers to the distance between elements on the final display. It includes the space between images and text, the space between lines and paragraphs of text, the space between individual words and even the space between individual characters. It also includes margins, indenting and space between headers and footers and the main body text. Fig 9.37 below details the effect of margin settings available within the current version of Microsoft Word. Fig 9.38 details settings in regard to line spacing together with indenting and character spacing settings within Microsoft Word.

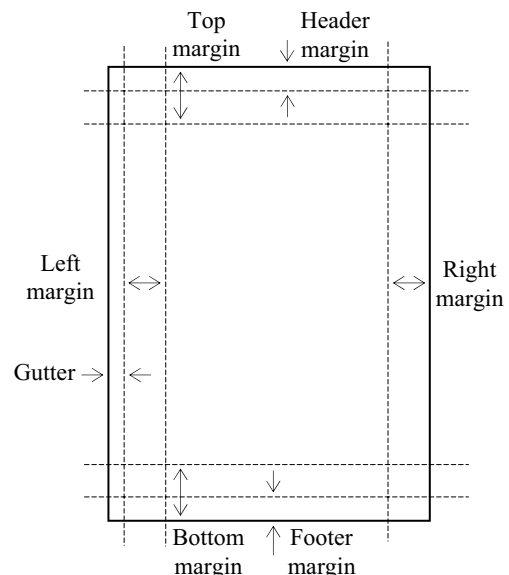
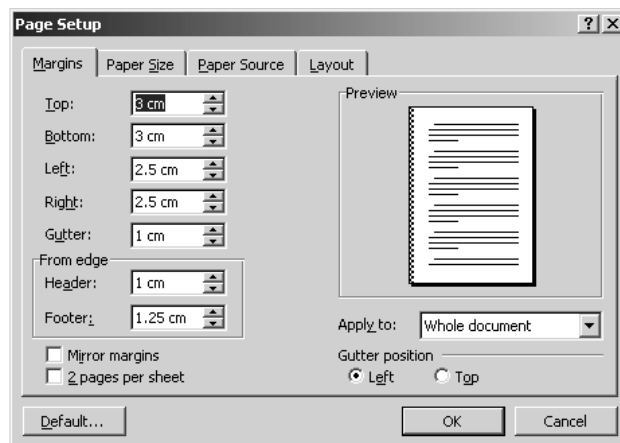


Fig 9.37

Page setup dialogue and margins within Microsoft Word.



GROUP TASK Activity

Investigate the margin settings within a word processor. Determine the relationship between top and header margins, bottom and footer margins, and also between the gutter and left/right margins.

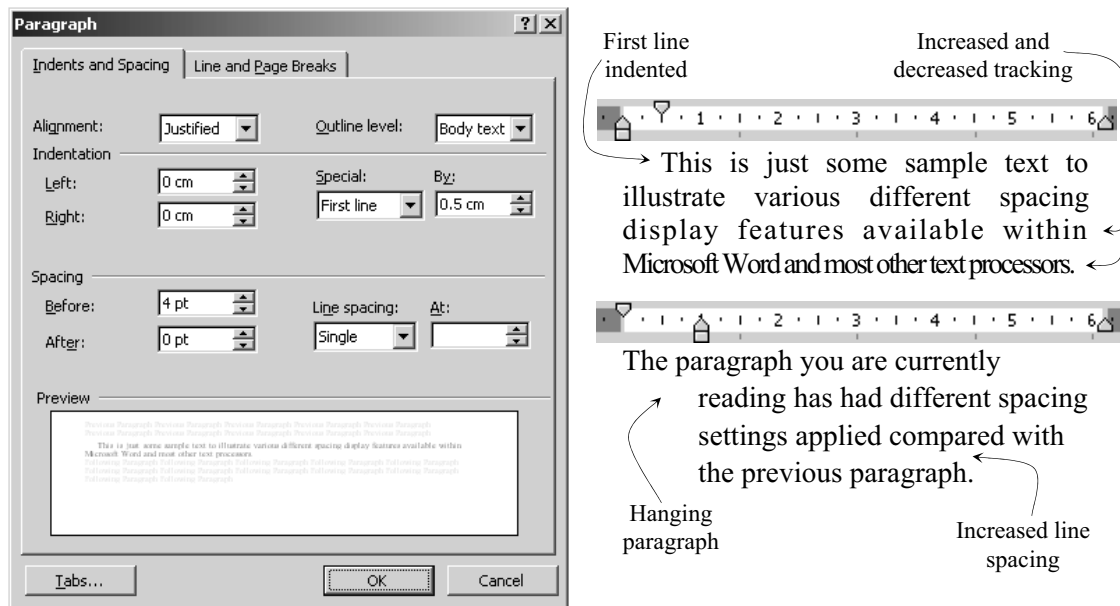


Fig 9.38
Common paragraph and character spacing display features.



GROUP TASK Discussion

The dialogue in Fig 9.38 displays the settings for the top paragraph of text. Describe alterations to these settings that would result in the spacing shown in the second paragraph.



Consider the following:

In Fig 9.36 not only the point size but also the X-height and line spacing are shown. Two fonts of identical point size appear to be of different sizes when their X-heights differ significantly. Similarly line spacing has a marked effect on readability. Too much line spacing can cause the reader to lose their place and too little makes it difficult to maintain the eye on the current line of text. In general the line spacing for body text should be approximately 120% of the font's point size. In most word processors this is the percentage used when single line spacing is specified. For example, 12-point line spacing is used for 10-point fonts.

The horizontal space between individual characters can also be altered. For example, Fig 9.38 shows examples of increasing and decreasing the tracking or character spacing. It is also possible to adjust the space between particular pairs of characters; this is known as kerning (see Fig 9.39). Kerning is generally only required for larger font sizes such as those used to format headings. Most current word processors contain basic automatic kerning features whilst desktop publishing applications provide much more precise control over kerning.

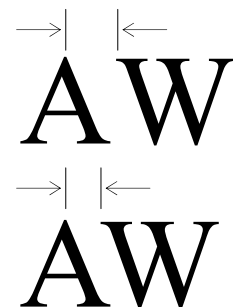


Fig 9.39
Kerning adjusts the space between individual character pairs.



GROUP TASK Activity

Investigate the effect of altering the line spacing, tracking and kerning when using various different typefaces at a variety of different point sizes. Comment on the relative differences between each typeface's X-height.

• Layout

All the text and image elements on a page or screen combine to communicate the information to the final user. Although the content of each element is obviously critical, the layout of the elements in relation to each other enhances the visual appeal of the display and also guides the user's eye through the information. It is therefore important to understand some of the basic principles of page layout.

Page layout is primarily the task of graphic designers, it is a creative process and as such there are no hard and fast rules. Clearly this course is not about graphic design, therefore we restrict our discussion to some of the basic guidelines worth considering when laying out page elements for display.

Balance and symmetry affects the formality of a document. The optical centre of a page is not in the physical centre, rather it is a position approximately three-eighths down from the top page. When elements are arranged symmetrically around the optical centre the display appears more formal than asymmetrical layouts. For example, in *Fig 9.40* the content is identical in both layouts, however the layout on the left appears more formal whilst the layout on the right has a more creative and informal feel. Each design achieves a slightly different purpose yet the information is identical.

Contrast between elements creates variety within the design. It can be used to add emphasise to some elements whilst reducing the impact of others. Contrast can be introduced by using different sized or weighted fonts, by altering the position or orientation of elements, or by using contrasting colours. Conversely overuse of contrast can be distracting – it is important that all elements are linked and work together in harmony. For example, in *Fig 9.40* the heading has been rotated in the right hand layout, this adds contrast, however the font used is the same typeface as the subheading hence these two headings remain linked.

Research indicates that most readers tend to scan a page using a Z pattern. Presumably such research was based on English speakers where reading occurs left to right and then scans down to the left to commence reading the next line. It makes sense to use this knowledge and position important design elements according to this Z pattern. Placing strong elements in other arrangements forces the reader to do a 'double take' to overcome their natural reading tendencies. For example, placing the strongest heading or image within the lower portion of the layout forces the reader's eye to that element against their natural tendencies. Furthermore, the reader must then scan the page again to locate the next most significant element.

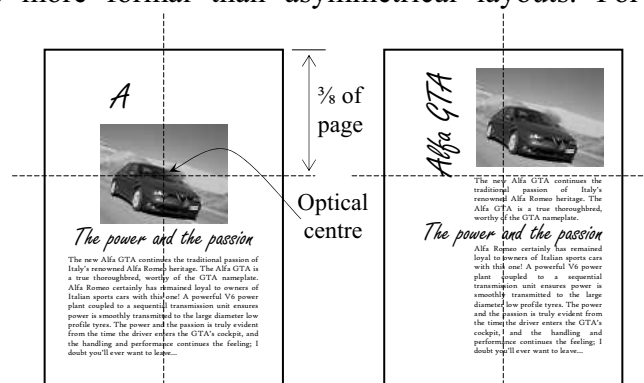


Fig 9.40

Balance and symmetry affects formality.



GROUP TASK Practical Activity

Create a single page advertisement for a product of your choosing. Your design is to include at least one heading, one image and one block of text. Swap your advertisement with one of your classmates and have them comment on the layout in terms of balance and symmetry, contrast and harmony and also the natural reading order of the elements.

Designing and developing simple web pages

Designing for the web introduces problems not present when designing for print. The capabilities of each end user's display are unknown, hence the design must be flexible so that it will display appropriately within a wide range of browsers and on a wide range of screens. HTML (Hypertext Markup Language) is the language of the web, its primary purpose being to overcome the differences between browsers and screens. In Chapter 4 we briefly introduced HTML in relation to creating hyperlinks, in this section we shall concentrate on two examples of common page layout elements used within HTML web pages, namely tables and frames. We restrict our discussion to a brief introduction to the HTML tags used to create these display elements.

• HTML Tables

Tables are used to split the display into distinct cells. Each cell can contain text, images, hyperlinks or any combination of HTML elements. HTML tables must be created within the body of an HTML document, that is between the `<body>` and `</body>` tags. The `<table>` tag begins a table, this is followed by specifying the detail of each horizontal row. Each new row commences with a `<tr>` tag followed by a sequence of `<th>` or `<td>` tags. A `<th>` tag specifies a table heading and a `<td>` tag indicates table data. Each of these tags is followed by the data to be displayed within the current cell. In general each tag should be concluded with its corresponding end tag, however in most browsers many end tags, such as the `</th>` and `</td>` tags, are optional.

```
<html>
<body>
<table border=1 cellpadding=10 cellspacing=20 width="100%" height="100%">
  <tr>
    <th colspan=2>This heading spans 2 columns</th>
    <th>Column 3</th>
    <th>Column 4</th>
  </tr>
  <tr>
    <th rowspan=3>This heading spans 3 rows</th>
    <td width="25%">A</td>
    <td width="25%">B</td>
    <td width="25%">C</td>
  </tr>
  <tr align="center">
    <td>D
    <td>E
    <td>F
  </tr>
  <tr>
    <td>G
    <td align="right">H
    <td>I
  </tr>
</table>
</body>
</html>
```

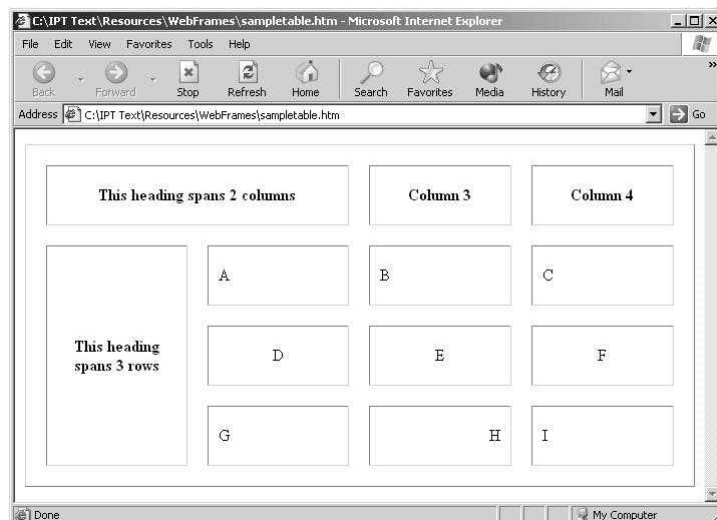


Fig 9.41
Sample HTML table code and the result displayed in Internet Explorer.

Notice that in Fig 9.41 above the table tag includes various other settings. The border setting, `border=1`, results in the thin border visible in the final browser display. If `border=0` was used then no border at all would be displayed. Web designers often use tables with invisible borders simply to specify the position of screen elements

precisely. The width and height settings are used to determine either the percentage of the screen used by a table or the precise number of pixels. These settings can also be specified for individual rows and cells. Using percentages for these settings allows the web page to adjust to the resolution of each individual user's screen. Furthermore, the page will reformat appropriately should the user resize the browser window.



GROUP TASK Practical Activity

Enter the HTML code from *Fig 9.41* into a text editor and view the result in a web browser. Alter the cellpadding, cellspacing, colspan and rowspan settings to determine their purpose.

• HTML Frames

Frames split a single browser window into individual sections called frames, where each frame displays a different HTML document. The `<frameset>` tag is used to specify the number of columns (or rows) together with the percentage of screen space or the precise number of pixels each frame is to occupy. In *Fig 9.42* below four frames have been specified. There are two rows where the first row occupies 25% of the available height and the second 75% of the height. The data following this first frameset tag specifies the contents of the first row, in this case two columns have been specified. The first of these columns, which is in the top left corner, is a frame whose source document is the HTML file `lefttop.htm`. The remaining three frames are similarly specified.

A single frame can change without affecting the content of other frames. This is particularly useful for menus as well as header/footer information that is common to a number of web pages.

```
<html>
<head>
  <title>Frame example
</title>
</head>
<frameset rows="25%,75%">
  <frameset cols="70%,30%">
    <frame src="lefttop.htm">
    <frame src="righttop.htm">
  </frameset>
  <frameset cols="85%,15%">
    <frame src="leftbot.htm">
    <frame src="rightbot.htm">
  </frameset>
</frameset>
</html>
```

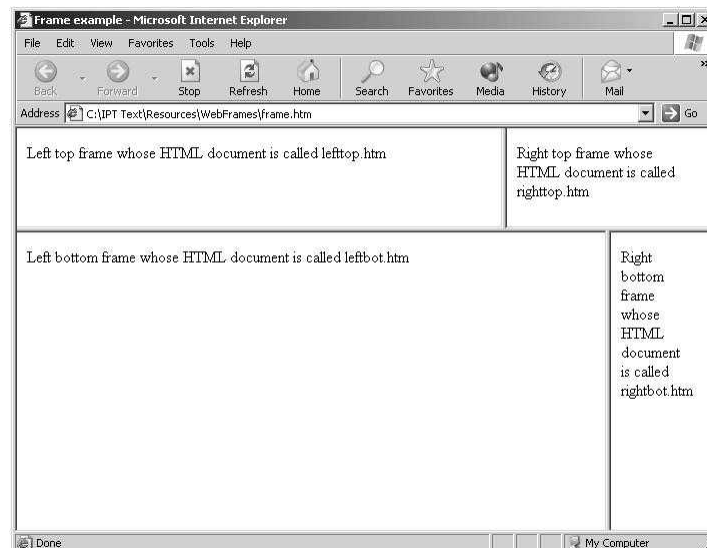


Fig 9.42

Sample HTML frame code and the result displayed within Internet Explorer.



GROUP TASK Practical Activity

Enter the HTML code from *Fig 9.42* into a text editor and save the file. Create a series of four simple HTML documents to display within the frames. Include one or more images in these documents using the `<img src=>` tag. View the result in a web browser. Adjust the size of the frames appropriately to suit your images.

Combining information from different sources

Many software applications are able to combine information from a variety of sources. For example, an image can be included within a text document, and records from a database can be used to produce form letters. In this section, we consider three commonly used techniques for combining information from different sources, embedding, linking and mail-merge.

- **Embedding**

In most applications it is possible to import files created within a variety of other applications into an existing file. The existing file is known as the “destination file” and the file being imported is known as the “source file”. This process is known as embedding as information within the source file becomes part of the destination file. In essence a copy of the source file is inserted into the destination file. For example, the paste command within most applications embeds a copy of the information currently on the clipboard into the current document. The current document is the destination file and the content of the clipboard is the source.

Once the embedding process is complete there is no connection maintained between the original source file and the destination file. The effect being that any future changes made to the original source file will not be reflected within the destination file. The embedded data can be edited from within the destination file using either the same software application or a similar software application to that used to create the original source file.

- **Linking**

Linking does not make a copy of the source file, rather it establishes a connection within the destination file to the source file. Therefore any alterations that are made to the original source file will automatically be reflected within the destination file. For example, a linked spreadsheet within a word processor file will automatically be updated to reflect any alterations made within the source spreadsheet.

HTML hyperlinks are an example of linking. The HTML document is the destination document that contains tags that specify the location and name of the linked source file. When a user clicks on a hyperlink the browser responds by retrieving and displaying the linked source file.



Consider the following:

Within Microsoft Word the options Insert, Link to File, and Insert and Link are available when importing an image file. Insert is the same as embedding, link to file obviously creates a link, but what does insert and link do? It actually embeds the source image and also maintains a link to the original source file.

The bottom screen in *Fig 9.43* is used to edit and update any links within the current document. For example, if the location of a source file is changed then the Change Source... button allows the new location to be specified.

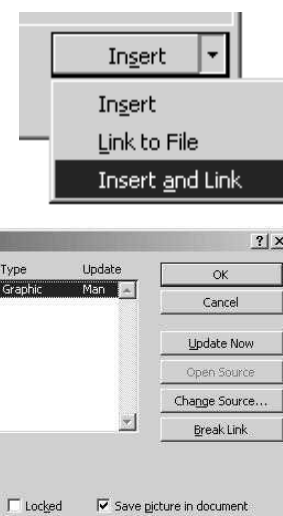


Fig 9.43

Linking and embedding images within Microsoft Word.

**GROUP TASK Practical Activity**

Try embedding and also linking to a single image file within a word processor, that is, the image will appear twice within the document. Alter the original source image and confirm the expected result occurs in the destination document.

**GROUP TASK Discussion**

Identify and discuss situations where linking is suitable and a similar list where embedding is suitable.

**GROUP TASK Discussion**

Explain reasons why it may be useful for a single image to be both embedded and linked within a destination file.

- **Mail-merge**

Mail-merge is a process where information from a database or other list is inserted into a standard document to produce multiple personalised copies. In most cases the standard document is produced within a word processor and fields from the source database or list are inserted. Each record within the data source is used to produce a single personalised copy of the standard document.

In most word processors the standard document is called the main document. The personalised letters produced by the mail-merge process are called form letters, however mailing labels, envelopes and lists can also be produced.

The following processes are completed to perform a mail-merge:

1. Identify or create a data source. Commonly this involves connecting to a database and creating a query to retrieve the desired records and fields.
2. Create the main document. Field codes are inserted within the text of the main document. Most word processors also include functions that allow different text to be displayed based on the value of a particular field. For example, a field may indicate whether an account is overdue, this data could be used to generate an appropriate overdue account message.
3. The final form letters are produced and displayed. Commonly, form letters are printed, however it is possible for them to be emailed or faxed directly from most word processors.

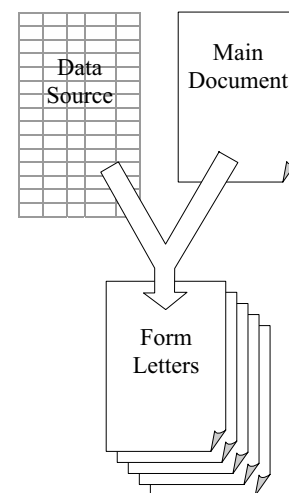


Fig 9.44

Mail merging creates personalized form letters.

**GROUP TASK Discussion**

Much of the mail and email received from businesses and government departments has been mail-merged. List examples you have encountered over the past week.

**GROUP TASK Practical Activity**

Use a word processor to perform a mail-merge using an existing data source. Describe the procedure using a system flowchart.

SET 9B

1. The final bitmap for each displayed screen is generated by:
 - (A) the operating system and the software applications currently executing.
 - (B) the software applications currently being executed.
 - (C) software within the screen itself.
 - (D) the screen's device driver in conjunction with the video card's software.
2. The displayed output from a database is commonly called a:
 - (A) form.
 - (B) query.
 - (C) template.
 - (D) report.
3. The space between lines of text is known as:
 - (A) line spacing.
 - (B) leading.
 - (C) point size.
 - (D) tracking.
4. A printed personalised letter from a government department is most likely to have been produced using which of the following processes?
 - (A) Using a word processor to mail-merge data from a database.
 - (B) Embedding records within a word processor.
 - (C) Linking a database to a document produced using a word processor.
 - (D) Including HTML hyperlinks within a document.
5. With regard to HTML tables, which of the following is true?
 - (A) Each column is defined, followed by each cell within that column.
 - (B) Each cell displays a different HTML document.
 - (C) Each row is defined, followed by each cell within that row.
 - (D) The table must always occupy the entire window within the browser.
6. In regard to top and header margins, which of the following is true?
 - (A) Both margins are measured vertically from the top of the page.
 - (B) The header margin must always be less than the top margin.
 - (C) The sum of the top and header margin determines the vertical start of the body text.
 - (D) The top minus the header margin determines the vertical start of the body text.
7. Which of the following terms describes the ability of a printer to print on both sides of a piece of paper?
 - (A) Simplex
 - (B) Duplex
 - (C) Layout
 - (D) Orientation
8. The optical centre of a page is:
 - (A) the same as the physical centre.
 - (B) below the physical centre.
 - (C) above the physical centre.
 - (D) above and to the right of the physical centre.
9. Courier 12-point italic is an example of:
 - (A) a font.
 - (B) a typeface.
 - (C) a font style.
 - (D) a serif font.
10. A chart is included within a word processor document. This chart changes when data in the underlying spreadsheet is edited. Which of the following is most likely?
 - (A) The chart is embedded within the word processor document.
 - (B) The chart was copied and pasted into the word processor document.
 - (C) The chart has been mail-merged using the word processor.
 - (D) The chart is linked to the spreadsheet from the word processor document.

11. Define each of the following terms?
 - (a) font
 - (b) leading
 - (c) serif
 - (d) footer
 - (e) tracking
 - (f) typeface
 - (g) X-height
 - (h) gutter
 - (i) kerning
 - (j) optical centre
12. Compare and contrast:
 - (a) HTML tables with HTML frames.
 - (b) linking and embedding.
 - (c) reporting within database applications with mail-merging within word processors.
 - (d) the resolution of screen displays with the resolution of printed displays.
13. Identify the software used and describe the processes occurring once the print command has been issued from within a software application.
14. Designing displays for the web introduces problems that are not present when designing displays for printing. Identify these problems and discuss possible solutions.
15. The following statement was made by a graphic designer:

“The design of displays is at least as important as the information they contain”

Discuss the validity of the above statement. Use examples of both printed and screen displays to assist your discussion.

NON-COMPUTER TOOLS FOR DISPLAYING

Today's computer tools are capable of displaying information of all types however there remain many non-computer tools that are used for display or that assist in the design of displays. For example, traditional books are still purchased and used in preference to their computer-based alternatives. Pen and paper is used for personal letters and inter-office memos despite the availability of computer-based alternatives. System models, such as dataflow diagrams are often hand drawn. Similarly the initial screen designs for many computer-based displays are developed using hand drawn storyboards.

In this section, we first consider storyboarding as a technique for designing screens, we then consider some examples of traditional methods of display that remain popular.

STORYBOARDS

Story boarding is a technique that was first used for the creation of video information, including film, television and animation. These storyboards show a hand drawn sketch of each scene together with a hand written description.

Video data by its very nature is linear, that is scenes are arranged into a strict sequence. Similarly printed information of all types must also be linear. However screen displays are different, they provide the ability for users to navigate screens in a variety of different sequences. As a consequence, storyboards created for computer-based screen display are typically composed of two primary elements; the individual screen layouts and descriptions, together with a diagram illustrating the possible navigation paths between these screens.



Consider the following:

The diagram below shows a hierarchical screen structure for a company's web site. Although the structure is basically hierarchical, a link is provided on each screen to return the user to the home page.

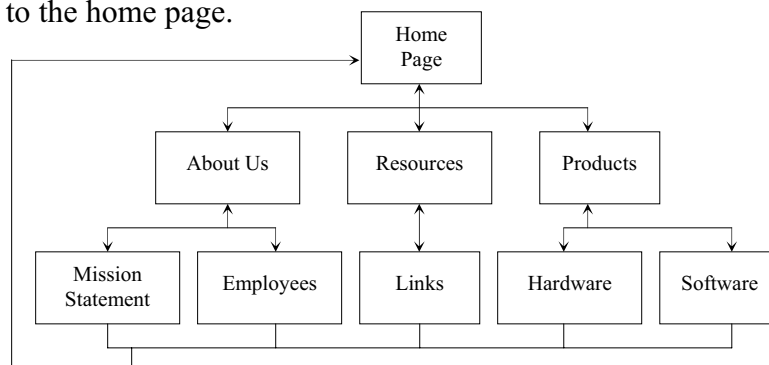


Fig 9.46

Hierarchical navigation between screens is common for web sites.

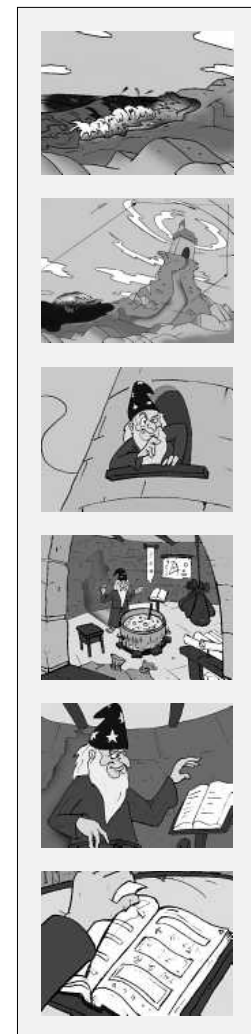


Fig 9.45
Video storyboards are always linear.



GROUP TASK Practical Activity

Examine a company's web site and create a diagram similar to the one shown above in Fig 9.46 describing the links between pages on the site.

TRADITIONAL METHODS OF DISPLAY

Although computer-based displays are enormously popular it is unlikely they will ever totally replace traditional methods of display. An original painting, a live concert or even a simple handwritten note conveys information that is difficult to reproduce electronically. Indeed it is often difficult to even identify the nature of this added information, so how can it possibly be reproduced on a computer?

In this section our aim is to consider and appreciate the advantages of traditional display techniques compared to computer display techniques. Many information systems produce information suited to computer display whilst other information is better displayed using traditional display techniques. In many cases a combination of both computer and non-computer based display is appropriate. The remainder of this section presents a number of scenarios for discussion.



Consider the following:

- A company receives approximately 300 queries by fax. Currently the large majority of these queries are dealt with immediately by simply handwriting and faxing back a reply within minutes of the original fax being received. A current employee suggests the current system should be replaced with one whereby replies are professionally typed and faxed directly from a computer – the management disagrees.
- You are experiencing difficulty with one of your subjects. An interactive online tutoring service is available at minimal cost, however you decide to use a real tutor. Using the real tutor costs substantially more and requires you to catch a bus to the tutor's home and back.
- An area manager purchases and commences using a personal digital assistant (PDA) to maintain their appointment diary and other business notes. After 6 months the manager reverts to using a traditional pen and paper diary system.
- A school is unable to create a class for a particular HSC course despite 15 students choosing that course. The school organises for the students to study the course via a largely online correspondence system. In the end only 2 of the initial 15 students decide to enrol in the course.
- Aunt Lucy has always sent out handwritten birthday cards to everyone. In the past 12 months she has begun sending emails instead of handwritten cards. At a recent family function many people comment on how disappointed they were to receive an impersonal email.
- Many people when performing Internet research print out items of interest rather than reading them directly on the screen. Some take the printouts to read in different locations but many simply prefer to read from paper.



GROUP TASK Discussion

Identify a computer and a non-computer based display solution suggested within each of the above scenarios.



GROUP TASK Discussion

Compare and contrast the displays created with a computer with those created without a computer for each of the above scenarios.

SOCIAL AND ETHICAL ISSUES ASSOCIATED WITH DISPLAYS

Information output from an information system is ultimately intended for a human audience. As a consequence this information must be presented in a manner that can be understood by the intended audience. This is a primary objective of the displaying information process. Often this audience includes people with a variety of different disabilities and from different age groups and backgrounds. As computer-based displays are heavily reliant on sight, those with visual impairment require particular consideration.

Some issues of particular importance to the displaying information process include:

- communication skills of those presenting displays,
- appropriate displays for the visually impaired, and
- displays suitable for young children.

We examine these issues and then conclude with a brief examination of past, present and emerging trends in displays.

COMMUNICATION SKILLS OF THOSE PRESENTING DISPLAYS

It is often said that the communication skills of the presenter are at least as important as the information they are presenting. If this is indeed true then what are the essential qualities of a good presenter? Let us consider answers to this question by first examining common mistakes and then thinking about some of the ingredients present in successful presentations.

Common presentation mistakes include:

- Reading every word of the presentation. It is not possible to interact with an audience when reading continuously. Furthermore, the audience is left wondering why they are there – surely they could simply have been handed the notes.
- Using too many slides within a PowerPoint presentation. Every slide should have a specific well-defined purpose. Slides should be used to emphasise important points or to illustrate points using graphs and images. Too many slides causes the audience to focus on the slides rather than the presenter, this creates a barrier between the presenter and the audience.
- Including large amounts of textual information can exhaust the audience as their time is spent reading rather than listening. Furthermore, most people can read much faster than the presenter can speak, hence the audience is likely to read ahead causing the words spoken by the presenter to become redundant.
- Using too many and inappropriate design features. For example, overusing animation, sound, video and other effects. These features can easily detract from the actual information rather than enhance its delivery. Such features should support the information presented.
- Attempting to deliver too much information can result in less actual information being understood and absorbed by the audience. Further detail can be given in the form of references or printed notes.
- Presenting information that is too complex or too trivial for the audience. The level and depth of treatment must suit the audience and the time available.



GROUP TASK Discussion

Think about presentations you have viewed that you consider to have been successful. List and describe aspects of these presentations that contributed to their success.

APPROPRIATE DISPLAYS FOR THE VISUALLY IMPAIRED

Most information is output visually from information systems, using either screens or printers. As a consequence various hardware and software tools are available to assist in the display of information to visually impaired people. Software is available to alter the colour and size of displays. For example, increasing contrast and enlarging fonts. Such tools allow those with limited sight or particular visual impairments such as colour blindness to view information in a more suitable visual form. For those with very limited or no sight, alternative display hardware is needed. Braille screens and embossers replace standard screens and printers. Furthermore, speech synthesis software can be used to convert textual information into speech.



GROUP TASK Practical Research

Identify features designed to assist the visually impaired within software on either your home or school computer.

Braille Screens

Braille screens usually include both input and output elements. Such screens are generally positioned underneath a standard keyboard (see Fig 9.47). The input elements allow users to navigate around the screen and to control various speech synthesis features. The display is composed of a row of Braille display cells similar to the one shown in Fig 9.48. Each Braille cell contains a grid of pins that rise and fall to create the Braille symbols. Traditionally each character is represented using a grid of 6 pins, however many Braille computer displays use 8 pins to enable the display of 256 different characters.

The Alva Satellite in Fig 9.47 contains two 'satellite' pads on either side of the Braille display. These pads are used to perform navigation functions similar to those of a traditional mouse. A row of touch switches is located directly above the row of Braille displays cells. These touch switches are used to quickly request speech feedback about the information displayed by the corresponding Braille display cell. With the appropriate software the display cells are able to create texture maps of images as well as Braille symbols for text and numeric information.

The Alva Satellite range of Braille displays connect to a standard personal computer via a USB port. For standard software applications, including web browsers, word processors and spreadsheets, there is no need to install any additional software or alter any settings on the computer. The display is simply connected and configured via communication over the USB interface.



Fig 9.47

The Alva Satellite 584 Pro Braille screen.

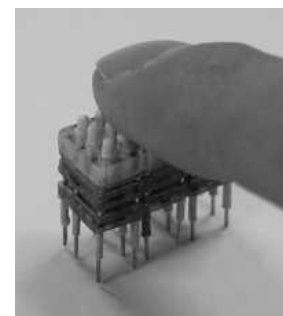


Fig 9.48

Braille display cell.



GROUP TASK Research

Using the Internet, or otherwise, determine how characters are represented using the Braille system.

Braille embossers

Braille embossers operate using a series of small hammers and anvils. Each hammer produces a single raised Braille dot when it strikes the paper and collides with its corresponding anvil. The Index Everest Braille embosser shown in *Fig 9.49* contains 13 hammers and is able to emboss on standard copier paper at speeds up to 95 characters per second.

The controls keys on most current embossers include Braille characters and many models provide synthesised speech as feedback should a setting be changed or an error occur.



Fig 9.49
The Index Everest Braille embosser.

Speech Synthesis

Synthetic speech systems are used for a variety of applications, including voice mail systems, warning systems and in particular for reading applications. Clearly it is reading where speech synthesis systems are of particular use to the visually impaired. Applications that require a limited vocabulary are able to use digital voice recordings.

However the ability to read text requires an unrestricted vocabulary. These are known as text to speech (TTS) systems and they are far more complex than systems based on digital voice recordings. In the last 10 years speech synthesis systems have progressed from science fiction to the point where they are now included within many mainstream operating systems – *Fig 9.50* is a screen shot from within Windows XP. More fully featured packages are available for use by visually impaired users.

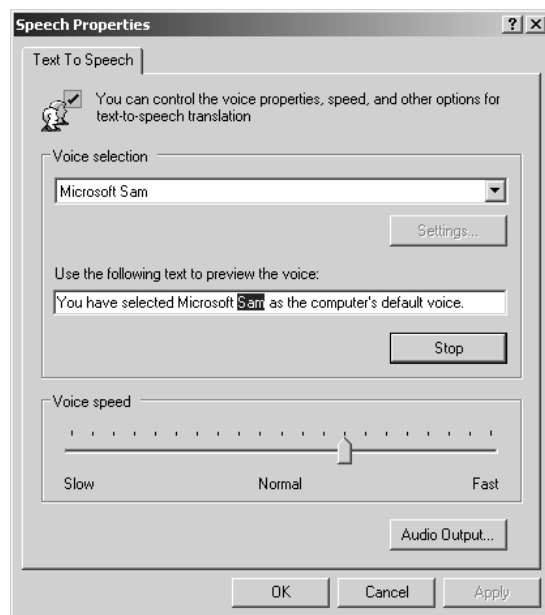


Fig 9.50
Basic speech synthesis is included with Microsoft's Windows XP operating system.

Almost all speech synthesis systems are software based. The speech synthesis software operates between other software applications and the operating system. Such systems allow the functionality of the speech synthesis software to be available to a variety of different software applications in much the same way as device drivers provide an interface between hardware devices and software applications. Currently the Speech Application Programming Interface (SAPI) is the standard within most operating systems that defines the way in which software applications communicate with speech synthesis software. Software applications that are SAPI compliant should in theory operate with any SAPI compliant speech synthesis software.

Most speech synthesis packages integrate with scanners and optical character recognition software to enable hardcopy to be read by the software. It is also common for these packages to include speech recognition capabilities. Speech recognition converts audio voice samples into text. Speech recognition combined with speech synthesis allows people with hearing impairment to communicate using speech.



Consider the following:

On the 7th of September 2004 the World Wide Web Consortium (W3C) released its recommendation for a new web standard called Speech Synthesis Markup Language (SSML). The essential role of SSML is to provide authors of synthesisable web content a standard way to control aspects of speech such as pronunciation, volume, pitch and rate across different synthesis-capable platforms.

SSML uses tags similar to those used within HTML documents to instruct speech synthesis systems in regard to the pronunciation of text elements. One example within the recommendation discusses the question “How do you pronounce 1/2?” Should it be read as “one half” or “1st of February” or “2nd of January” or “one divided by two”. The tags within SSML aim to resolve such dilemmas.



GROUP TASK Discussion

Discuss implications of SSML for those with visual and hearing impairments and also those involved in the creation of web content.



GROUP TASK Discussion

SSML is not just about providing access to web content for the visually impaired. One of its primary aims is to allow content to be accessed using standard mobile and fixed telephones. Identify reasons why such access would be desirable.

DISPLAYS SUITABLE FOR YOUNG CHILDREN

Young children possess a natural motivation to learn and absorb new information. They learn best when they are having fun. They enjoy exploring and discovering in an unstructured random order, yet they also appreciate familiar content upon which new knowledge is built. Indeed most young children enjoy repeating the same activity many times. For example, preschoolers enjoy viewing their favourite video hundreds of times. Therefore computer displays for young children require a balance between familiar elements and opportunities for exploring new areas of interest.

Young children are in the process of developing their reading and fine motor skills. They quickly become disinterested when they are unable to understand how to operate software or have difficulty using the keyboard or mouse. Screen displays for young children must take account of these realities.

Let us consider some general criteria worth considering when creating or selecting software for young children.

- The program should be fun to use and it should have educational value. The educational value should not distract from its enjoyment value.
- Reading ability should not be assumed. Rather, large graphical icons depicting familiar and intuitive items are more appropriate. Many displays replace drop down menus with a complete screen of say a bedroom. Hotspots within the image initiate different activities.
- The use of colour, animation, speech and sound greatly improves the experience for young children. However these aspects of the displays should always be appropriate to the information being displayed. Young children are easily distracted by elements not directly related to the content.

- Inappropriate inputs, such as randomly pounding on the keyboard or holding down a key should not cause errors. In general, such inputs should be ignored.
- Children should be able to use the program without assistance. This allows them to choose their own path through the software or perhaps to repeat favourite parts. For this to occur requires that navigation between screens be as simple as possible.
- Feedback in response to inputs should be quick and it should be clear. The highly interactive nature of software is a primary motivator. Incorrect inputs should be dealt with in non-judgemental ways. Learning occurs through mistakes, so responses should always be positive.
- It should be easy to return to activities or get out of activities. Even within individual activities it should be simple to repeat sections or go back a step. In essence the child should have control over the order of display.



Consider the following:

Millie's Math House (Edmark Software)

Kids, ages 3-7, develop a love of math when they play in Millie's Math House.

Through 7 fun filled activities that feel like play children learn about numbers, counting, addition, subtraction, patterns, problem solving, size, geometric shapes and much more.

Kids count critters, build mouse houses and create crazy-looking bugs while learning early math concepts. Kids learn to identify and compare shapes and sizes, create and complete patterns, learn numbers to 30 and practice addition and subtraction.



Harley the Horse wants some cookies. At the Cookie Factory kids decorate cookies with 0 to 20 jellybeans to make Harley a tasty treat. Kids can make cookies as they choose or play a fun game. Harley will tell players how many jellybeans he wants on his cookies.

Dorothy Duck is asking What's My Number? Players must place the same number of objects on the stage as Dorothy has on hers. Kids explore simple addition and subtraction with Dorothy. Problems are illustrated with objects making understanding easy.

Little, Middle and Big need shoes. There are large, small and medium shoes. Kids have tons of fun choosing just the right shoes for these three crazy characters.

The Number Machine is lots of fun. Kids click on a number and watch as crazy critters pop-up and are counted out loud. Children can count 9 worms, 23 penguins and more. Kids can also play a fun game where they are asked to find a number.

Build-A-Bug is a zany game that kids go crazy over. Children choose bug parts, eyes, antennae, ears, feet, tails and spots and then a number to build their bug. In mode two kids are challenged to follow written and verbal directions to build just the right bug.

Players will also explore shapes and discover patterns as they build mouse houses and play with Bing and Bong.

Activities have two modes, Explore and Discover and Question and Answer. Kids learn while exploring freely in the first mode and can then play fun games in the second mode to test their knowledge.

Learning math is a blast with Millie and her friends in Millie's Math House.

(Source: kidsclick.com)



GROUP TASK Discussion

Read the information above describing “Millie’s Math House”. Identify aspects of this program that make it suitable for children aged 3–7 years.

PAST, PRESENT AND EMERGING TRENDS IN DISPLAYS

The enormous increase in processing power, access speed and capacity of display memory, together with advances in display hardware technologies has had a profound effect on the nature of computer based displays. This includes both screen and printed displays.

Past trends

In the early 1950s computers used punched cards for both collecting and displaying data. *Fig 9.51* shows the IBM 533 Card Punch Reader which was the primary input and output device for IBM’s 650 computer produced from 1954-1962. During the mid 1950s typewriter style devices became common for both input and output. The IBM 838 Inquiry Station shown in *Fig 9.52* was used to transmit simple requests to the computer – the resulting responses being typed out. At the time this was seen as a major advance towards humanising computers.

It was not until the early 1970s that CRT based terminals became widely available. These terminals allowed many remote users the ability to share the computer’s processing and storage resources. The most successful example was Digital Equipment Corporation’s VT100 terminal (see *Fig 9.53*), whose monochrome screen was able to display 24 lines of 80 column text. Most modern terminal emulation software still communicates using the VT100 standard.

Prior to the release of laser printers in the late 1970s, all printers were based on impact principles. Drum and daisy wheel printers used preformed characters to impact an inked ribbon onto the paper. Dot matrix printers used a series of pins to impact a ribbon and form characters (and also images) on paper. Dot matrix printers remained popular until the early 1990s. It wasn’t until 1988, when the first inkjet printer was released by Hewlett Packard, that the popularity of dot matrix printers declined. Although later examples of dot matrix printers were able to produce colour output the quality was extremely crude compared to modern colour inkjet and laser output.



Fig 9.51
The IBM 533 Card Punch Reader with inset showing the punched cards used by the device.



Fig 9.52
The IBM 838 Inquiry Station used for both input and printed output.



Fig 9.53
Digital Equipment Corporation’s VT100 terminal was released in 1978.

Present and emerging trends

Display devices continually improve in terms of resolution, colour reproduction and speed. Physically both screens and printers are smaller in size, yet they contain much faster processors and larger amounts of storage. Furthermore, prices continue to drop, for example in the mid 1980s a simple dot matrix printer retailed for approximately the same price as a full colour laser printer costs today.



Consider the following:

Some present and future display trends at the time of writing include:

- Flexible screens or ‘electronic paper’ that can be rolled or folded. At the time of writing operational prototypes had been released by various companies. According to Philips, mass production is likely to commence within a few years. *Fig 9.54* shows an example of one of Pioneer’s prototypes. Currently the most promising designs are based on organic electroluminescence display (OELD) technology.
- 3D printers form objects based on 3-dimensional models. Some construct objects by laying down layers of fine plastic particles. The print head then deposits a glue-like substance on areas that will form the final object. Once all layers have been printed the loose particles are removed to reveal the final object. Other 3D printers use high powered lasers to burn away material and reveal the final object. Although commercial 3D printers have been available for a number of years there are now models under development that are aimed at the home market.
- Holograms are 3 dimensional projections that appear to float in space. They are often used in science fiction movies, however at the present time such displays are not a reality. Currently holographic type images are generated using laser beams and some sort of transparent screen or gas filled structure. It seems likely that eventually true holographic technologies will be developed.

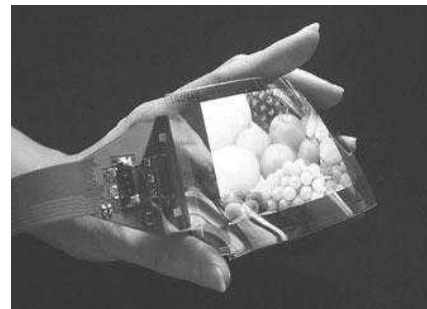


Fig 9.54
A prototype of Pioneer’s flexible full colour screen.



Fig 9.55
Hologram type image of an actress projected onto a transparent screen during a live stage production.



GROUP TASK Discussion

Identify and discuss possible applications of each of the display technologies described above.



GROUP TASK Research

Using the Internet, or otherwise, identify recent advances in standard screen and printer technologies.

SET 9C

1. Which of the following is best suited to computer-based display processes?
 - (A) original paintings
 - (B) museum artefacts
 - (C) survey results
 - (D) stage production
2. A storyboard includes:
 - (A) the layout of each screen.
 - (B) descriptions of each screen.
 - (C) navigation between screens.
 - (D) All of the above.
3. Braille screens primarily rely on:
 - (A) sight
 - (B) hearing
 - (C) touch
 - (D) smell
4. Successful slideshow presentations are likely to include which of the following elements?
 - (A) Each slide has a specific well-defined purpose.
 - (B) The slides contain extra detailed information that cannot be addressed during the presentation.
 - (C) Each individual slide includes some type of special effect.
 - (D) The presenter reads each slide to ensure all the information is heard by the audience.
5. Which of the following is FALSE in regard to most TTS systems?
 - (A) They have an unrestricted vocabulary.
 - (B) They are primarily software based.
 - (C) Simple examples are included in many operating systems.
 - (D) They require many digital voice recordings.
6. The Braille system represents characters:
 - (A) using a grid of pins that rise and fall.
 - (B) using different combinations of raised dots.
 - (C) using hammers and anvils.
 - (D) by embossing the outline of each character.
7. In regard to displays for young children, which of the following is true?
 - (A) Background music and sound effects are important motivators.
 - (B) The software should ensure that activities cannot easily be repeated.
 - (C) Data entry errors should be dealt with firmly.
 - (D) Navigation between screens should be simple.
8. Which of the following is FALSE in regard to VT100 terminals?
 - (A) They were released and used with IBM's 650 computer.
 - (B) They included a CRT based display.
 - (C) Most terminal emulation software still communicates using the VT100 standard.
 - (D) They were released in 1978.
9. Punched cards were used for:
 - (A) data input.
 - (B) program input.
 - (C) output.
 - (D) All of the above.
10. The ability to convert digital voice samples into text is known as:
 - (A) speech recognition.
 - (B) speech synthesis.
 - (C) TTS
 - (D) SAPI
11. Define each of the following terms?

(a) storyboard	(d) speech synthesis
(b) linear	(e) Braille
(c) hierarchical	(f) punched card
12. Identify and describe features present within the operating system on your home or school computer that are designed to assist those who are:
 - (a) visually impaired.
 - (b) hearing impaired.
 - (c) mobility impaired.
13. Identify and research the operation of a specific model of display device used during the:

(a) 1950s	(b) 1960s	(c) 1970s
14. Identify, research and describe an experimental display device.
15. Download a free or trial version of a software product designed for young children. Evaluate the suitability of this product in terms of its screen designs.

CHAPTER 9 REVIEW

1. Which of the following is true of displaying information processes?
 - (A) Information is altered.
 - (B) Output from the system is generated.
 - (C) Data is transformed into information.
 - (D) Information from a variety of sources is combined.
2. Which list contains only display devices?
 - (A) speaker, inkjet printer, LCD monitor.
 - (B) microphone, keyboard, digital camera.
 - (C) Video card, AGP bus, VGA cable.
 - (D) CPU, RAM, ROM.
3. Different colours are displayed on a CRT by:
 - (A) varying the current to the magnetic steering coils.
 - (B) varying the current to each pixel's transistor.
 - (C) firing the electron beams at different intensities.
 - (D) increasing or decreasing the refresh rate.
4. In regard to LCD screens, which of the following is true?
 - (A) They require a separate light source.
 - (B) Light is produced within the polarizing panels.
 - (C) Each TFT produces its own light.
 - (D) The liquid crystals glow when electrical current is applied.
5. Examples of MEM devices include:
 - (A) TFTs, LCDs and LCOS chips.
 - (B) Lasers and piezo crystals.
 - (C) inkjet nozzles and stepper motors.
 - (D) DMDs and GLVs.
6. The type of ink used is less critical when an inkjet printer is based on which technology?
 - (A) Thermal
 - (B) Piezoelectric
 - (C) CMYK
 - (D) Bubblejet
7. Which of the following is true in regard to the size of speaker diaphragms?
 - (A) Larger diaphragms reproduce higher frequencies and smaller diaphragms reproduce lower frequencies.
 - (B) Larger diaphragms reproduce lower frequencies and smaller diaphragms reproduce higher frequencies.
 - (C) Larger diaphragms reproduce all frequencies well whilst smaller diaphragms are unable to reproduce lower frequencies.
 - (D) Smaller diaphragms reproduce all frequencies well whilst larger diaphragms are unable to reproduce higher frequencies.
8. In regard to serif and sans serif fonts, which of the following is generally true?
 - (A) Serif fonts are used for printed hardcopy, sans serif fonts are used in most other circumstances.
 - (B) Sans serif fonts are used for printed hardcopy, serif fonts are used in most other circumstances.
 - (C) Serif fonts are used for body text, sans serif fonts are used for titles.
 - (D) Sans serif fonts are used for body text, serif fonts are used for titles.
9. HTML tags are acted upon by:
 - (A) web servers.
 - (B) network software.
 - (C) web browsers.
 - (D) web pages.
10. A word processor document is emailed. The receiver finds that the images are not displayed. What is the most likely reason?
 - (A) The images were embedded.
 - (B) The images were linked.
 - (C) The images were embedded and linked.
 - (D) The images were mail-merged.
11. What does each of the following abbreviations stand for?

(a) CRT	(d) DVI	(g) IWB	(j) LCOS	(m) GLV
(b) LCD	(e) HDMI	(h) TFT	(k) MEM	(n) CMYK
(c) DAC	(f) GPU	(i) dpi	(l) DMD	(o) DSP
12. Explain how images are produced by:

(a) Inkjet printers	(c) CRT monitors	(e) Transmissive LCD projectors
(b) Laser printers	(d) LCD monitors	(f) DMD based projectors
13. In most applications each character typed is displayed on the screen. Create a list to describe the sequence of processes occurring.
14. Compare and contrast common measures of screen resolution with common measures of printer resolution.
15. Create a set of guidelines worth considering when laying out page elements for:
 - (a) printed display
 - (b) screen display

In this chapter you will learn to:

- recognise and apply appropriate stages in their project work
- read and interpret the requirements for a new system in terms of:
 - the needs of the users of the information system
 - who the participants are
 - the data/information to be used
 - required information technology
 - information processes
- use a variety of design tools to help plan the structure of an information system
- use an information system to generate information
- read a set of specifications
- understand the need for a time schedule
- interpret Gantt charts
- understand the need for journals and diaries
- recognise the resources that are relevant, available and required for use in developing the system
- modify or extend an existing system according to specifications
- test and evaluate an existing system to see if it meets requirements and specifications
- recognise different roles of people and how they communicate throughout different stages of the project
- produce a report stating the need, and how an information system will meet it
- diagrammatically represent the information system in context
- document the relationship between the new system, user of the information system and their need(s)
- analyse and customise user interfaces and other tasks in applications software forming part of the solution
- identify the training needs of users of the information system
- document the procedures to be followed by participants
- implement systems that pay as much attention to the needs of participants as they do to information technology

In this chapter you will learn about:

Traditional stages in developing a system

- understanding the problem
- planning
- designing
- implementing
- testing, evaluating and maintaining

Complexity of systems

- systems for individuals
- systems for organisations
- systems developed by individuals
- systems developed by teams

Roles of people involved in systems development

- different roles played by individuals in the team and communication between them
- strengths and weaknesses of individual team members
 - communication
 - interpersonal
 - technical
 - organisational

Social and ethical issues

- machine-centred systems simplify what computers do at the expense of participants
- human-centred systems as those that make participants' work as effective and satisfying as possible
- how the relationships between participants change as a result of the new system
- ensuring the new system provides participants with a safe work environment
- awareness of the impact the system may have on the participants, including:
 - opportunities to use their skills
 - meaningful work
 - need for change
 - opportunities for involvement and commitment

Which will make you more able to:

- describe the nature of information processes and information technology
- classify the functions and operations of information processes and information technology
- identify and describe the information processes within an information system
- recognise and explain the interdependence between each of the information processes
- identify and describe social and ethical issues
- describe the historical developments of information systems and relate these to current and emerging technologies
- select and ethically use computer based and non-computer based resources and tools to process information
- analyse and describe an identified need
- generate ideas, consider alternatives and develop solutions for a defined need
- recognise, apply and explain management and communication techniques used in individual and team-based project work
- use and justify technology to support individuals and teams.

DEVELOPING INFORMATION SYSTEMS

INTRODUCTION TO SYSTEM DEVELOPMENT

New information systems are developed when either an existing system no longer meets the needs of its users or new needs are identified that could be met by an information system. In this chapter we begin with an outline of the traditional structured method of developing information systems. We then examine the complexity of information systems including various tools, resources and documents used and produced to assist the management and development process. Throughout our discussions we introduce the different roles of development personnel and the strengths required to fulfil these roles.

There are various other alternative system development approaches, many of which are presented in the HSC course. These alternative approaches still include similar activities to those present in the traditional structured approach, however they perform these activities in different sequences and with different emphasis. In the Preliminary course we restrict our discussion to the traditional structured approach not because it is the best approach, but because it provides a structured introduction to the activities performed during the development of all information systems.

The traditional structured approach to system development specifies distinct stages or phases which are completed in sequence. These stages combine to describe the activities or processes needed to develop an information system from an initial idea through to its final implementation and ongoing maintenance. The complete development process is known as the 'System Development Life Cycle (SDLC)' or simply the 'System Development Cycle (SDC)'. In this text we will use the abbreviation SDLC. The SDLC is closely linked to the concept of structured systems analysis and design, where a series of distinct steps are undertaken in sequence during the development of systems.

During each stage of the SDLC a specific set of activities is performed and each stage produces a specific set of outputs. These outputs are commonly called 'deliverables'. For example, a requirements report is an example of a deliverable that describes what must be done to achieve the system's purpose. In general the deliverables from each stage of the SDLC form the inputs to the subsequent stage. For example, the requirements report provides crucial input when comparing the feasibility of different proposed solutions. In IPT the SDLC is split into five distinct stages, namely Understanding the problem, planning, designing, implementing and finally testing, evaluating and maintaining.



GROUP TASK Discussion

The five stages of the IPT SDLC are similar to the stages used to develop all types of systems. Consider building a new house. Outline what would occur during each of these five stages when building a new house.

TRADITIONAL STAGES IN DEVELOPING A SYSTEM

The particular stages or phases within the SDLC differ depending on the needs of the organisation and also on the nature of the system being developed. As a consequence different references split the SDLC into slightly different stages. In the IPT syllabus the SDLC is split into five stages, namely

1. Understanding the problem
2. Planning
3. Designing
4. Implementing
5. Testing, evaluating and maintaining

In this chapter we outline the activities occurring during each stage. The overall activities performed are similar regardless of the number of distinct stages. The five stages specified in the IPT syllabus describe one method of splitting the SDLC, but of course there are numerous other legitimate ways of splitting the SDLC into stages.



Consider the following sets of SDLC stages

The SDLC policy (1999) of the U.S. House of Representatives specifies and describes the following seven phases:

1. Project Definition
2. User Requirements Definition
3. System/Data Requirements Definition
4. Analysis and Design
5. System Build
6. Implementation and Training
7. Sustainment

The HSC Software Design and Development (SDD) course focuses on the creation of software rather than total information systems. In terms of information systems the development of software is just one part of the solution. In the SDD syllabus the version of the SDLC used is called the Software Development Cycle and is split into the following five stages:

1. Defining and understanding the problem
2. Planning and design of software solutions
3. Implementation of software solutions
4. Testing and evaluation of software solutions
5. Maintenance of software solutions

Many Systems Analysis and Design references use SDLC stages similar to one of the following:

1. Investigation	1. Planning	1. Requirements
2. Design	2. Analysis	2. Analysis
3. Construction	3. Design	3. Design
4. Implementation	4. Build	4. Construction
	5. Implementation	5. Testing
	6. Operation	6. Acceptance



GROUP TASK Discussion

Compare and contrast each of the above lists of SDLC stages with the stages specified in the IPT syllabus.

Before we discuss each stage of the SDLC let us briefly identify the activities occurring and the major deliverables produced during each stage of the IPT syllabus version of the SDLC. The dataflow diagram in *Fig 10.1* shows each stage as a process, and the significant deliverables as the data output from each process. The deliverables from all previous stages are used during the activities of each subsequent stage. To improve readability these dataflows have not been included on the diagram. The grey circular arrow behind the diagram indicates the sequence in which the stages are completed.

Users are included on the diagram as their input is central to the successful development of almost all information systems. Indeed it is often ideas from users that initiate the system development process in the first place. Furthermore, the needs of users largely determine the requirements of the new system. As a consequence feedback from users is vital during the SDLC if the requirements are to be met and are to continue to be met.

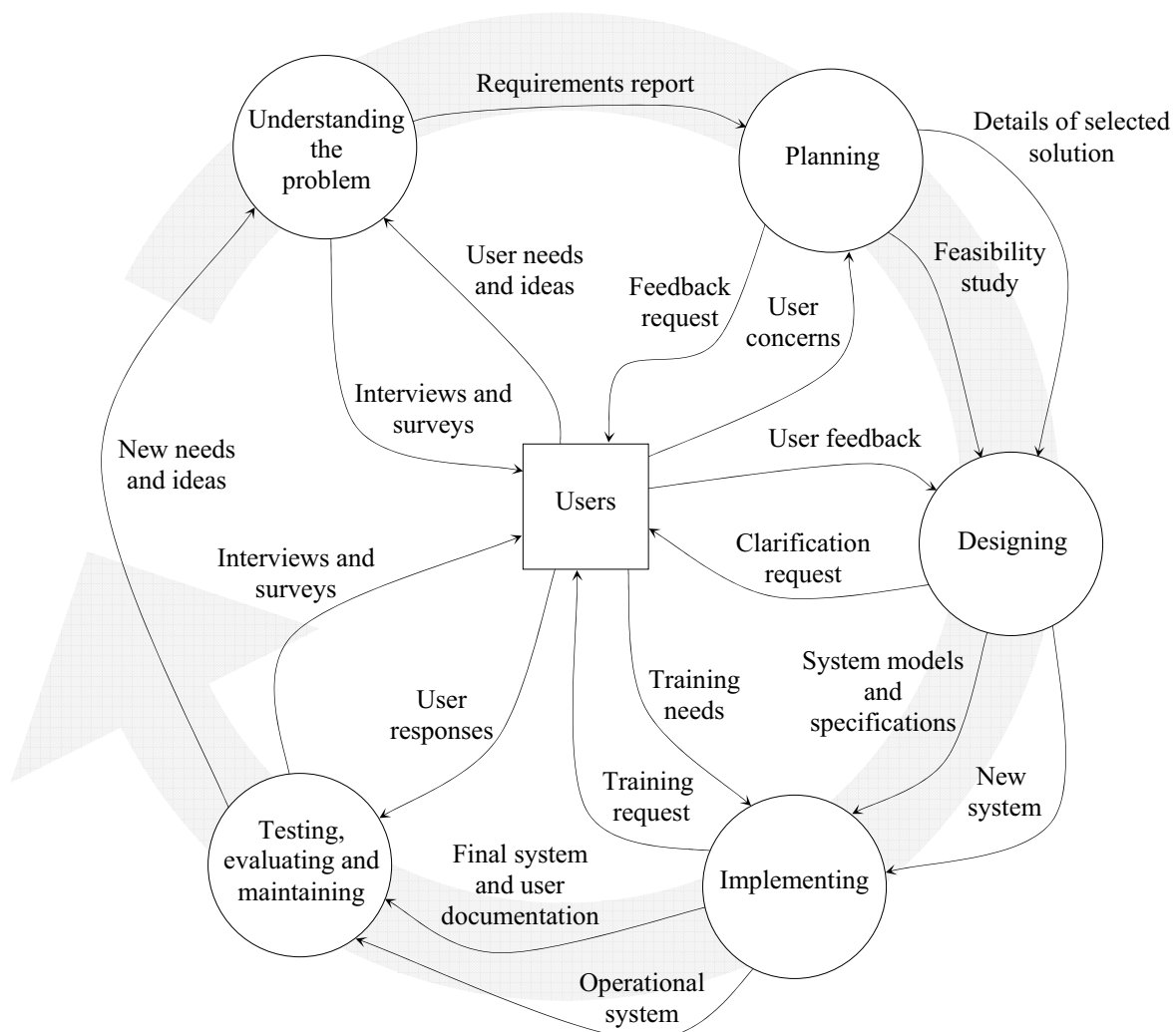


Fig 10.1
The version of the System Development Lifecycle (SDLC) used in IPT.



GROUP TASK Discussion

The above diagram implies some of the activities occurring during each stage of the SDLC. Identify and discuss the general nature of the activities occurring during each stage.

UNDERSTANDING THE PROBLEM

The primary aim of this first stage of the SDLC is to determine the purpose and requirements of a new system. It is not until the requirements have been established that possible solutions can be considered. The Requirements Report is therefore the essential deliverable produced by this stage. A Requirements Report defines the precise nature of the problem to be solved. In essence this stage determines what needs to be achieved to make the system a success.

A systems analyst is a person who analyses systems, determines requirements and then plans information systems. They are problem solvers who possess strong analytical and communication skills. In relation to 'understanding the problem' the systems analyst completes and/or manages activities which aim to determine the needs of the users, participants and management and the processes performed by the current system. This information allows the analyst to determine the purpose of the new system which is further refined with the help of system models to finally develop a list of requirements in the form of a Requirements Report. The new system must achieve all the requirements if it is to meet the identified purpose.



Systems Analyst

A person who analyses systems, determines requirements and designs new information systems.

DETERMINING NEEDS AND UNDERSTANDING THE EXISTING SYSTEM

So what type of activities are commonly performed to determine needs and also to understand the existing system? For systems used and designed by individuals this is a relatively simple process. Presumably the individual already understands their own needs and can develop the system accordingly. For larger systems (or where a team of developers is involved) more structured activities are required. Some common activities include:

Interviewing users including management and participants.

Interviews are performed either in person or over the phone. Interviews can be formal with prepared questions designed to obtain specific information, informal where the interviewee leads the discussion according to their needs or a combination of both.

Often the initial idea to consider a new system occurs because of problems or inadequacies with an existing system. The general nature of these issues is first expressed during an initial informal interview with management. Further detail and new issues will be uncovered during subsequent and more formal interviews with participants and other users of the existing system. In most cases interviews are an excellent means of identifying new needs, but often they do not uncover needs that are already being met by the existing system. Users and management often assume new systems will include all the functionality present within their existing system. The analyst must determine both new and existing needs. Other strategies, such as task analysis, more effectively determine needs that are already being met.

Surveying users including participants.

Surveys are paper or electronic questionnaires. As such they must be prepared in advance. This means the responses can only hope to gather information about issues the analyst has already anticipated. Surveys are often distributed to a large number of people and the results are summarised using a spreadsheet or other analysis tool. Such surveys are designed so the responses are easy to analyse using computers. For example, rating agreement on a scale from 1 to 5, multiple choice questions or yes/no

questions. Open ended questions, such as “Any further comments?” are often ineffective as people rarely complete such questions in sufficient detail, so they are better dealt with in an interview.

Performing task analysis activities.

Task analysis involves observing and questioning participants whilst they work. This uncovers the order in which current processes are undertaken and the time devoted to each task. It also provides the analyst with first hand knowledge about the precise data/information, participant roles, technology and information processes operating within the existing system. The system analyst will use this information to determine current needs and also to determine the relative importance of any new needs in the context of the current system.



Consider the following

During an initial interview with a system analyst the manager of a retail store expresses the need to create a series of specific reports so they can improve profits. They know the current system cannot produce these reports, so they are looking for a new system that can. In one report the manager would like details of the total profit for each product sold during each week so they can modify their marketing to maximise profit. This leads the analyst to investigate further. They interview participants, in this case sales staff, to identify the data entered into the system. The analyst finds that two systems are operating; one that records sales and another that records orders for stock. Currently these two systems operate well but independently.



GROUP TASK Discussion

The store manager intends to remove all the existing systems and replace them with a new single system that meets their reporting needs. Assume you are the analyst. How would you respond to the manager? To make an informed decision, what further information is required and how could this information be gathered?

DETERMINING THE PURPOSE AND REQUIREMENTS OF THE NEW SYSTEM

The information gathered from the above activities allows the analyst to formulate a list of needs. Only needs which the client has agreed upon are included. The final list of needs is used to create a statement identifying who the information system is for and what it must achieve. This statement is the purpose of the information system. Commonly the Requirements Reports commences with this statement of the purpose followed by the list of identified needs. In some cases it is appropriate to include a list of needs that will not be addressed by the system. This ensures both the developers and the client are clear about the scope of the project, that is the boundaries for the system are determined to clearly identify what will be included and what will not be included. Often the analyst creates models of the existing system, such as dataflow diagrams and data dictionaries, to assist and inform their efforts.

The needs are refined to create a list of achievable requirements. In general terms, a requirement is a feature, property or behaviour that a system must have. If a system satisfies all its requirements then all the identified needs will be met and the system's purpose will be achieved.



Requirements

Features, properties or behaviours a system must have to achieve its purpose. Each requirement must be verifiable.

It is necessary to verify that all requirements have been met if we are to accurately evaluate the success of the project. For this to occur, all requirements must be expressed in such a way that they can be verified or tested. Consider the statement ‘Customers are to be informed if there will be a delay delivering items’. This is a satisfactory need and may well form part of the system’s purpose, however it is difficult to verify if it has been achieved. It is a subjective statement and is therefore unsuitable as a requirement. Now consider the statement ‘The system shall email customers if an item cannot be delivered within 5 working days of receiving an order’; this statement can easily be tested and is therefore a suitable requirement. In essence it must be possible to design a test which verifies that a requirement has or has not been met.

Some requirements will address existing items that must be used by the new system. For instance, existing participants and their skills, hardware and software that will remain or details of an existing network. Any data accessed from other systems should be specified – a context diagram is often useful. The set of requirements should address everything the system must do and everything the system must use. Notice the words “must do” and “must use”; the requirements do not specify details of how the requirements are to be achieved and they don’t specify items unless they must be included. It is important that the requirements do not restrict the range and nature of possible solutions unless it is unavoidable. This particularly important for large complex systems where the Requirements Report will be used to obtain quotations and possible solutions from a number of different IT providers.

The Requirements Report should be expressed in such a way that it is understandable to the client and also useful as a technical specification for the new system’s developers. In most instances these two parties have a very different view of the system, hence it is often appropriate for two different versions of the requirements report to be produced. Each version contains the same content organised into a form that meets the specific needs of each party. In essence the Requirements Report forms a communication interface between the client and the system’s technical developers. Ensuring each party understands the Requirements Report is absolutely essential as all subsequent stages of the SDLC rely on its content.



GROUP TASK Discussion

The Requirements Report is a particularly critical document when developing systems for larger organisations or when a team of developers will be used. Identify reasons why this is true.

PLANNING

In this, the second stage of the system development cycle, the aim is to determine possible solution options and then make a decision on which option, if any, should be designed and implemented. Feasibility studies are undertaken to make such decisions. Once (and if) a proposed solution has been chosen then the Requirements Report can be updated to reflect the detail of the chosen solution. Finally plans with regard to how the project will be managed can be created.

DETERMINE POSSIBLE SOLUTION OPTIONS

For smaller systems developed by individuals or for individuals, the process of determining possible solution options is often a relatively simple task. Often the time and money available is limited and so too are the skills of the individual. This limits or constrains the range of possible solutions. For example, often existing hardware must be used or developing custom software is too expensive. When an individual is

developing a system for themselves they often already have a solution in mind based on their knowledge and experience. Nevertheless it is still valuable to research alternative solutions prior to committing to a particular path.

For larger systems developed by teams and for larger organisations determining possible solution options is more complex and time consuming. Often proposed solutions are sought from a number of outside IT companies. Each proposal must detail how each of the requirements will be achieved. This includes details of the required participants, data/information, information technology and information processes.



Consider the following

- Jack is a carpenter who wishes to automate his quotation and invoicing system. He has some experience using Microsoft's Excel spreadsheet and the software is already installed on his home computer. Jack intends to develop the system himself using Excel.
- Madge is a professional photographer. She maintains copies of the digital photographs she takes so that customers can order copies and enlargements. Customers often order some years after the event. Actually printing the required photos is a simple task, but locating the correct digital file is taking forever. Currently she has many photos archived on CD-ROMs, others on her portable hard disk drive and others on flash drives. Madge has already created a simple customer database, so she now intends to extend this database to store the photos along with various fields to describe each photo.
- Big R is a chain of 15 retail stores that operates a central warehouse to store and distribute goods using a fleet of trucks. After thorough analysis management has decided to consider updating the warehouse and distribution system. The purpose of the new system is to minimise stock held in the warehouse and minimise the time taken to deliver stock to each store. The warehouse manager previously worked with a particular system and he is intent on implementing this same system within Big R's warehouse.



GROUP TASK Discussion

The above scenarios indicate one proposed solution. Discuss possible problems apparent in each of these proposed solutions and suggest how other alternative possible solutions could be determined.

DECIDE WHICH SOLUTION OPTION (IF ANY) TO DEVELOP

The ability of each possible solution to meet all the requirements must be assessed fairly – the Requirements Report plays a major role in this regard. Without a common and complete set of requirements it would be impossible to make a fair comparison between different solution options. However what if a number of solutions are able to meet the requirements, then on what basis can a decision be made? Feasibility studies are largely concerned with addressing criteria upon which the answer to this question is based.



Feasible

Capable of being achieved using the available resources and meeting the identified requirements.

So what is a feasibility study? To answer this question requires an understanding of the word ‘feasibility’. Consider making some large purchase – say a new car, a new computer or some new piece of furniture. Prior to making such a purchase you ask yourself various questions. What kind do I want? What features do I want? Will it do what I need it to do? What will it cost and can I afford it? Will it require maintenance and what will that cost? And finally should I actually buy it? In essence you are performing an informal mini-feasibility study. Asking and answering similar questions is the essence of all feasibility studies. The ultimate aim is to determine the feasibility of each possible solution and then recommend the most suitable solution. Remember it is possible, and reasonably common for no feasible solution to be recommended, meaning the existing system will remain.

Most feasibility studies consider the following four areas:

- **Technical feasibility.** Is the information technology (hardware and software) available? Will the information technology work with existing technologies? Do the participants possess the required technical skills?
- **Economic feasibility.** Will the new system be cost effective? How long will it take for the cost of the new system to be recovered as a result of increased profits? Could the money invested in the new system be more effectively used elsewhere?
- **Schedule feasibility.** Can the solution be completed on time? What are the consequences if it is not completed on time? Are strict deadlines required and if so how will they be enforced? What training is needed, how long will it take and how will existing duties be performed whilst training occurs?
- **Operational feasibility.** Will the system work in practice? Are management and employees in favour of the new system? Will ongoing support and training be available in the future? Will the system operate well with existing systems?



GROUP TASK Discussion

Identify from whom and explain how answers to the above questions could be determined.

UPDATE REQUIREMENTS REPORT TO REFLECT THE CHOSEN SOLUTION

Once a particular solution has been selected additions to the Requirements Report can be made to include more specific detail. Areas particular to the chosen solution will likely include:

- Detailing the participants who will use the new system, including their roles, current skills, and any new skills that will be required. Perhaps new personnel will be needed or some people’s role will change.
- Information technology should be specified This includes existing hardware and software that will be used and also and new items. Possible suppliers can then be identified and quotations obtained.
- Examples or detailed specifications of the data required and the information the system will create. If data is sourced from another existing system then samples can be obtained for use during the design and testing stages. Depending on the nature of the system, samples of reports and other information output from the system are often valuable.
- An outline of the information processes that will form the solution, perhaps as a high level dataflow diagram. The design of each information process is part of the designing stage, but identifying the general nature of the information processes is useful to help identify the development tasks needed to produce the solution.

DETERMINE HOW THE PROJECT WILL BE MANAGED

Project management aims to ensure the system development lifecycle results in a system that achieves its purpose on time and within budget. It is the process of balancing the allocation of time and money so that development tasks achieve the requirements effectively and efficiently.

During the planning stage it is critical to plan how the project will be managed. Without a plan, time and budget overruns are likely and furthermore the problems leading to such overruns are difficult to detect until it is too late. Lack of planning is a major reason for project failure, indeed poor planning can lead to projects being abandoned altogether. Project management plans must recognise that virtually all projects encounter problems at some stage. As a consequence project management is not a static one off process; rather it is ongoing throughout project development and it must adapt and change to reallocate tasks, money and time in an effort to overcome problems. For larger projects a professional project manager is appointed.

General areas that require consideration when planning how a project will be managed include scheduling, funding and communication.

- Scheduling of development tasks, including techniques for monitoring the completion of tasks. Gantt charts, journals and diaries are common tools for planning and monitoring the progress of development tasks. Often a journal is shared amongst the development team and may require subtask to be signed off as they are completed. Space for comments with regard to any issues are also included. Diaries provide a more detailed view of the work as it progresses. Often individuals maintain their own diary to document work to be done, work that is complete and any issues encountered.

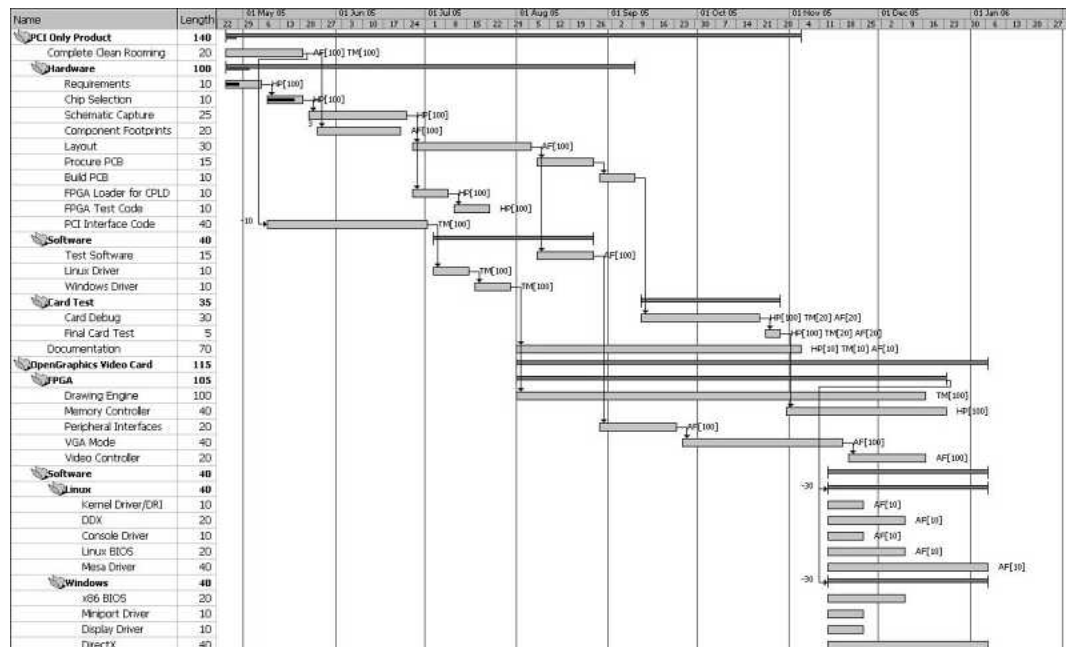


Fig 10.2

Sample Gantt chart for an Open Source Graphics Card's development.

Gantt charts are horizontal bar charts used to graphically schedule and track individual tasks. Fig 10.2 shows a sample Gantt chart used to track the progress of the development of an Open Source Graphics Card. Each horizontal bar represents an individual task. The length of the bar is the time allocated and the start of the bar indicates when the task is to begin. Most project management software applications are able to create Gantt charts, for example Microsoft Project.



GROUP TASK Practical Activity

Create a simple Gantt chart to describe the sequence of each of the SDLC activities we have examined so far in this chapter.

- How funds will be allocated to tasks, including mechanisms to ensure funds are spent wisely. Commonly a funding management plan is created. This plan details how funds are allocated to tasks, mechanisms to ensure money is spent wisely, who is accountable for each task's budget and procedures for reallocating funds during development.
- Lines of communication are required between development personnel and with the client, users and other stakeholders. Typically a communication plan is produced. The plan will specify the communication mediums to be used (e.g. email, telephone, project journal, etc). It also outlines how team members can obtain answers to questions that may arise (e.g. one team member liaises directly with the client and all other team members must obtain answers through this person). In addition methods for monitoring progress and refining or adding requirements will require communication. Often regular team meetings are scheduled where such issues can be addressed and communicated to all involved.



GROUP TASK Discussion

Clearly for large complex development projects formal project management strategies are needed. But is it really necessary for smaller projects? Discuss.

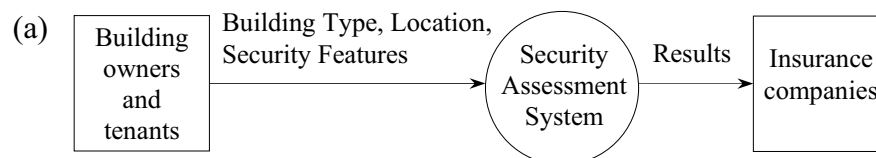


HSC style question:

Development of a new information system is being considered. The system is to assess the impact of various types of security used to protect different types of residential and commercial buildings in various locations. The system aims to provide information to insurance companies to enable them to accurately modify the amount their customers pay for insurance based on the insured building's security features, location and type. The results from the system will be sold to insurance companies who can then download updated information each month.

- Draw a context diagram for the proposed system based on the above description.
- Describe techniques that could be used to determine the precise data required and the information to be created by such a system.

Suggested Solution



- Surveys could be produced and distributed to each insurance company to determine their needs in regard to the information they would find useful. The results of the surveys would then be collated and analysed to determine the format and type of information the system should produce, this being the output from the system. Working backwards would enable the precise data (and processes) to be identified that are needed to generate the required outputs.

SET 10A

1. A person who determines the requirements for a new system is called a:
 - (A) systems analyst.
 - (B) developer.
 - (C) programmer.
 - (D) project manager.
2. An interview is completed:
 - (A) in person.
 - (B) over the phone.
 - (C) on paper.
 - (D) Either (A) or (B).
3. Context diagrams are used when understanding the problem to:
 - (A) define how the requirements will be developed.
 - (B) define the data entering and leaving the existing system.
 - (C) determine all the information processes used by the existing system.
 - (D) identify how the data is transformed into information by the system.
4. Analysing the consequences of not meeting deadlines is part of:
 - (A) technical feasibility.
 - (B) economic feasibility.
 - (C) schedule feasibility.
 - (D) operational feasibility.
5. Which of the following is true of all system requirements?
 - (A) They must be verifiable.
 - (B) They specify information technology.
 - (C) They must describe the behaviour of the system.
 - (D) They must describe a user need.
6. Writing down each step a participant performs to complete a task is part of:
 - (A) developing a participant survey.
 - (B) creating a system model.
 - (C) ensuring tasks are completed.
 - (D) task analysis.
7. A software application that is suitable for use within a new information solution will not be upgraded in the future. This will be considered when examining the solution's:
 - (A) technical feasibility.
 - (B) economic feasibility.
 - (C) schedule feasibility.
 - (D) operational feasibility.
8. Fair comparisons between different solutions are possible because each solution option aims to meet the same:
 - (A) feasibility criteria.
 - (B) requirements.
 - (C) user needs.
 - (D) budgetary constraints.
9. Which of the following is an advantage of direct observation of users over interviews and surveys?
 - (A) Results are easier to compile.
 - (B) Observation requires less time.
 - (C) Observation better determines existing needs.
 - (D) People work efficiently when observed.
10. Gantt charts are project management tools whose primary function is:
 - (A) communication.
 - (B) scheduling.
 - (C) funding.
 - (D) requirements definition.
11. Define each of the following terms.

(a) survey	(c) requirement	(e) feasible
(b) interview	(d) Gantt chart	(f) systems analyst
12. Identify and describe the purpose and essential features of a typical
 - (a) Requirements Report.
 - (b) Feasibility Study
13. Describe strategies and techniques for determining and confirming all the requirements for a new information system.
14. Suggest types of information technology that could be used to assist in the delivery of surveys and interviews. Use specific examples to illustrate your responses.
15. Make a list of the tasks performed during the:
 - (a) Understanding the Problem stage
 - (b) Planning stage

DESIGNING

This third stage of the system development lifecycle (SDLC) is where the actual solution is designed and built. This includes describing the information processes and specifying the system resources required to perform these processes. A diagram such as the “Information System in Context Diagram” in *Fig 10.3* can be used to outline all the system’s components. The resources used by the new information system include the participants, data/information and information technology. Information technology includes all the hardware and software resources used by the system’s information processes. Some new information systems may require completely new hardware and software, whilst others may utilise existing hardware and software to perform new information processes – in fact any combination of new and existing information technology is possible, it depends on the requirements of the new system and the detail of the chosen solution’s information processes.

The design process commences by describing the detail of the system’s information processes. System models are created, such as context diagrams and dataflow diagrams. During the system modelling process, the data and information used and produced by the system is determined and clearly defined using a data dictionary. Once the information processing and data/information is understood the particular detail of the information technology that will perform the processes can be determined. Depending on the individual system, it may be necessary to have new software developed, existing software modified or specific hardware components assembled. Furthermore, specifications and suppliers for required outside communication lines, network cabling, furniture, off-the-shelf software and standard hardware are determined in preparation for their purchase and/or installation. Throughout the entire process consultation with both users and participants should be ongoing. It is essential that the needs and concerns of all people affected by the final operational system remain central to the design process.

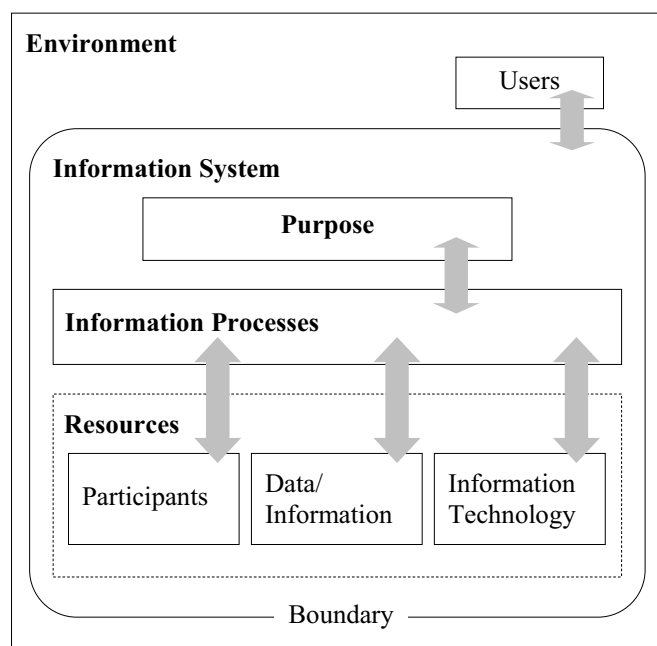


Fig 10.3
Diagrammatic representation of an
information system in context.



GROUP TASK Discussion

Discuss techniques that could be used throughout the design process to ensure user and participant’s needs and concerns are not overlooked.



GROUP TASK Discussion

A variety of different people and organisations are involved in the design stage. This is particularly true for more complex larger systems. Make a list these people and organisations, and outline their role during the design stage of the system development lifecycle.



Consider Just-in-Time Taxis Booking Allocation Information System

A new information system is being developed for Just-in-Time Taxis. Customers' book taxis online and the central administration system allocates bookings efficiently to the taxi closest to the customer. Each taxi will have a PDA-based system installed that utilises Telstra's 3G mobile network to communicate via the Internet and includes a GPS receiver. The GPS functionality allows the location of each taxi to be relayed back to the central administration system where it is used to optimise the allocation of taxis to customers.

The following Information System in Context diagram has been prepared.

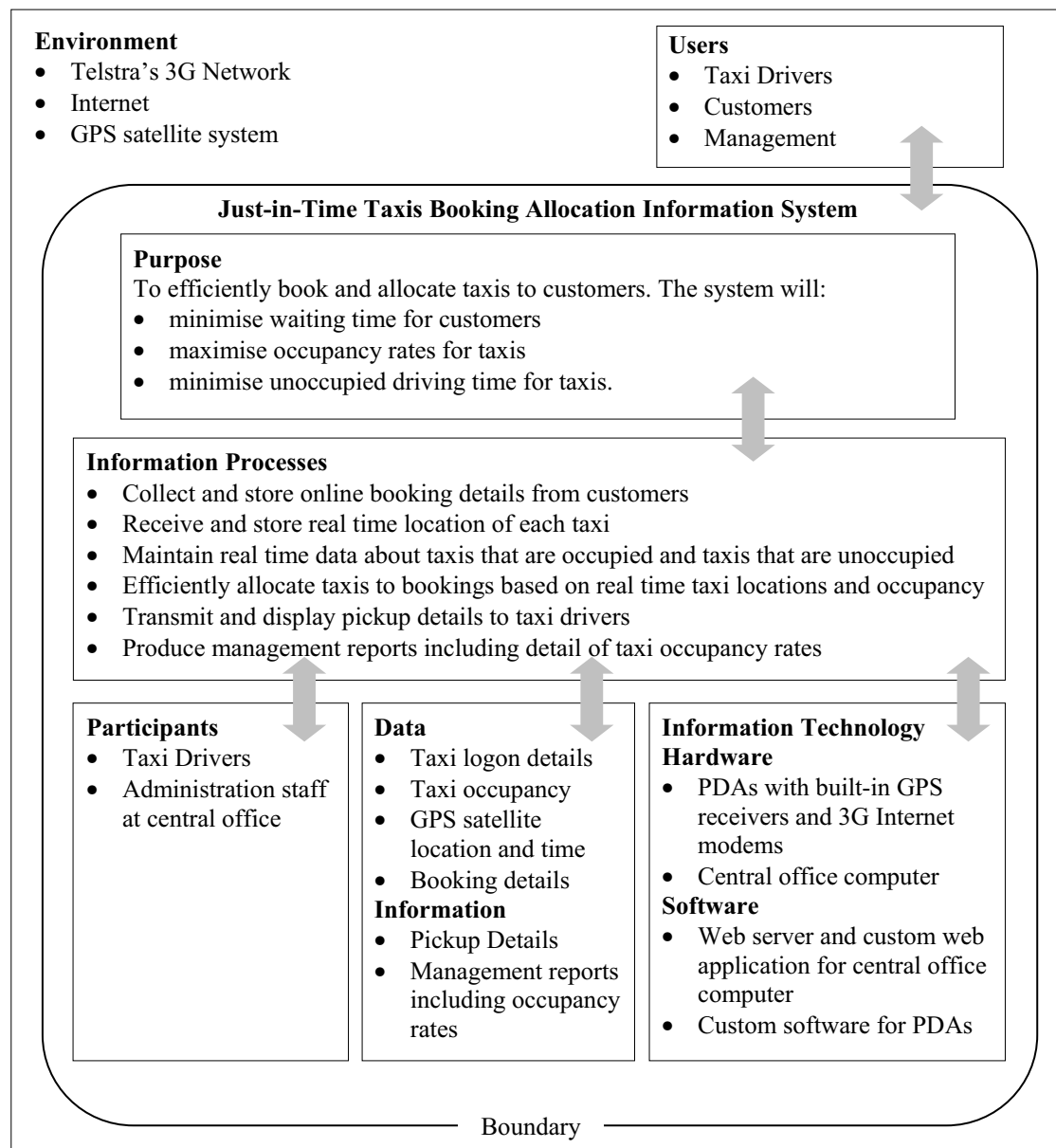


Fig 10.4

Just-in-Time Taxis Information System in Context Diagram



GROUP TASK Discussion

Based on the above diagram, discuss which items will need to be designed for this specific system and which items could be purchased with minor or no modification.

MODELLING INFORMATION PROCESSES AND DATA/INFORMATION

The vital link between all the system's resources is the information processes. Describing the detail of such processes is critical to all aspects of the design – including hardware purchases. As a consequence detailed models of the solution should be produced. We have already examined and used context diagrams, dataflow diagrams and data dictionaries in earlier chapters. Dataflow and context diagrams were formally introduced in Chapter 7 (p269-271) – it would be worthwhile reviewing these pages. In this section we revisit these techniques and emphasise their use as tools to assist in the design of a new system.

Context diagrams

Context diagrams represent the entire new system as a single process. They do not attempt to identify the information processes within the system, rather they identify the data entering and the information leaving the system together with its source and its destination (sink). The sources and sinks are called “external entities”. As is implied by the word external, these entities are outside the system within the system's environment.

So how does a context diagram assist the design process? Context diagrams indicate where the new system interfaces with its environment. They define the data and information that passes through the boundary and in which direction it travels. Descriptions of the data and information is then further detailed within a data dictionary. Ultimately the data entering the system from all its sources must be sufficient to create all the information leaving the system to its sinks.



Consider Just-in-Time Taxis Booking Allocation Information System

The following context diagram has been prepared for Just-in-Time's new system:

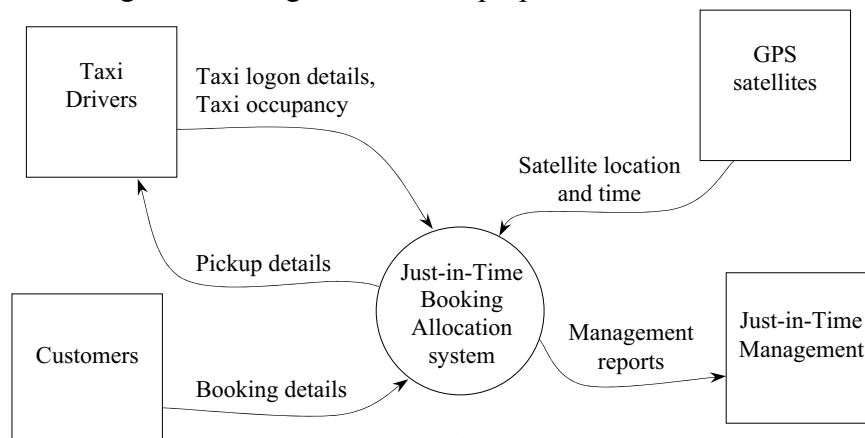


Fig 10.5

Context diagram for Just-in-Time Taxis Booking Allocation Information System



GROUP TASK Discussion

Compare the above context diagram (Fig 10.5) to the information system in context diagram (Fig 10.4). In particular, describe extra information contained on the context diagram.



GROUP TASK Discussion

Research the operation of GPS receivers to explain why “GPS satellites” are included as external entities in Fig 10.5.

Dataflow diagrams (DFDs)

The development of DFDs is a critical step in the design of the solution. The DFDs specify the data entering each information process and the processed data or information leaving each information process. The designer must consider the consequences of their design for the hardware and software that will ultimately perform the processing. For example, a process that involves the transmission of a large file will be inefficient over a slow communication link. Furthermore, if the transfer needs to be complete in order for other information processes to commence then these processes will also be slow.

A series of progressively more detailed dataflow diagrams are created to refine the system into its component sub-processes. Eventually the lowest level DFDs will contain processes that can be solved independently. Breaking down a system's processes into smaller and smaller sub-processes is known as 'top-down design'. The component sub-processes can then be solved and even tested independent of other processes. Once all the sub-processes are solved and working as expected they can be combined to solve the larger problem.



Consider Just-in-Time Taxis Booking Allocation Information System

In this particular information system there are two sub-systems that each performs particular information processes, namely the PDAs in each taxi and the central office computer. Therefore the Level 1 dataflow diagram splits the system into these two sub-systems or processes.

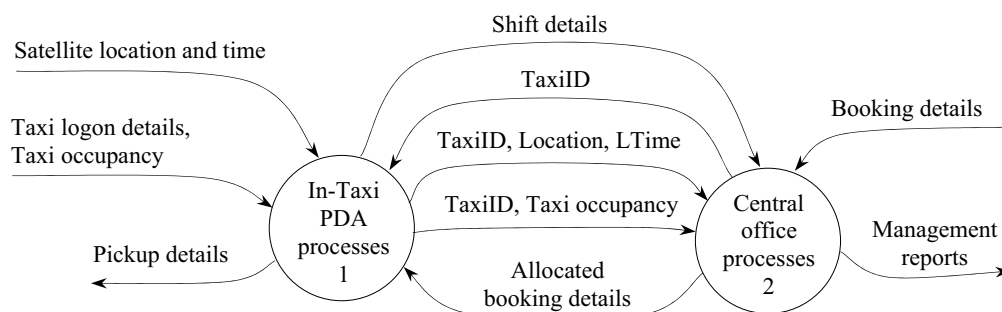


Fig 10.6

Level 1 data flow diagram for Just-in-Time Taxis Booking Allocation System



GROUP TASK Discussion

Consider each of the six general information processes listed on the information system in context diagram in Fig 10.4. Where will each of these processes take place, in the PDAs and/or in the central office?

Data Dictionaries

Data dictionaries are used to detail each of the data items used by the system. They are tables where each row describes a particular data item and each column describes an attribute or detail of the data item. Clearly the name or identifier given to the data item must be included, together with a variety of other details such as its data type, storage size, description and so on.

Data dictionaries are often associated solely with the design of databases where they are used to document details of each field. Commonly such details include at least the field name, data type, data format, field size, description and perhaps an example. However data dictionaries are also used in conjunction with many design tools. For

instance a data dictionary can be used to specify details of each data flow used on context and data flow diagrams. The details specified for each data item should be selected to suit the purpose for which the data dictionary is created. Context diagrams describe an overall view of the system and hence specifying the data type, a description and perhaps an example will likely suffice. When designing a database much more detailed specifications are needed, including the previously mentioned details and possibly other additional detail such as data validation, default value, whether it is key field and so on. Software developers also use data dictionaries to document all the variables and data structures within their code.



Consider Just-in-Time Taxis Booking Allocation Information System

The following data dictionary describes each data flow included on the Level 1 data flow diagram in *Fig 10.6*.

Data Flow Name	Media/Data type	Description
Satellite location and time	Numeric	Data transmitted continuously from GPS satellites
Taxi logon details	Record	Entered by driver to specify PDA password, the vehicle, driver and their shift start and end time.
Taxi occupancy	Boolean	Flag indicating if the taxi has a passenger or not.
Pickup details	Record/Images	Address and GPS navigation to next pickup location which is displayed on the driver's PDA.
Booking details	Record	Includes customer address, destination, number of passengers and desired pickup time.
Management reports	Various	Various reports requested by management, such as details of occupancy rates.
Shift details	Record	Includes ID of driver and vehicle together with shift start and end times.
TaxiID	Integer	Unique identifier for each taxi/driver shift
Location	Integers	Precise latitude and longitude of taxi
LTime	Time	Precise time the taxi's Location was determined
Allocated booking details	Record	Copy of Booking details sent to allocated taxi

Fig 10.7

Data dictionary accompanying Just-in-Times Level 1 data flow diagram.



GROUP TASK Discussion

Analyse the Level 1 DFD (*Fig 10.6*) in conjunction with the above data dictionary (*Fig 10.7*). Does the data dictionary improve your ability to make sense of the DFD? Discuss.



GROUP TASK Activity

With reference to the information processes on the *Fig 10.4* information in context diagram, and using the *Fig 10.6* DFD and data dictionary above, design Level 2 DFDs for the Just-in-Time Taxi Booking Allocation system.

IDENTIFYING AND BUILDING THE INFORMATION TECHNOLOGY

The creation of the system models was about designing the information processes and defining the detail of the data and information used and produced by the system. The information processes, data/information and also the information technology all works together with the participants to achieve the system's purpose. The data is processed into information using the system's information technology – the hardware and software. Hence the hardware and software must be chosen or built to maximise the efficiency of the information processing.

The ability of the hardware and software to perform the system's information processes is of course essential, however there are various other factors that also require consideration. Many of these factors are likely to be specified within the Requirements Report. Some possible questions that address these factors include:

- Is it maintainable? Are there regular upgrades and will these upgrades continue?
- Are spare parts available for hardware now and in the future?
- Is software user friendly? Is it easy to learn? Does it use appropriate terminology?
- Does the user interface behave similarly to other applications known to the users? Can the user interface be modified to suit the user's needs?
- Is software human centred rather than machine centred? Does it enable tasks to be performed according to the user's preferences not the machines?
- Will it meet future needs and expansion? How easily can it be modified or expanded?
- Can it be customised to meet new and emerging requirements?
- Is there a large customer base for the technology? Are their many users proficient in the use of the technology? Do they recommend the technology or do they have complaints?
- Is the technology mature? That is, is it stable with few errors or is it experimental?
- Are the human interface devices ergonomically sound?
- Is the furniture and environment in which users will work ergonomically sound?

The nature of the system and its requirements will determine which of the above questions are relevant. For example, when designing a website the hardware and the furniture used by end-users is beyond the scope (or control) of the system. It is important to confirm answers to each question using sources other than the manufacturer or distributor. Existing customers who have used the technology for some time are often in the best position to confirm claims made by manufacturers and distributors.



GROUP TASK Research

Using the Internet, or otherwise, identify and then recommend PDAs that would be suitable for use within the Just-in-Time's new system. Consider the above dot points as you make your assessment.



GROUP TASK Practical Activity

We considered the basic principles of user interface design in Chapter 3 (p107). Choose either a software application or a web site and analyse its user interface based on the set of principles on page 107.



GROUP TASK Practical Activity

Design screens for collecting the data used by the Just-in-Time Taxi Booking Allocation system.

IMPLEMENTING

This fourth stage of the system development lifecycle (SDLC) is where the new system is installed and finally commences operation. The old system ceases operation and is replaced with the new system. There are various different methods for performing this conversion. However, all these conversion methods require a similar set of tasks to be completed prior to the system commencing operation. These steps include:

1. Installing network cabling and outside communication lines.
2. Purchasing and installing new hardware and software.
3. Configuring the new hardware.
4. Installing, customising and configuring the software.
5. Converting data from the old system to the new.
6. Training the users and participants.

In this section we consider four common methods of implementing or converting from an old system to a new system. We then consider techniques for training users and participants to operate and understand the new system.



GROUP TASK Discussion

Do the 6 steps above need to be completed in the precise order they are listed? Justify your answer using examples relevant to the implementation of the Just-in-Time Taxis Booking Allocation system.

METHODS OF IMPLEMENTATION OR CONVERSION

There are a number of methods of introducing a new system and each of these methods suits different circumstances. Usually implementation of a new system includes converting from an old system, hence these methods of implementation are called methods of conversion.

We consider the following four methods of conversion:

- Direct conversion
- Parallel conversion
- Phased conversion
- Pilot conversion

Direct Conversion

This method involves the old system being completely dropped and the new system being completely implemented at a single point in time. The old system is no longer available. As a consequence, you must be absolutely sure that the new system will operate correctly and meet all its requirements. This conversion method is used when it is not feasible to continue operating two systems together. Any data to be used in the new system must be converted and imported from the old system. Users must be fully trained in the operation of the new system before the conversion takes place.

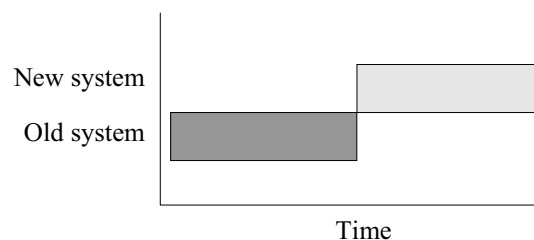


Fig 10.8

Direct conversion method of implementation.

Parallel Conversion

The parallel method of conversion involves operating both the old and new systems together for a period of time. This allows any major problems with the new system to be encountered and corrected without the loss of data. Parallel conversion also means users of the system have time to familiarise themselves fully with the operation of the new system. In essence, the old system remains operational as a backup for the new system. Once the new system is found to be meeting requirements then operation of the old system can cease. The parallel method involves double the workload for users as all tasks must be performed on both the old and the new systems.

Parallel conversion is especially useful when the product is of a crucial nature. That is, dire consequences would result if the new system were to fail. By continuing operation of the old system, the crucial nature of the data is protected.

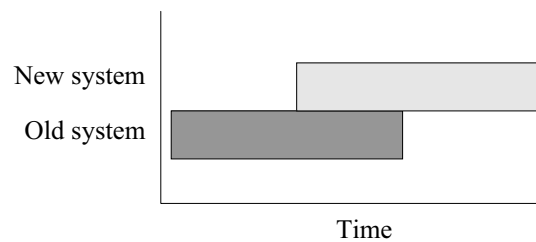


Fig 10.9

Parallel conversion method of implementation.

Phased Conversion

The phased method of converting from an old system to a new system involves a gradual introduction of the new system whilst the old system is progressively discarded. This can be achieved by introducing new parts of the new product one at a time while the older parts being replaced are removed.

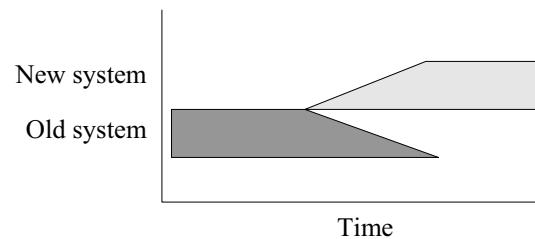


Fig 10.10

Phased conversion method of implementation.

Often phased conversion is used because the system, as a whole, is still under development. Completed sub-systems are released to customers as they become available. Phased conversion can also mean, for large organisations, that the conversion process is more manageable. Parts of the total system are introduced systematically across the business, each part replacing a component of the old system. Over time the complete system will be converted.

Pilot Conversion

With the Pilot method of conversion the new system is installed for a small number of users. These users learn, use and evaluate the new system. Once the new system is deemed to be performing satisfactorily then the system is installed and used by all. This method is particularly useful for systems with a large number of users as it ensures the system is able to operate and perform correctly in a real operational setting. The pilot method also allows a base of users to learn the new system. These users can then assist with the training of others during the systems full implementation. The pilot conversion method can be used as the final acceptance testing of the product. Both the developers and the customer are able to ensure the system meets requirements in an operational environment.

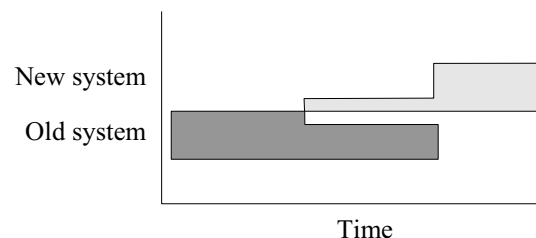


Fig 10.11

Pilot conversion method of implementation.



Consider Just-in-Time Taxis Booking Allocation Information System

Taxis are privately owned and each owner enters into a contract with Just-in-Time Taxis. Although there are many taxis already signed with Just-in-Time who use their existing manual phone based system it is anticipated that many more taxis will be required if the new system is to realise its potential and recover the development costs in a reasonable amount of time.

In addition, some of the existing taxi owners are reluctant to convert their cabs over to the new system. They claim the additional cost of the PDA and the ongoing Internet charges are difficult to justify.



GROUP TASK Discussion

Identify and justify a suitable method of converting the old Just-in-Time system to the new system. Note that it is possible for any combination of conversion techniques to be used.

IMPLEMENTING TRAINING FOR PARTICIPANTS AND USERS

Successful training requires motivated learners. Even the best trainers, using fantastic training techniques and materials will fail if the learners are simply not motivated. For example, nearly all of us complete subjects at school that we are really not enthused about. As a consequence learning in these subjects is an effort. In contrast, even the most unmotivated student is able to learn incredible amounts of information about their favourite hobby or sport. When people are motivated about a subject they actively seek out information, often without prompting. This is not to say that the training methods used are insignificant, rather the point is that motivated learners are vital if the training methods are to be a success.



GROUP TASK Discussion

Choose a subject where some of the class is motivated to learn whilst others are not. Identify reasons for each individual's level of motivation. (Don't choose IPT, as no doubt everyone is highly motivated!)

In regard to new information systems, the learners are the participants and the users. These people are likely to be motivated learners when they:

- are open to change.
- understand how the new system will meet their needs.
- have provided input that has been acted upon during the development of the system.
- have an overall view of the larger system and how their particular tasks will assist in achieving the system's purpose.

These characteristics are achieved through continuous two-way communication throughout the SDLC. For example, if a user has provided an idea during the development process then they should receive feedback regardless of whether the idea has been implemented or not. Indeed feedback on ideas that have not been included is particularly important. Most people will accept rejection if they can see their ideas were considered and that there is a logical reason their ideas were not included.

Let us assume the participants and users are on the whole motivated. We still need to implement some formal training to enable them to commence operating the new system. Some possible training techniques include:

- **Traditional group training sessions**

The trainer can be a member of the system development or an outsourced specialist trainer. If the software has been purchased with little modification then an outsourced training specialist is likely to provide a better service due to their intimate knowledge of the software. If the software has been customised then a member of the development team is perhaps a better choice. In either case the training can be performed onsite or at a separate premises. Onsite group training can often lead to problems as apparently urgent, but unrelated matters, often interrupt the sessions. Off site training allows participants to focus more fully on the training.

- **Peer training**

One or more users undergo intensive training in regard to the operation and skills needed by the new system. These users are also trained in regard to how to train others to use the system. The trained users are then used to train their peers. Peer training is often a one-to-one process. The trained user is a sort of expert who works alongside and assists other users as they learn the skills to operate the new system. This technique allows users to learn skills, as they are required over time.

- **Online training such as tutorials and help systems**

Online tutorials and help systems allow users to learn new skills at their own pace and as they are needed. It is common for larger systems to be provided with a complete tutorial system. Such systems include sample files and databases that can be manipulated and changed without fear of altering or deleting the real data. Many help systems are now context sensitive. This means they display information relevant to the task being completed.

- **Traditional printed manuals**

Printed manuals contain similar information to online tutorial and help systems. However they provide the flexibility to be read away from the computer or to be browsed in a less structured manner. Often procedural information in regard to the operation of a particular system or the completion of particular tasks is documented using printed manuals. Online help is generally specific to particular software rather than to the system. Also, manuals for hardware are usually supplied in printed form.



Consider Just-in-Time Taxis Booking Allocation Information System

Taxi owners sign up to use Just-in-Time's services on a regular basis. Clearly a training strategy is needed to familiarise the new owners with the operation of the system. In addition, taxi owners do not drive their cab 24 hours a day 7 days a week. Rather taxi drivers rent the cab from taxi owners for different shifts. Each of these drivers will also require training – sometimes with very little notice. To further complicate matters, for many taxi drivers English is not their first language.



GROUP TASK Discussion

Propose suitable training strategies that could be used or developed by Just-in-Time to ensure all taxi drivers can use the system and also to ensure the drivers appreciate the advantages of the system.



HSC style question:

Currently the solicitors within a legal firm dictate all their legal letters into small hand held voice recorders. Each solicitor's legal secretary later listens to the recording as they type each letter using an appropriate template. The secretaries save each letter to a network drive using a strict naming and directory system. The network drive is backed up each evening. A draft of each letter is printed and proof read by the solicitor. The secretary then makes any changes, prints the final letter, has it signed and finally posts or faxes the letter.

Each solicitor already has a computer on their desk, hence the firm's owner has decided to dispense with the hand held voice recorders and have the solicitors type their own letters directly. The secretaries will still print the final letters, have them signed and then post or fax the letters.

To implement the new system, the owner of the firm simply states that solicitors will now type their own letters and immediately collects all the hand held voice recorders.

- (a) Identify likely problems that will be encountered.
- (b) Propose a more suitable strategy for implementing the new system that includes training and evaluation.

Suggested Solutions

- (a) Possible problems include:

- Solicitors may save letters to their local drive, which is not backed up. Letters not saved to the backed up network drive can be lost if a storage device fails.
- As no training has occurred the solicitors will not use the current naming/directory system, hence it will be difficult to locate letters for editing and in the future.
- Solicitors may not be familiar with WP software, in particular templates, hence formatting will be incorrect and will not conform to current practices.
- Solicitor must manually inform their secretary each time they type a letter so the secretary can print/post them. Some letters will no doubt be missed.
- Letters are no longer being proofed in draft form, hence more errors are likely to occur in posted letters.
- Legal secretaries were previously typing the letters, which also involved a certain level of proofing. This no longer occurs so more errors are likely.
- As no consultation has occurred some participants will likely feel annoyed with the new arrangements.
- There will be less total work for the legal secretaries; hence job losses or reduction in hours may result.

- (b) Strategies could include:

- Thorough training for solicitors prior to actual conversion with regard to:
 - the file naming and directory system
 - use of network drive and importance of backups
 - use of the word processor and in particular existing templates
 - an accurate system for informing secretaries when a letter is ready for printing
 - proof reading techniques

- There is no need for the direct cut over described in the question. A phased or even pilot conversion would be better suited where some (or all) solicitors use the new system some of the time and assess the effect on work routines, efficiency and accuracy.
- The new system should be formally evaluated once fully operational to ensure cost savings have resulted. It is possible that the opposite will occur as the higher paid solicitors are now spending more time on less highly skilled tasks.

TESTING, EVALUATING AND MAINTAINING

Testing, evaluating and maintaining is the fifth and final stage of the software development lifecycle (SDLC). Unlike the previous stages of the SDLC, aspects of this final stage continue throughout the life of the system.

Tasks included in the testing, evaluating and maintaining stage include:

- acceptance testing to ensure the system meets requirements,
- ongoing evaluation to monitor performance,
- ongoing evaluation to review the effect on users and participants, and
- maintaining the system to ensure it continues to meet requirements.

Problems identified during any of the above tasks will require modifications to the system. For each modification the SDLC commences again. Even if the modification is relatively minor each stage of the SDLC should be completed. This is necessary to ensure the modification works correctly with all parts of the existing system and also to ensure all documentation continues to reflect the current operational system.



GROUP TASK Discussion

Testing and evaluation occurs throughout all stages of the SDLC. Identify examples of testing and evaluation used during each preceding stage.

ACCEPTANCE TESTING TO ENSURE THE SYSTEM MEETS REQUIREMENTS

The testing, evaluating and maintaining stage of the SDLC commences with formal testing of the system to ensure it meets the requirements specified in the Requirements Report – this is known as acceptance testing. Once the tests confirm the requirements have been met the system is signed off as complete. The client and the system developers agree to use the results of the acceptance tests as the basis for determining completion of the new system. If the tests are successful then the client makes their final payment and the system analyst and developer's jobs are complete.



Acceptance Tests

Formal tests conducted to verify whether or not a system meets its requirements.

Acceptance testing enables the client to determine whether or not to accept the new system.

For large-scale information systems acceptance testing is best performed by an outside specialist testing organisation. Even for smaller systems it is preferable for acceptance tests to be performed by people who were not involved in the system's development. People involved in the system development process are likely to be biased. They have designed and implemented the new system, so clearly they will feel the requirements have been met. Furthermore they will, unsurprisingly, view their particular solution as superior to other possibilities.

Although using outside testers is preferable, it is not unusual for the client to perform their own acceptance tests prior to finally accepting and signing off the new system. This is understandable, given that all systems are ultimately developed to meet the needs of clients. Unfortunately disagreement between the clients view of an acceptable system can differ from the views of the developers. It is preferable to agree on the precise nature of the testing and who will perform the tests early in the SDLC – in terms of our highly structured SDLC this should occur during the creation of the Requirements Report. This can easily become a significant problem with less structured development approaches.

ONGOING EVALUATION TO MONITOR PERFORMANCE

There are two essential factors to consider in regard to monitoring the performance of a system. Performance can be monitored from a technical viewpoint – is the system continuing to achieve its requirements? Or the system's performance can be monitored from a financial viewpoint – is the system resulting in improved profits? Each of these factors requires ongoing examination to determine the extent to which the system is meeting expectations. This is the process of evaluation.



Evaluation

The process of examining a system to determine the extent to which it is meeting its requirements.

Technical performance monitoring

Technical performance monitoring aims to evaluate the continuing achievement of the systems evolving requirements. Notice we say 'evolving' requirements. Some old requirements may go down in priority over time or even become irrelevant. Other totally new requirements will emerge and existing requirements will change. This is the nature of virtually all information systems – they change over time. Ongoing evaluation of technical performance aims to verify that requirements continue to be met and identify any changes that may require modifications to the system.



Consider the following:

Some common issues uncovered when performing ongoing technical system evaluation relate to the following factors:

- As the amount of data in the system grows, storage and retrieval processes slow. For instance, when we first purchase a new computer it seems even large video files can be accessed almost instantly, over time the hard drive fills and access slows markedly.
- As the number of transactions increase, response times decrease. For example, making a withdrawal from a bank is fast at 10am in the morning, however at 4pm on a Friday afternoon transactions are intolerably slow.
- As users gain more experience their tolerance of poor performance and usability issues decreases. In other words 'familiarity breeds contempt'. For example, a user interface that generates a simple warning message after each new record is added may be acceptable and even useful to new or irregular users. When entering large quantities of data, experienced users will find responding to such messages hundreds of times a day very irritating.



GROUP TASK Discussion

One example is given for each of the above dot point. Identify and describe further examples of each dot point.

Financial performance monitoring

During the 'Planning' stage of the SDLC a feasibility study was undertaken. This study included analysis of the system's economic feasibility. Financial performance monitoring is largely about evaluating the accuracy of the real economic situation against the economic predictions made in the feasibility study. The aim being to evaluate the extent to which the new system is achieving its economic goals.

Data collected during the evaluation should therefore be sufficient to produce accurate comparisons with the expected results within the feasibility study. Consider the graph in *Fig 10.12*, it shows the results of an original break-even analysis compared to the actual situation for a particular project. A simple analysis of this graph indicates that the project ran slightly over budget when it first became operational some 4½ years ago. Despite this the system managed to reach its break-even point a month prior to expectations. Furthermore, according to the graph, the system has failed to realise its expected economic potential over the last 12 months. Although all of the preceding comments are true of the graph, they are not necessarily true of the system. Perhaps a new competitor entered the market a year ago? Maybe 2 years ago there was a major recession? Environmental factors such as these should be considered when performing financial performance monitoring on an information system.

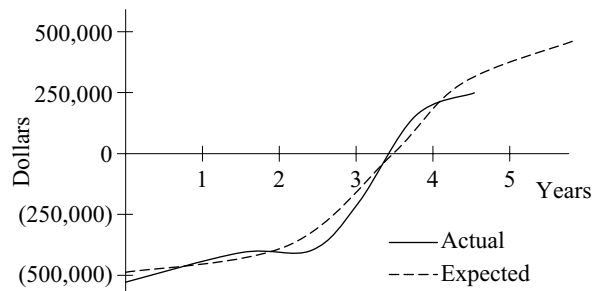


Fig 10.12

Business performance monitoring evaluates actual compared to expected performance.

ONGOING EVALUATION TO REVIEW THE EFFECT ON USERS AND PARTICIPANTS

Have you ever participated in market research, been interviewed about a product or service, or completed a survey? If so then it is likely you were part of an ongoing user evaluation. Similar techniques can be used to gauge the effect of information systems on their users and participants. Users and participants are the most critical elements of an information system. If these people are positive about the system then it more than likely to be a success, however the opposite is also true.

Following is a list of some of the effects of information systems on users and participants. Those that have already been discussed in the text include the relevant page numbers. The remaining three items are discussed following the list. All these items are worth considering when creating evaluation tools.

- Decreased privacy (p18, 113)
- Unwanted changes in the type of employment (p23)
- Unwanted changes in the way work is undertaken (p24)
- Health and safety concerns (p27)
- Ergonomic concerns (p121)
- Poor user interfaces (p107)

- Centralised processing can result in disempowerment of people (p275)
- Unsuitable displays (p374)
- Little or no sense of accomplishment
- Deskilling
- Loss of social contact

All people need to feel a sense of accomplishment. There should be a well-defined purpose to every task they perform. Also, each task should have a distinct start and end point. For example, it is most demoralising to work within a system where a single task is continuous, extra work is always present and no end is ever in sight. This is an extreme example however unfortunately many existing information systems do indeed include such tasks. Evaluation should identify such occurrences so that modifications can be made.

Deskilling can occur when the information system performs processes that were once performed by participants. For example, when desktop publishing software revolutionised the printing industry the trade of “typesetting” changed almost overnight. All the existing typesetting skills required to manually set lead type were no longer needed. These workers had to either leave the industry or retrain to use the new software. Deskilling can also occur when an information system restricts participants to particular tasks and excludes them from others.

Loss of social contact is becoming a common issue. Efficient communication systems allow more and more people to work from home. There is no doubt that this has many advantages, however people are social creatures and they need to develop and maintain relationships with each other. Loss of social contact can also occur when an information system requires participants, particularly those involved in data entry, to spend long periods of time at a computer.



Consider Just-in-Time Taxis Booking Allocation Information System

Each taxi is privately owned and the owner enters into a contract with Just-in-Time Taxis to provide them with a minimum number of bookings per week in exchange for a set weekly fee. A significant number of the taxi owners claim that since introducing the online system the number of bookings they receive has been less than anticipated. Just-in-Time Taxis confirms this to be true and states that this is a consequence of maximising the speed at which taxis fulfil bookings. According to Just-in-Time, some taxis are just lucky as they are more often in the right place at the right time.



GROUP TASK Discussion

Identify and distinguish between the needs of the taxi owners and the needs of Just-in-Time Taxis. Propose modifications that may improve the ability of the system to better meet the needs of both Just-in-Time and the taxi owners.



GROUP TASK Discussion

List and describe evaluation techniques that could be used to identify the effects of a new system on its users and participants.

MAINTAINING THE SYSTEM TO ENSURE IT CONTINUES TO MEET REQUIREMENTS

Information systems require regular maintenance if they are to continue to meet their requirements. In this regard information systems are just like any other system. For example, a car requires regular servicing if it is to continue to function correctly. However even cars that have been serviced according to the manufacturer's specifications do break down. It is the same with information systems. Therefore maintaining an information system involves:

1. regular maintenance, and
2. repairs when faults occur.

Let us briefly consider typical maintenance tasks performed during the operation of an information system.

- Maintaining a hardware and software inventory. An inventory is a detailed list of all the hardware, software and any other equipment used by the system. It should include where each item is located, when it was purchased and how much it cost.
- Perform backups of the system's data and ensure these backup copies are secured in a safe location. Restore data from backups should a fault occur.
- Protect against viruses by ensuring virus protection software is used and updated. If a virus is detected then initiate processes to remove the virus and protect the rest of the system from infection.
- Ensure illegal software is not installed and that all required software is correctly licensed. Should unlicensed or illegal software be found it should be removed.
- Maintain hardware by carrying out all recommended cleaning and other maintenance tasks.
- Ensure stock of all required consumables is at hand. Consumables include printer toner cartridges, disks, recordable CDs and tapes.
- Install and configure replacement or additional hardware and software.
- Setup network access for new users. This includes assigning data access rights together with installing the hardware.
- Monitoring the use of peripheral devices.
- Purchasing and replacing faulty hardware components as problems occur.
- Ensuring new users receive training in regard to the operation of the new system.



GROUP TASK Discussion

Consider Just-in-Time Taxis Booking Allocation system. Are all the above dot points relevant to the maintenance of this new system? Use examples from the Just-in-Time system to justify your responses.



GROUP TASK Discussion

Consider the maintenance of your home or school computer system. What steps do you take to address each of the above dot points? Discuss.

SET 10B

1. Data dictionaries are used to:
 - (A) describe the detail of the system's information processing.
 - (B) determine the information technology required by the system.
 - (C) define the data and information used and produced by the system.
 - (D) ensure participant's needs remain central to the design process.
2. Which of the following is true of all external entities?
 - (A) They are part of the system.
 - (B) They provide data to the system.
 - (C) They receive information from the system.
 - (D) Both (B) and/or (C).
3. Progressively breaking a system's processes into more and more detailed sub-processes is known as:
 - (A) system modelling.
 - (B) bottom-up design.
 - (C) top-down design.
 - (D) system design.
4. What is the function of information technology within an information system?
 - (A) To secure the system's data.
 - (B) To support the information processes.
 - (C) To interface with participants.
 - (D) To perform the information processes.
5. System models aim to describe the detail of the system's:
 - (A) processes and data/information.
 - (B) hardware and software.
 - (C) users, including participants.
 - (D) data and information.
6. New custom software needs to be developed as part of a new information system. When would this occur?
 - (A) Throughout the entire SDLC.
 - (B) During the implementation stage.
 - (C) During the designing stage.
 - (D) Prior to the SDLC commencing.
7. Some users use the system for a while and then they train other users. What is the most likely form of conversion being used?
 - (A) direct conversion
 - (B) parallel conversion
 - (C) phased conversion
 - (D) pilot conversion
8. Why is ongoing evaluation necessary?
 - (A) Because all systems require regular maintenance.
 - (B) To ensure original requirements are met by the new system.
 - (C) To check the system continues to meet its evolving requirements.
 - (D) To correct problems that were incorrectly implemented.
9. Which of the following is the most important if training is to be successful?
 - (A) A motivated trainer.
 - (B) Motivated learners.
 - (C) Motivational training materials.
 - (D) An interruption free training environment.
10. A new system contains four distinct sub-systems. Each sub-system is implemented progressively over a 12 month period. What type of conversion is being used?
 - (A) direct conversion
 - (B) parallel conversion
 - (C) phased conversion
 - (D) pilot conversion
11. Explain how each of the following would be used during the last three stages of the SDLC.
 - (a) Information system in context diagram
 - (b) Context diagram
 - (c) Data flow diagrams
 - (d) Data dictionaries
 - (e) Requirements report
 - (f) Feasibility study

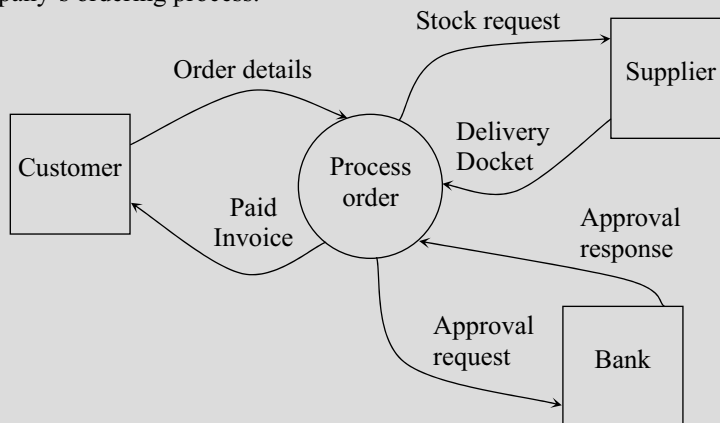
12. Define the following terms.

- (a) Direct conversion
- (b) Parallel conversion
- (c) Phased conversion
- (d) Pilot conversion
- (e) Acceptance testing

13. Identify and justify a suitable method of conversion for each of the following new systems.

- (a) A bank upgrading to a new model ATM throughout Australia.
- (b) A school implementing a new student reports system.
- (c) A company implementing modifications to their website.

14. The context diagram below has been produced to model the flow of data to and from a company's ordering process.



- (a) Describe, in words, the processing and data movements implied by the above context diagram.
- (b) Create a Level 1 dataflow diagram and associated data dictionary to model your written answer to part (a).
- (c) Context diagrams and dataflow diagrams are closely related. Describe this relationship using examples based on the above context diagram to illustrate your answer.

15. Make a list of the tasks performed during the:

- (a) Designing stage
- (b) Implementing stage
- (c) Testing, evaluating and maintaining stage.

SOCIAL AND ETHICAL ISSUES

In this section we examine a number of articles that discuss issues relevant to developers of new information systems. The overriding theme of each article is the push to humanise how information systems operate. Information systems exist to support the information needs of people, therefore people should be in control of the system rather than the system being in control of the people. That is, information systems should be human-centred, not machine-centred.

The ultimate aim of human-centred methodologies is to create systems that are unobtrusive and that respond intuitively to the needs of users and participants. Such systems require less training, less effort to use and furthermore they deliver the information people want. Human-centred systems improve the effectiveness of participants' work and they increase job satisfaction. They allow people to work the way they want to work, not the way the system wants them to work. When designing human-centred systems the important questions to ask concern

Machine-centred systems do the opposite of human-centred systems; they simplify what computers do at the expense of people. Machine-centred systems control the behaviour of users and participants. They force people to follow procedures dictated by the system. Participants must learn how the system works rather than the system operating the way they work.



Consider the following:

The machine-centred mind set

At the Chicago world fair of 1933, the official motto was: "Science Finds - Industry Applies - Man Conforms". To many of us today this seems quite shocking - yet it has been the driving force of much development in the last century.

In particular, if you look at the rise of computing over the last 50 years, you will see that, on the whole, development has been extraordinary, but fairly straightforward: it can be caricatured as trying to make "faster and faster machines fit into smaller and smaller boxes".

Starting from the time of the ENIAC, one of the colossal computers of the 1940s, most IT progress has been driven from the point of view of the machine. Since then things have changed - but perhaps not really that much. Even if computers can today calculate many times over what was possible a few years ago, and the machines have become somewhat less obtrusive, much of the "mind set" has stayed the same. It is the visions of huge calculating machines spanning massive rooms, trying to recreate an absolute artificial intelligence, that still haunt much of the thinking of today.

Clearly, it is difficult to shake off old mind sets.

Fig 10.13

Extract from the foreword of *"Inhabited Information Spaces: Living with your data"*
by D. Snowden, E. Churchill and E. Frécon, 2004.



GROUP TASK Discussion

Do you agree that the 'machine-centred mind set' has not changed that much? Debate both sides.

Human Centred Systems

The history of IT systems development is plagued by recurring failure. These failures may sometimes be attributable to failures in the software engineering process, such as functional inaccuracies or failure and lack of robustness, arising from insufficient rigour in the development process or from insufficient system testing. However, it is now well recognised that system failure more often results from an inadequate consideration of the interaction between the IT system and its users, resulting in lack of system 'usability'.

Fig 10.14

Introduction to a course on 'Human-Centred Systems' at Sheffield Hallam University in the UK.



GROUP TASK Discussion

How could a 'lack of system usability' lead to system failure? Provide examples to justify your response.

Introduction to Ambient Intelligence

"Ambient Intelligence" (AmI) refers to a vision of the future information society stemming from the convergence of ubiquitous computing, ubiquitous communication and intelligent user-friendly interfaces as envisaged in the ISTAG-Scenarios of Ambient Intelligence in 2010 (ISTAG 2001). It puts the emphasis on user-friendliness, user-empowerment and support for human interactions.

Information and Communication Technologies-based artefacts and computers would fade into the background. People would be surrounded by intelligent and intuitive interfaces embedded in all kinds of objects. The environment would recognize individuals and their needs and wants, as well as changes in individuals, changes in needs and wants, and changes in the environment. It would respond in a seamless, unobtrusive and often invisible way, nevertheless remaining under the control of humans. Intelligent agents would eventually make decisions that automatically serve a person or notify a person of a need to make a decision or to carry out an action (adapted from SRI *Pervasive Computing* 2001).

In short, computers would conform to and serve the needs of humans rather than require people to conform to computers by learning specific skills and performing lengthy tasks. Interactions between humans and computers would become relaxing and enjoyable without steep learning curves (ISTAG 2001).

The vision of "Ambient Intelligence" as being developed in the ISTAG report is far-reaching and assumes a paradigm shift in computing from machine-centred towards human-centred computing. It argues for placing the user at the centre of future development. Technologies should be designed for people rather than making people adapt to technologies.

Fig 10.15

Introduction to 'Ambient Intelligence in Everyday Life: A function-Oriented Science & Technology Road Mapping Project' by O. DaCosta.



GROUP TASK Discussion

The concept of 'Ambient Intelligence' described above certainly sounds desirable, but will it assist the information systems of real world organisations to better achieve their purpose? Debate both sides.



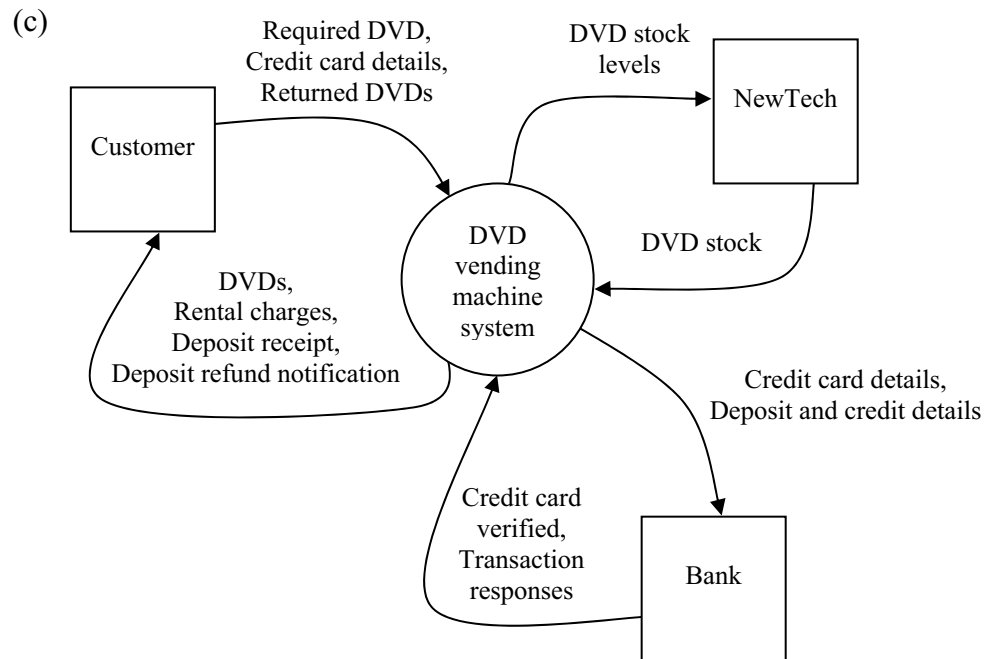
HSC style question:

NewTech Vending machines is currently designing a new vending machine system for DVD rentals. The machines will take the place of traditional video stores and will be installed in service stations. These DVD vending machines will look a lot like traditional soft drink vending machines with the following significant differences:

- Users must swipe a credit card. The credit card is verified and \$100 dollars is credited as a deposit to NewTech's account. The deposit, less the rental fee, is credited back to the customer's credit card account when the DVD is returned to the machine.
 - Upon return of a DVD the vending machine reads security data on the DVD to verify the disk's identity and to ensure it is not damaged. If a problem is encountered then the deposit is not refunded, the disk is ejected and the customer must phone NewTech directly to resolve the issue.
 - Each vending machine electronically reports its DVD stock levels back to NewTech on an hourly basis. A van is dispatched to replenish machines that report low stock levels.
- (a) Would you describe these DVD vending machines as human-centred or machine-centred information systems? Justify your response.
 - (b) List the data collection devices present in one of these DVD vending machines. Identify the data collected by each device you list.
 - (c) Construct a context diagram for the DVD vending machine system.
 - (d) Discuss social and ethical issues that are likely to result should these DVD vending machines prove to be popular.

Suggested Solutions

- (a) Machine-centred. The vending machine directs the precise order in which processes occur and the customer (user) has little or no control over this order. Presumably the user enters a code corresponding to their desired DVD and swipes their credit card. The machine responds by verifying the credit card and dispensing the DVD. Similarly when returning a DVD, the machine accepts the disk, verifies it and then refunds the deposit less the rental fee. The machine directs the user at all times and hence it is a machine centred system.
- (b) Collection devices and data collected includes:
 - Keypad – for collecting DVD codes from the customer in order to identify the desired movie.
 - Magnetic stripe reader – for collecting the credit card details from the customer.
 - DVD drive – to read returned DVDs to verify their identity and that they are not damaged.



(d) Some possible social and ethical issues include:

- Some traditional video stores may have to close as the new vending machines become popular. As a consequence, video store employees will lose their jobs.
- Only customers who have a credit card will be able to rent DVDs. Even if the system is extended to allow accept other types of cards, those without bank accounts are still excluded.
- How will returns be processed if and when a machine breaks down? Perhaps the customer will be held responsible and be expected to travel to a machine that is still operational.
- Many credit cards do not have associated PINs, rather a signature is used. How does the system protect against fraudulent credit card usage? That is, how can it detect stolen credit cards?
- What if a customer finds a DVD to be defective? The system currently assumes damage to DVDs is always caused by the customer and the system keeps their deposit.
- Such a system could improve the availability of DVD rentals. Machines can be installed in remote areas. Also they can be used after hours and on public holidays when traditional stores are closed.

CHAPTER 10 REVIEW

1. The IPT syllabus version of the SDLC contains:
 - (A) 3 stages.
 - (B) 4 stages.
 - (C) 5 stages.
 - (D) 6 stages.
2. The purpose of the SDLC is to:
 - (A) plan, design and implement software applications.
 - (B) analyse existing systems.
 - (C) analyse possible solutions to problems.
 - (D) plan, design and implement systems.
3. Process 1 completes and then at some later time Process 2 uses the data created by Process 1. What information processes are needed for this to occur?
 - (A) collecting
 - (B) storing and retrieving
 - (C) transmitting and receiving
 - (D) displaying
4. System models or diagrams are primarily used to:
 - (A) describe the system's information processes and data/information.
 - (B) explain and justify how the requirements will be met.
 - (C) provide detailed information to software developers.
 - (D) specify the interactions with users in order to develop training materials.
5. The cost of training participants so they possess the necessary technical skills to use a system would be part of assessing:
 - (A) technical feasibility.
 - (B) economic feasibility.
 - (C) schedule feasibility.
 - (D) operational feasibility.
6. What is the essential reason for continuing to evaluate the effect of a system on its users?
 - (A) Because poor interactions between computer systems and people are a primary reason for system failure.
 - (B) To ensure users are aware of why the system operates the way it does.
 - (C) To ensure information technology used by the system is maintained and upgraded.
 - (D) Because users will be the first to identify critical maintenance tasks that require attention.
7. Which of the following is the primary deliverable from the 'planning' stage?
 - (A) Requirements Report
 - (B) Project Plan
 - (C) Feasibility Study Report
 - (D) Models of the existing system
8. Which of the following is largely determined by the needs of the system's users?
 - (A) Information technology
 - (B) Information processes
 - (C) The system's purpose
 - (D) Data/Information.
9. Which of the following is NOT a characteristic of a human-centred system?
 - (A) Increased job satisfaction.
 - (B) Intuitive response to user needs.
 - (C) Flexibility in the way people work.
 - (D) Simpler to design and build.
10. All requirements are verified during which stage of the SDLC?
 - (A) Making decisions
 - (B) Designing solutions.
 - (C) Implementing solutions.
 - (D) Testing, evaluating and maintaining.
11. Describe the tasks performed by each of the following personnel during the SDLC.

(a) systems analyst	(c) client	(e) users, including participants
(b) project manager	(d) programmers	(f) system developers
12. Describe the tasks performed during a feasibility study.
13. Identify and describe techniques that can be used to collect data in order to prepare a Requirements Report.
14. Describe THREE training resources that could be used to assist participants to work with a new information system.
15. New systems should be designed so they are easy to maintain in the future. Propose techniques that could be used during the SDLC to improve the maintainability of new information systems.

GLOSSARY

acceptance testing	Formal tests conducted to verify whether or not a new system meets its requirements.
accumulator	A register within the CPU. It stores the result of the latest computation carried out by the CPU.
active high	A pin or wire whose normal rest state is a low voltage. To activate the function the voltage is raised to high. Compare with active low.
active low	A pin or wire whose normal rest state is a high voltage. To activate the function the voltage is decreased to low. Compare with active high.
ADC	Analog to digital converter.
address bus	Communication lines used to transfer memory locations from the CPU to main memory and the I/O systems. A component of the system bus.
ADSL	Asymmetrical digital subscriber line. A common implementation of DSL.
AGP	Advanced graphics port. A bus standard allowing video cards to directly access main memory independent of the CPU.
ALU	Arithmetic logic unit.
amplitude	The height of a wave. For audio the amplitude determines the volume or level of the sound.
analog	Continuous. Analog data can take any value within its range.
analysing	The information process by which data is interpreted, transforming it into information. The process by which data can be represented and summarised so that humans can better understand it.
application software	Software that performs a specific set of tasks to solve specific types of problems.
ASCII	American Standard Code for Information Interchange.
asymmetrical	Not symmetrical. Communication in each direction occurs, or can occur, at a different speed.
asynchronous	Not in time. Communication that does not attempt to synchronise the sender and receivers clock signals. Also called 'start-stop' communication.
ATA	Advanced technology attachment. A series of standards specifying communication between a drive's controller and the interface on the motherboard.
audit trail	A system that allows the details of any transaction to be traced back to its origin.
backup copy	A copy of files made to protect against the possible loss of the original files. Usually made on a regular basis.
bandwidth	The difference between the highest and lowest frequencies in a transmission channel. Colloquially, bandwidth refers to the speed of transmission.
baud rate	The number of signal events occurring each second. Equivalent to the number of symbols per second.
Bayer filter	A filter used on many CCD based digital cameras. Bayer filters alternate red and green rows with blue and green rows.
Bezier curve	A curve described using a sequence of nodes. Each node contains two points - an anchor point and a control point.

bias	An inclination or preference towards an outcome. Bias unfairly influences the outcome.
BIOS	Basic input output system. Provides a standard way for the operating system to communicate with the system's hardware interfaces and devices.
bit	Binary digit, either 0 or 1.
bitmap image	A method of representing an image as individual picture elements (pixels).
block based encoding	A system for compressing video data.
boundary	The delineation between a system and its environment.
bps	Bits per second. A measurement of the speed of communication.
Braille	A system for displaying text to the blind.
break-even point	The point in time when a new system has paid for itself and begins to make a profit.
broadband	A transmission medium that carries more than one transmission channel. Each channel occupies a distinct range of frequencies.
browser	A software application that interprets HTML code into words, graphics and other elements seen when viewing a web page from a web server.
buffer	A storage area used to assist the movement of data between two devices operating at different speeds.
bus	An intricate network of connections and wires used for communication between devices on the motherboard. Examples include the system bus and external buses.
byte	8 bits.
cable modem	A modem used to connect to a broadband coaxial network. Commonly the network is shared with cable television channels.
cache	A small amount of faster memory that is used to speed up access times to a larger and slower type of memory.
CCD	Charged coupled device.
CCITT	International telegraph and telephone consultative committee. The organisation responsible for maintaining the rules for encoding fax transmissions.
CD-R	Recordable compact disk that can only be written to once.
CD-RW	Rewriteable compact disk.
cell	The intersection of a row and a column within a spreadsheet.
centralised processing	A single computer performing all processing for one or more users.
chipset	A single chip that combines the functions of many different chips. Commonly a chipset on the motherboard combines the controlling circuits for most of the systems hardware interfaces.
CHS	Cylinder, head, sector. A system for addressing each block on a hard disk.
client-server processing	A form of distributed processing where multiple CPUs operate sequentially. The server provides processing resources to the clients.
CMOS	Complementary metal oxide semiconductor. CMOS chips are used to store configuration settings used the BIOS.

CMTS	Cable modem termination system. The device that connects a number of cable modems to an ISP.
CMYK	Cyan, magenta, yellow and key. Key refers to black ink. CMYK is a system for representing colour, also known as four-colour process.
collecting	The information process by which data is entered into or captured by a computer system, including deciding what data is required, how it is sourced and how it is encoded for entry into the system.
context diagram	A systems modelling technique describing the data entering and leaving a system together with its source and sink.
control bus	Communication lines used by the CPU to control the operation of main memory and the I/O systems. A component of the system bus.
copyright	The sole legal right to produce or reproduce a literary, dramatic, musical or artistic work, now extended to include software.
Copyright Act 1968	A legal document used to protect the legal rights of authors of original works.
CPU	Central processing unit.
CRT	Cathode ray tube.
CTS	Carpal tunnel syndrome.
CU	Control unit.
DAC	Digital to analog converter.
data	The raw material used by information processes.
data bus	Communication lines used to transfer data into and out of the CPU. A component of the system bus.
data dictionary	A table identifying and describing the nature of each data item. Data dictionaries are used in many areas of system design, including the design of databases.
data flow	A labelled arrow on context and data flow diagrams describing the nature and direction of data movement.
data integrity	Occurs when data is correct and accurately reflects its source.
data mining	An analysis process that discovers new unintended relationships amongst data.
data quality	Data that is accurate, timely and accessible.
data store	Where data is maintained prior to or after it has been processed. Data stores are represented as open rectangles on data flow diagrams.
data validation	Checks to ensure data is reasonable and meets certain criteria as it is entered. For example HSC marks must be between 0 and 100.
data verification	Checks to ensure data is correct. For example ensuring a customer's address is accurate.
DBMS	Database management system.
decryption	The process of decoding encrypted data using a key.
demodulation	The process of decoding a modulated analog wave back into its original digital signal. The opposite of modulation.
device driver	A program that provides the interface between the operating system and a peripheral device.

DFD	Data flow diagram. A system modelling technique describing the movement of data between processes.
dial-up modem	A modem used to transfer data over a traditional voice telephone line.
digital	Discrete. Digital data is coded and represented as distinct numbers. Computers use binary digital data.
direct conversion	Completely replacing an old system with a complete new system at a particular point in time. Also called direct-cutover.
display adapter	Synonym for video card.
displaying	The information process that controls the format of information presented to the participant or user. The method by which information is output from the system to meet a purpose.
distributed processing	Multiple CPUs used to perform processing tasks, often over a network and transparent to the user.
DMA	Direct memory access. A system that allows devices to communicate directly with main memory without the assistance of the CPU.
DMD	Digital micromirror device.
DMT	Discrete multitone. A modulation standard used by ADSL to dynamically assign frequencies.
DNS	Domain name server. A server that determines the IP address associated with a domain name.
DOCSIS	Data over cable service interface specifications. The standards specifying communication over a cable network.
dot pitch	The width of each pixel in mm. Commonly used to describe the resolution of screens.
dpi	Dots per inch. A measure of screen or printer output resolution.
draw software application	A software application for manipulating vector images.
DSL	Digital subscriber line.
DSLAM	DSL access multiplexor. A device at the telephone exchange that combines multiple signals from ADSL customers onto a single line to ISPs, and extracts individual customer signals from a single line.
DSP	Digital signal processor
DVI	Digital video interface. Used to connect digital monitors to video cards.
Dvorak	A keyboard layout designed to increase typing speeds.
EBCDIC	Extended Binary Coded Decimal Interchange Code.
ECP	Extended capabilities port. A half-duplex parallel port standard.
EFM	Eight to fourteen modulation. A system that converts each byte into 14 fourteen bits such that all bit patterns include at least two but less than 10 consecutive zeros.
email	Electronic mail.
embedding	Importing a source file into a destination file. The source file becomes part of the destination file.

encryption	The process of making data unreadable by those who do not possess the decryption code.
environment	The circumstances and conditions that surround an information system. Everything that influences or is influenced by the system.
ergonomics	The study of the relationship between human workers and their work environment.
ethical	Dealing with morals or the principles of morality. The rules and standards for right conduct or practice.
evaluation	The process of examining a system to determine the extent to which it is meeting its requirements.
external buses	Buses used to transfer data between the system bus and other hardware devices.
external entity	A source or sink for data entering or leaving the system. External entities are not part of the system.
feasible	Capable of being achieved using the available resources and meeting the identified requirements.
fetch-execute cycle	The cycle of events, which a computer carries out to perform each machine code instruction.
fibre optic link	A transmission medium that uses light to represent digital data.
file management software	Software for logically organising files on secondary storage devices.
file server	A computer (including hardware and software) dedicated to the function of storing and retrieving files on a network.
flash memory	Electronic solid-state non-volatile memory.
floating point	A binary system for representing real numbers. Floating point does not represent all numbers exactly.
flow control	A system that controls when data can be transmitted and when it can be received.
font	A specific example of a particular typeface. For example Time New Roman Italic 12 point.
four-colour process	A printing system that uses cyan, magenta, yellow and black dots to form full colour images. Compare with spot colour.
FTP	File transfer protocol. A set of rules for transferring files across a network.
full duplex	Communication in both directions at the same time.
Gantt chart	A project management tool for scheduling and assigning tasks.
GB	Gigabyte.
Gb	Gigabit
GIF	Graphics interchange format.
GLV	Grating light valve.
GPU	Graphics processing unit.
group information system	An information system with a number of participants who work together to achieve the system's purpose.

gutter	Extra margin to allow for binding.
hacker	People who aim to overcome the security mechanisms used by computer systems.
half duplex	Communication in either direction but not at the same time.
handshaking	The process of negotiating and establishing the rules to be used for communication.
hard copy	A copy of text or image based information produced on paper.
hard disk	A random access magnetic secondary storage device. A type of disk in which the platters are solid and the mechanism is sealed inside a container.
hardware	The physical units that make up a computer or any device working with the computer.
heat sink	Commonly an aluminium covering designed to radiate heat away from the CPU.
helical	A type of magnetic tape system where multiple tracks are written at an angle to each other. Helical technology is also used within VCRs.
HID	Human interface device. A standard that forms part of the USB standard. HID drivers are included as part of most operating system.
hot swap	The ability to connect and disconnect devices whilst the system is operating.
HSL	Hue, saturation and luminance. A system for representing colour.
HTML	Hypertext markup language.
HTTP	Hypertext transfer protocol.
hub	A device for connecting nodes on a LAN. Messages are repeated to all attached nodes.
hypertext	Bodies of text that are linked in a non-sequential manner.
I/O	Input/Output.
IDE	Integrated drive electronics. An interface used to transfer data between the system bus and secondary storage devices. A term used to describe storage devices that contain their own controller, rather than it being on the motherboard.
IMAP	Internet message access protocol. A protocol used to download email messages from an email server to an email client.
information	The output displayed by an information system. Knowledge is acquired when information is received.
information processes	What needs to be done to transform the data into useful information. These actions coordinate and direct the system's resources to achieve the system's purpose.
information technology	The hardware and software used by an information system to carry out its information processes.
instruction register	A register within the CPU that holds the next instruction to be executed.
integers	Whole numbers. Includes negative and positive whole numbers and zero.
internal bus	See system bus.
Internet	Global communication network. The Internet is a medium used to connect computers together.
IP	Internet protocol. Each machine on a network, including the Internet, has a unique IP address.

IRQ	Interrupt request line. A direct line from a device to the CPU. Used by device to get the attention of the CPU.
ISP	Internet service provider. A connection point to the Internet. An ISP provides connection to the Internet for many customers.
IX	Internet exchange. Another name for a NAP.
IWB	Interactive Whiteboard. A collection device for delivering presentations to groups of people which often works in conjunction with a projector or large monitor as the display device.
Kb	Kilobit.
KB	Kilobyte.
kerning	Altering the horizontal space between particular character pairs.
LAN	Local area network. A network connecting devices over small physical distances and using the same rules of communication.
laser	Light amplification by stimulated emission of radiation.
LBA	Logical block addressing. An addressing system where each block of data on a hard disk is assigned a sequential number.
LCD	Liquid crystal display.
LCOS	Liquid crystal on silicon.
leading	The distance between lines of text. Measured from the bottom of the descenders on one line to the top of the ascenders on the next line. Pronounced 'ledding' as prior to digital typesetting strips of lead were used.
LED	Light emitting diode.
line spacing	The distance between the bottom of the descender on one line and the bottom of the descender on the next line.
linking	Establishing a connection between a source and destination file. Alterations to the source file will be reflected in the destination file.
liquid crystal	A substance in a state between a liquid and a solid.
MAC address	Media access controller address. A unique address hardwired into NICs and other network devices.
machine language	Instructions that are understood and can be executed by the CPU. Each machine language instruction is part of the CPU's instruction set.
mail-merge	A process where information from a database or other list is inserted into a standard document to produce multiple personalised copies.
Mb	Megabit.
MB	Megabyte.
MEM device	Micro-electromechanical device.
microfiche	A small sheet of clear film onto which a miniature image of each page of a publication has been exposed. A magnifying device (microfiche reader) is used to read microfiche cards.
microwave	High frequency electromagnetic waves that travel in straight lines.
MIDI	Musical Instrument Digital Interface

mirroring	A process performed by various RAID implementations where the same data is simultaneously stored on multiple hard drives. Mirroring improves read access times but not write times.
mixing software application	A software application used to manipulate and combine sampled audio data.
model	A representation of something. Computer models are mathematical representations of systems and objects.
modem	Shortened form of the terms modulation and demodulation. A device whose primary function is to modulate and demodulate signals.
modulation	The process of encoding digital data onto an analog wave by changing its amplitude, frequency and/or phase.
monitor	A dynamic display device.
motherboard	The main printed circuit board in a computer that contains the bus lines. It is equipped with sockets to which all processors, memory modules, plug-in cards, daughterboards, or peripheral devices are connected.
mouse	A mechanical or optical input device used to move a pointer on a screen.
MPEG	Moving Pictures Expert Group.
MR effect	Magneto-resistance effect. A soft magnetic material that conducts electricity well when in the presence of a magnetic field but is otherwise a poor conductor.
NAP	Network access point. A NAP connects many ISPs to high-speed lines to other NAPs. Also called an Internet exchange (IX).
narrowband	A transmission medium that supports a single transmission channel. Compare with broadband.
NIC	Network interface card. The interface between a computer and a LAN.
NPP	National privacy principle. 10 NPPs are contained within the Privacy Act 1988.
NPV	Net present value. A measure of the predicted real cost benefits of an investment.
OCR	Optical character recognition.
optical centre	A point approximately three eighths down a page but horizontally in the centre.
organising	The information process by which data is structured into a form appropriate for the use of other information processes, such as the format in which data will be represented.
OSI model	Open systems interconnection model. A set of standards developed by the International Standards Organisation (ISO). The OSI model is a seven-layer model of communication ranging from the application layer down to the physical layer.
paint software application	A software application for manipulating bitmap images.
parallel conversion	A method of converting to a new system where both the old and new systems operate together for a period of time.
parallel port	A port that transfers bytes of data using 8 parallel wires.
parallel processing	A form of distributed processing where multiple CPUs operate simultaneously to execute a single program or application.

parallel transmission	Method of communication where bits are transferred side by side down multiple communication channels.
participants	A special class of user who carries out (or initiates) the information processes within an information system. Compare with users.
password	A secret code used to confirm that a user is who they claim to be.
PCI	Peripheral component interface. An external bus standard.
PCIe	PCI Express. An external bus standard often used to connect graphics cards.
PDA	Personal digital assistant.
personal information system	An information system with a single participant who is also the sole end user.
phased conversion	A gradual conversion from an old system to a new system.
Piezo crystal	A crystal that expands and contracts as electrical current is altered.
pilot conversion	A small number of users are converted to the new system prior to complete conversion.
pipelining	Multiple instructions being at different stages of execution at the same time.
pixel	Picture element. The smallest element of a bitmap image.
plasma	A state of matter often known as ionised gas.
platter (hard disk)	A single precision aluminium or glass disk within a hard disk.
PnP	Plug and play. A system where permanent registers within a device provide information to the system so the system can automatically allocate the device its required resources.
points	A typesetting measure. There are 72 points per inch.
polarizing panel	A panel that only allows light to enter at a particular angle.
POP	Post office protocol. A protocol used to download email messages from an email server to an email client.
PoP	Point of presence. The devices at an ISP that connect users to the Internet.
privacy	An individual's right to feel safe from observation or intrusion into their personal lives. Consequently individuals have a right to know who holds their personal information and for what purpose it can be used.
Privacy Act 1988	The legal document specifying requirements in regard to the collection and use of personal and sensitive information in Australia.
procedure	The series of steps required to complete a process successfully.
processing	A method by which data can be manipulated in different ways to produce a new value or result (e.g. calculating a total, filtering an email, changing the contrast of an image, changing the volume of a wave file).
program counter	A register within the CPU that holds the address of the next instruction to be executed. In most cases the program counter is incremented to point to the next instruction in memory.

public key encryption	An encryption system where one key (the public key) is used to encrypt the data and a second key (the private key) is used to decrypt the data. Also known as asymmetrical encryption.
punched card	Cards used for both input and output during the 1950s and 1960s.
purpose	A statement identifying who the information system is for and what it needs to achieve. The purpose fulfils the needs of those for whom the system is created (the users).
QAM	Quadrature amplitude modulation. A common modulation technique where the amplitude and phase of the wave are altered.
QWERTY	Popular keyboard layout. Named after the first six characters of the top row.
RAID	Redundant array of independent disks.
RAM	Random access memory.
random access	Data can be stored and retrieved in any order.
raster scan	A technique for drawing or refreshing a screen row by row.
redundant	Repetition exceeding what is necessary.
reflective projector	A projector that reflects light off a smaller reflective image.
refresh rate	The number of times per second that a screen is redrawn.
register	A fast temporary memory location within the CPU and other devices.
requirements	Features, properties or behaviours a system must have to achieve its purpose. Each requirement must be verifiable.
requirements report	The requirements document for a system. A 'blue print' of what the system will do.
RGB	Red, green and blue. A system for representing colour.
ROI	Return on investment. The percentage increase of an investment over time.
router	A device that directs messages to the intended receiver over the most efficient path. Routers can communicate between networks that use different protocols.
RS232	An asynchronous serial standard used by most serial ports.
RSI	Repetitive strain injury.
RTF	Rich text format. A method for organising text data.
sampling (Audio)	The level, or instantaneous amplitude, of an analog audio signal recorded at precise intervals.
sans serif	Without serifs. Refers to a font that does not include serifs.
SAR	Successive approximation register. A component within an ADC that repeatedly produces digital numbers.
SATA	Serial advanced technology attachment. A serial version of the ATA standard.
satellite	A transponder in orbit above the earth.
screen	A dynamic display device.

screen resolution	The number of horizontal pixels by the number of vertical pixels on a screen. Screen resolution can also be measured in dots per inch (dpi) or dot pitch (width of each pixel in mm).
SDLC	System development life cycle. Sometimes abbreviated to SDC.
search	To look through a collection of data in order to locate a required piece of data.
secondary storage	Non-volatile storage. Examples include hard disks, CD-ROMs, DVDs, tapes and floppy disks.
secret key encryption	An encryption system where a single key is used to both encrypt and decrypt data. Also known as symmetrical encryption.
sequential access	Data must be stored and retrieved in a linear manner.
sequential file	Files that can only be accessed from start to finish. Data within a sequential file is stored as a continuous stream.
serial port	A port based on the RS232 standard.
serial transmission	Method of communication where bits are transferred one after the other.
serif	Small strokes present on the extremities of characters in serif typefaces.
simplex	Communication in a single direction only.
simulation	The process of imitating the behaviour of a system or object. A specific application of a model.
sink	An external entity that is the recipient of output from an information system.
SMTP	Simple mail transfer protocol. A protocol used to send email from an email client to an SMTP server and also to transfer email between SMTP servers.
social	Friendly companionship. Living together in harmony rather than isolation.
software	The instructions that control the hardware and direct its operation.
sort	To arrange a collection of items in some specified order.
sound card	A device that converts digital audio to analog and vice versa.
source	An external entity that provides data (input) to an information system.
speech synthesis	The process of producing speech from text using a computer.
spot colour	A printing system that uses one or more inks of a predetermined colour. Compare with four-colour process.
SPP	Standard parallel port. A simplex parallel port standard.
spreadsheet software application	A software application for manipulating numeric data. Spreadsheets combine input, processing and output within a single screen.
SQL	Structured query language.
SSML	Speech synthesis markup language.
start-stop communication	See asynchronous.

stepper motor	A motor that repeatedly turns a precise distance then stops for a precise period of time.
storing and retrieving	The two-step process by which data can be saved and reloaded to allow for other processing to take place, a temporary halt in the system, backup and recovery, and/or the transfer of data or information.
storyboard	A technique that illustrates each screen layout, together with the links between screens.
streaming	The process of delivering data at a constant and continuous rate. Streaming is necessary when delivering audio and video data.
striping	A process performed by various RAID implementations where data is split into chunks and each chunk is simultaneously stored (and retrieved) across multiple hard drives. Striping improves data access times.
switch	An intelligent device for connecting nodes on a LAN. Messages are directed to the intended receiver.
synchronous	Communication where a single clock signal is used to ensure data is received precisely in time with when it was sent.
system	Any organised assembly of resources and processes united and regulated by interaction or interdependence to accomplish a common purpose.
system bus	Communication lines linking the CPU, main memory and the I/O systems. The system bus is composed of a data bus, address bus and control bus. Also known as the internal bus.
system clock	A clock located on the motherboard that provides a constant regular pulse. The system clock is used to synchronise the operation of all devices on the motherboard.
systems analyst	A person who designs and manages the development of information systems.
systems flowchart.	A systems modelling technique describing the logic and flow of data, together with the general nature of the hardware tools.
TCP/IP	Transport control protocol internet protocol. A set of protocols used for communication across networks, including the Internet.
technology	The result of scientific knowledge being applied to practical problems.
TFT	Thin film transistor.
track	On a hard disk each track is a concentric circle on the surface of the disk. Optical disks contain a single continuous spiral track.
tracking	Adjusting the horizontal space between characters evenly within a block of text.
tracking beam	A laser used to ensure the read or write head remains in alignment with the data track on an optical disk.
transmissive projector	A projector that directs light through a smaller transparent image.
transmitting and receiving	The information process that transfers data and information within and between information systems.
transponder	A device that receives and transmits microwaves. A contraction of the words transmitter and responder.
TTS	Text to speech.
tweeter	A speaker designed to reproduce high frequency sound waves.

twos complement	An exact binary system for representing whole numbers or integers.
UART	Universal asynchronous receiver/transmitter. The controller within an RS232 serial port.
UPS	Uninterruptible power supply.
URL	Universal resource locator.
USB	Universal serial bus. A popular serial bus standard where up to 127 peripheral devices share a single communication channel.
user interface	Part of a software application that displays information for the user. The user interface provides the means by which users interact with software.
users	People who view or use the information output from an information system either directly (direct users) or indirectly (indirect users).
vector image	A method of representing images using a mathematical description of each shape.
VGA	Video graphics array (not adapter) supporting resolutions up to 640 by 480 pixels. The plugs that were first used with VGA are now called VGA connectors or adapters and are used extensively to connect analog monitors to video cards.
video card	An interface between the system bus and a screen. It contains its own processing and storage chips. Also called a display adapter.
virus	Software that deliberately produces some undesired or unwanted result.
volatile	In computers, refers to memory that requires power to maintain its data.
VRAM	Video random access memory.
W3C	World wide web consortium.
WAN	Wide area network. A network connecting devices over large physical distances.
woofer	A speaker designed to reproduce low frequency sound waves.
WWW	World wide web.
X-height	The height of the lower case letters that do not have ascenders or descenders.
Xon/Xoff	A software flow control system used by RS232 serial ports.

INDEX

- acceptance testing 405
- accumulator 251
- ADC 92, 96
- address bus 292
- ADSL 308
- AGP 341
- ALU 251
- ambient intelligence 413
- amplitude 137
- analog 54, 67, 305
- analysing 45, 169
- application software 105
- ASCII 55, 84
- asymmetrical 308
- asynchronous 283
- ATA 296
- ATM 7
- audit trail 20

- backup copy 20
- bandwidth 288, 305
- barcode scanner 88
- baud rate 287, 305
- Bayer filter 93
- Bezier curve 134
- bias 118, 275
- BIOS 209
- bit 92
- bitmap image 58
- block based encoding 63
- Blue-ray 72
- boundary 7
- bps 287
- Braille 374
- break-even point 407
- broadband 288
- browser 109
- buffer 227, 354
- bus 292
- byte 92

- cable modem 310
- cache 171, 256
- CCD 89, 92, 98
- CCITT 308
- CD-R 217
- CD-RW 218
- cell 149

- centralised processing 275
- chipset 296
- CHS 209
- CMOS 92, 97
- CMTS 310
- CMYK 132, 147, 351
- collecting 38, 81
- comparator 96
- condenser microphone 94
- context diagram 270, 392, 396
- control bus 294
- copyright 28
- Copyright Act 1968 28
- copyright laws 28
- CPU 170, 292
- CRT 343
- CTS 122
- CU 250
- currency 57

- DAC 96, 353
- data 12
- data bus 292
- data dictionary 155, 231, 398
- data flow 35, 269
- data integrity 106
- data mining 192
- data quality 21, 67
- data store 269
- data validation 21, 106
- data verification 21, 106
- DBMS 154, 185, 230, 266
- decryption 235
- demodulation 305
- desktop publishing 147
- device driver 103, 319
- DFD 35, 269, 397
- dial-up modem 306
- digital 54, 67, 305
- digital camera 92
- direct conversion 400
- display adapter 341
- displaying 50, 339
- distributed processing 275
- DMA 295
- DMD 347
- DMT 309
- DNS 324

- DOCSIS 310
- Dolby surround 62
- dot pitch 345
- downloadng 280
- dpi 345
- draw software applications 133, 262
- DSL 309
- DSLAM 309
- DSP 95, 353
- DVD 72
- DVI 341
- Dvorak 84
- dynamic microphone 94

- EBCDIC 55
- EFM 214
- email 324
- embedding 367
- encryption 20, 235
- environment 6
- ergonomics 27
- ethical 17
- evaluation 406
- external buses 296
- external entity 35, 269

- facsimile 73
- feasible 389
- fibre optic link 315
- file management software 229
- file server 198
- flash file 143
- flash memory 219
- flatbed scanner 89
- floating point 57
- font 361
- four colour process 147
- FTP 322
- full duplex 282

- Gantt chart 391
- GIF 142
- GLV 348
- gutter 362

- hacker 19
- half duplex 282
- handshaking 307
- hard copy 160
- hard disk 171
- hardware 12

- HDMI 341
- heat sink 252
- helical 210
- HID 105, 319
- hot swap 222, 301
- HSL 132
- HTML 109, 157
- hub 312
- human centred 412
- hyperlink 156
- hypertext 156

- I/O 292
- IDE 296
- IMAP 324
- index (database) 155
- information 5, 12
- information processes 9
- information system, data/information 12
- information system, environment 6
- information system, in context 6
- information system, information processes 9
- information system, information technology 12
- information system, participants 10
- information system, purpose 8
- information technology 12
- instruction register 250
- integer 57
- integers 57
- internal bus 292
- Internet 109
- interview 114
- IP 313, 318, 321
- IRQ 295
- ISP 155, 309, 310, 314
- IWB 248
- IX 314

- kerning 363
- kerning 148, 363
- keyboard 82

- LAN 276, 291, 312, 321
- laser 88, 213
- LBA 209
- LCD 342
- LCOS 347
- leading 148, 362
- leading 362
- LED 86, 88

- line spacing 362
- linking 367
- liquid crystal 342
- literature search 113

- MAC address 312
- machine centred 412
- mail-merge 368
- maintaining 409
- MB 171
- MEM device 347
- microfiche 239
- microphone 94
- microwave 315
- MIDI 61
- mirroring 221
- mixing audio 135
- mixing software application 135, 263
- model 179
- modem 305
- modulation 287, 305
- monitor 341
- motherboard 292
- mouse 86
- MPEG 62
- MR effect 206, 207

- NAP 314
- narrowband 209
- national privacy principles 18, 120
- NIC 312
- NPP 18, 120
- numbers 56

- OCR 100
- optical centre 364
- organising 44, 129
- OSI model 317

- paint software application 131, 262
- parallel conversion 401
- parallel processing 254
- parallel transmission 281
- participants 10
- password 20, 234
- PCI 296, 297
- PCIe 341
- PDA 372
- phased conversion 401
- Piezo crystal 352
- pilot conversion 401

- pipelining 254
- pixel 58
- plasma screen 345
- platter (hard disk) 171, 208
- PnP 297
- points 361
- polarizing panel 342
- POP 317, 324
- PoP 314
- presentation software 158
- privacy 18, 120
- Privacy Act 1988 18
- procedure 267
- processing 47, 247
- public key encryption 235
- punched card 378
- purpose 8

- QAM 288, 307, 308
- QWERTY 82

- RAID 221
- RAM 170
- random access 202
- raster scan 344
- real numbers 57
- reflective projector 346
- refresh rate 344
- register 250
- representing 44
- requirements 387
- requirements report 387, 388, 390
- RGB 132, 147
- router 313
- RSI 122
- RTF 145

- sampling (audio) 60
- sans serif 362
- SAR 96
- SATA 296
- satellite 315
- scan code 83
- scanner 88
- screen 341
- screen resolution 344
- SDLC 383
- search 174
- secondary storage 198
- secret key encryption 235
- security of data and information 19

- sequential access 202
- sequential file 202
- serial port 296
- serial transmission 281
- serif 362
- simplex 282
- simulation 179
- sink 35, 269
- slide 159
- SMTP 324
- SMTP 317, 324
- social 17
- social and ethical issues, accuracy of data and information 21
- social and ethical issues, appropriate information use 26
- social and ethical issues, changing nature of work 22
- social and ethical issues, copyright laws 28
- social and ethical issues, health and safety 27
- social and ethical issues, national privacy principles 18
- social and ethical issues, privacy of the individual 18
- social and ethical issues, security of data and information 19
- software 103
- Sony surround 62
- sort 176
- sound card 94, 353
- source 35, 269
- speech synthesis 375
- spot colour 147
- spreadsheet software application 149, 265
- SQL 185
- SSML 376
- start-stop communication 285
- stepper motor 351
- storing and retrieving 46, 197
- storyboard 371
- streaming 326
- striping 221
- structuring 44
- survey 114
- switch 313
- synchronous 283
- system 3
- system, diagrammatic representation of 4
- system bus 292
- system clock 294
- systems analyst 386
- systems flowchart 14, 37
- TCP/IP 318, 321
- technology 12
- telephone system 71
- text 55
- TFT 343
- timeline 140
- track (hard disk) 208
- tracking 363
- tracking 148, 363
- tracking beam 216
- transmissive projector 346
- transmitting and receiving 48, 279
- transponder 315
- trimming 141
- TTS 375
- tweeter 355
- twos complement 57
- uploading 280
- UPS 203
- URL 110
- USB 299
- user interface 107
- users 10, 385
- VCR 74
- vector image 59
- vehicle counting 99
- VGA 341
- VHS 72
- video card 341
- virus 19
- VRAM 293
- W3C 376
- WAN 292, 314
- woofer 355
- word processor 145
- X-height 362

