NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF MANAGEMENT SCIENCES

COURSE CODE:MBA 853

COURSE TITLE:DECISION SUPPORT SYSTEMS (DSS)

COURSE GUIDE

MBA 853
DECISION SUPPORT SYSTEMS (DSS)

Course Team Gerald C. Okereke (Developer/Writer) - ICT

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
No. 5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja
Nigeria

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published By:
National Open University of Nigeria

First Printed 2010

ISBN: 978-058-907-4

CONTENTS **PAGE**

Course Aim

This course is designed for information systems expertsand business managers, generally to understand the application of decision support systems in making timely and accurate business decisions. It is further aimed at giving a correct trend and perspective of DSS in the evolution of information systems. Also, it serves to encourage the usage of decision support systems by business managers. Supportive information systems are also examined from comparative and complimentary perspectives.

Course Objectives

The following are the major objectives of this course.

1. explain the different meanings of information systems
2. explain the concept of information systems and its evolution
1. classify DSS into various forms
2. identify the fields in which DSS are used
• identify different types of decision theory
• describe the applications of the different types of DSS
• compare and contrast the forms of DSS
1. distinguish the concept of e-business and decision support systems
3. describe the major e-business decision support applications that are being customised, personalised, and web-enabled for use in e-business and e-commerce
4. list different ways of applying predictive analysis in decision making in life and business
• identify the applications of spreadsheets as decision making tools
• define the concept of executive information system (EIS)
5. describe the various applications of EIS in sectors of business
6. explain the major indices to identify an effective EIS implementation in organisation
*0identify the major components of GIS
*1outline the steps to take in developing an effective GIS
6.0 define and understand the concept of expert systems
7.0 identify the advantages and disadvantages associated with expert systems
1. define the concept of artificial intelligence
7. identify AI characteristics that are appealing to business functions

Credit Units

This course attracts 3 credit units.

Study Units

The study units of this course are as follows:

Module 1

Unit 1 Introduction to Information Systems
Unit 2 Overview of Decision Support System
Unit 3 Decision Theory and Decision Making

Module 2

Module 3

Course Assessment

- Assignments = 30 %
- Examination = 70%

Course Code MBA 853
Course Title Decision Support Systems (DSS)

Course Gerald C. Okereke (Developer/Writer)

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Office
No. 5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja
Nigeria

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published By:
National Open University of Nigeria

First Printed 2010

ISBN: 978-058-907-4

CONTENTS PAGE

MODULE 1

UNIT 1 INTRODUCTION TO INFORMATION SYSTEMS

CONTENTS

1.0 INTRODUCTION

The term Information System (IS) sometimes refers to a system of persons, data records and activities that process the data and information in an organisation, and it includes the organisation's manual and automated processes. Computer-based information systems are the field of study for information technology, elements of which are sometimes called an "information system" as well, a usage some consider to be incorrect.

In this way, the term "information system" has different meanings:

- In computer security, an information system is described by three objects (Aceituno, 2004):

Structure

- Repositories, which hold data permanently or temporarily, such as buffers, RAM, hard disks, cache, etc.
- Interfaces, which exchange information with the non-digital world, such as keyboards, speakers, scanners, printers, etc.

Channels

- Which connect repositories, such as buses, cables, wireless links, etc. A Network is a set of logical or physical channels.

Behaviour

- Services, which provide value to users or to other services via messages interchange.
- Messages, which carry a meaning to users or services.

- In geography and cartography, a Geographic Information System (GIS) is used to integrate, store, edit, analyse, share, and display geo-referenced information. There are many applications of GIS, ranging from ecology and geology, to the social sciences.
- In knowledge representation, an information system consists of three components: human, technology, and organisation. In this view, information is defined in terms of the three levels of semiotics. Data which can be automatically processed by the application system corresponds to the syntax-level. In the context of an individual who interprets the data they become information, which correspond to the semantic-level. Information becomes knowledge when an individual knows (understands) and evaluates the information (e.g., for a specific task). This corresponds to the pragmatic-level.
- In Mathematics, in the area of domain theory, a Scott information system (after its inventor Dana Scott) is a mathematical 'structure' that provides an alternative representation of Scott domains and, as a special case, algebraic lattices.
- In Mathematics rough set theory, an information system is an attribute-value system.
- In sociology information systems are also social systems whose behaviour is heavily influenced by the goals, values and beliefs of individuals and groups, as well as the performance of the technology.
- In systems theory, an information system is a system, automated or manual, that comprises people, machines, and/or methods organised to collect, process, transmit, and disseminate data that represent user information.
- In telecommunications, an information system is any telecommunications and/or computer related equipment or

interconnected system or subsystems of equipment that is used in the acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of voice and/or data, and includes software, firmware, and hardware.

- In organisational informatics, an information system is a system of communication between people. Information systems are systems involved in the gathering, processing, distribution and use of information and as such support human activity systems.
- The most common view of an information system is one of Input-Process-Output.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define the different meanings of information systems
- describe the concept of information systems and its evolution
- outline the basic structure of information systems
- differentiate the different types of information systems and how they relate to themselves.

3.0 MAIN CONTENT

3.1 History of Information Systems

The study of information systems, originated as a sub-discipline of computer science, in an attempt to understand and rationalise the management of technology within organisations. It has matured into a major field of management that is increasingly being emphasised as an important area of research in management studies, and is taught at all major universities and business schools in the world.

Information technology is a very important malleable resource available to executives. Many companies have created a position of Chief Information Officer (CIO) that sits on the executive board with the Chief Executive Officer (CEO), Chief Financial Officer (CFO), Chief Operating Officer (COO) and Chief Technical Officer (CTO).The CTO may also serve as CIO, and vice versa.

3.2 The Evolution of the Concept of Strategic Information Systems

Since the 1950's, the concept of information systems has evolved from electronic accounting machines speeding paperwork and bureaucracy, to providing general information support in functional areas through the generation of routine performance reports in the 1960's (early Management Information Systems). In the 1970's and beginning of the 80's systems that provided customised information to improve decision making emerged (DSS and ESS). Since the mid-eighties, the role of information systems has taken a different turn. Computing power and networks have placed information systems at the heart of the business. The concept of Strategic Information Systems has gained acceptance, emphasising the fact that information is a strategic resource.

You must not confuse strategic information systems with strategic-level systems like ESS. Strategic information systems can be at any level. They are systems that fundamentally change the organisation itself. So viewed from this perspective, each of the six types of systems, Transaction Processing Systems (TPS), KnowledgeWork Systems (KWS), Office Automation Systems (OAS), Management Information Systems (MIS), Decision Support Systems (DSS), and Executive Support System (ESS) can potentially be regarded as strategic information systems if their impact transforms the goals, business processes, products and external relations of the companyto produce competitive advantage. In fact strategic information systems can change the very nature of the business. In the next section we wil applications of strategic information systems at the business level, firm level and industry level.

3.3 The Structure of Information System

The structure of an information system may be visualised as infrastructure plus applications. The applications have a conceptual structure based on the purpose or needs being met and the functions of the organisations that employ them. The three infrastructures that provide the general capabilities and capacity for information access and processing are technology, data and personnel. The infrastructures enable specific applications and activities.

2. The Technology Infrastructure consists of computer and communication hardware,software, and general purpose software systems. The computer hardware consists of computer and related storage, input and output devices. The communication hardware contains devices to control the flow of communications within internal networks and with external network providers. Computer hardware is made operational through system software and provides generalised functions necessary for applications. Computer operating systems, communication software, and network software is not specific to a single application but provides facilities for many different applications. An example is database management systems to manage databases and perform access and retrieval function for a variety of applications and users.

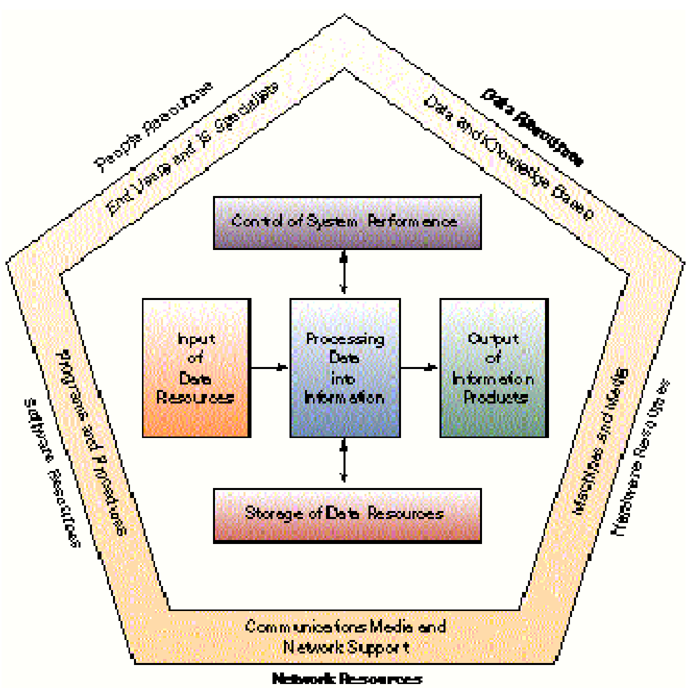


Figure 2: The components and structure of an information system.

3. The Database Infrastructure. The database forms a data infrastructure. They provide for storage of data needed for one or more organisational functions and one or more activities. There will be a number of databases based on organisational activities. Planning of the database infrastructures involves determining what should be stored, what relationships should be maintained among stored data, and what restrictions should be placed on access. The resul of database planning and implementation with database management system is a capacity to provide data both for application and ad hoc needs. Comprehensive databases designed for ad hoc use may be termed data warehouses.

4. **Personnel. The information systems personnel can be viewed as a** third infrastructure, which includes all personnel required to establish and maintain the technology, and the database infrastructures and the capacity to perform user support, development, implementation, operation and maintenance activities. The personnel may be divided between an MIS function and functional areas. There may be , for example, general purpose user support personnel in the MIS function and functional information management support personnel in the functional areas of the organisation.

All information systems use people, hardware, software, data, and network resources to perform input, processing, output, storage, and control activities that transform data resources into information products.

3.4 Types of Information Systems

Different Kinds of Systems

Information systems support different types of decisions at different levels of the organisational hierarchy. While operational managers mostly make structured decisions, senior managers deal with unstructured decisions; middle managers are often faced with semi-structured decisions.

For each functional area in the organisation, three or four levels of organisational hierarchy depending on the author, can be identified: the operational level, knowledge level, management level and strategic level. Each of these levels is served by different types of information systems.

Major Types of Systems

Transaction Processing Systems (TPS) record daily routine transactions such as sales orders from customers, or bank deposits and withdrawals. TPS are vital for the organisation, as they gather all the input necessary for other types of systems. Think about how one could generate a monthly sales report for middle management or critical marketing information to senior managers without TPS. TPS provide the basic input to the company's database. A failure in the TPS often means disaster for the organisation. Imagine what happens when the reservation system at Turkish Airlines fails: all operations stop, no transactions can be carried out until the system is up again. Long queues form in front of ATMs and tellers when a bank's TPS crashes.

Knowledge Work Systems (KWS) support highly skilled knowledge workers in the creation and integration of new knowledge into the company. Computer Aided Design (CAD) systems used by product designers not only allow them to easily make modifications without having to redraw the entire object (just like word processors for documents), but also enable them to test the product without having to build physical prototypes. Three dimensional graphical simulation systems like GRASP (Graphical Robotics Applications Simulation Package) are used by British Aerospace and Rolls Royce for evaluating and programming industrial robots. Architects use CAD software to create, modify, evaluate and test their designs; such systems can generate photo realistic pictures, simulating the lighting in rooms at different times of the day, perform calculations, for instance on the amount of paint required. Surgeons use sophisticated CAD systems to design operations.

Financial institutions are using knowledge work systems to support trading and portfolio management with powerful high-end PC's. These allow managers to get instantaneous analysed results on huge amounts of financial data and provide access to external databases.

Office Automation Systems (OAS) support general office work for handling and managing documents and facilitating communication. Text and image processing systems evolved from word processors to desktop publishing, enabling the creation of professional documents with graphics and special layout features. Spreadsheets, presentation packages like PowerPoint, personal database systems and note-taking systems (appointment book, notepad, cardfile) are part of OAS.

In addition OAS includes communication systems for transmitting messages and documents (e-mail) and teleconferencing capabilities.

Management Information Systems (MIS) generate information for monitoring performance (e.g. productivity information) and maintaining coordination (e.g. between purchasing and accounts payable).

MIS extract process and summarise data from the TPS and provide periodic (weekly, monthly, quarterly) reports to managers.
Today MIS are becoming more flexible by providing access to information whenever needed (rather than pre-specified reports on a periodic basis). Users can often generate more customised reports by

selecting subsets of data (such as listing the products with 2% increase in sales over the past month), using different sorting options (by sales region, by salesperson, by highest volume of sales) and different display choices (graphical, tabular).

Decision Support Systems (DSS) support analytical work in semi-structured or unstructured situations. They enable managers to answer "What if?" questions by providing powerful models and tools (simulation, optimisation) to evaluate alternatives (e.g. evaluating alternative marketing plans).

DSS are user-friendly and highly interactive. Although they use data from the TPS and MIS, they also allow the inclusion of new data, often from external sources, such as current share prices or prices of competitors.

Executive Support Systems (ESS) or Executive Information Systems (EIS) provide a generalised computing and communication environment to senior managers to support strategic decisions. They draw data from the MIS and allow communication with external sources of information. But unlike DSS, they are not designed to use analytical models for specific problem solving. ESS are designed to facilitate senior managers' access to information quickly and effectively.

ESS have menu driven user friendly interfaces, interactive graphics to help visualisation of the situation, and communication capabilities that link the senior executive to the external databases he requires (e.g. Dow Jones News/Retrieval, or the Gallop Poll).

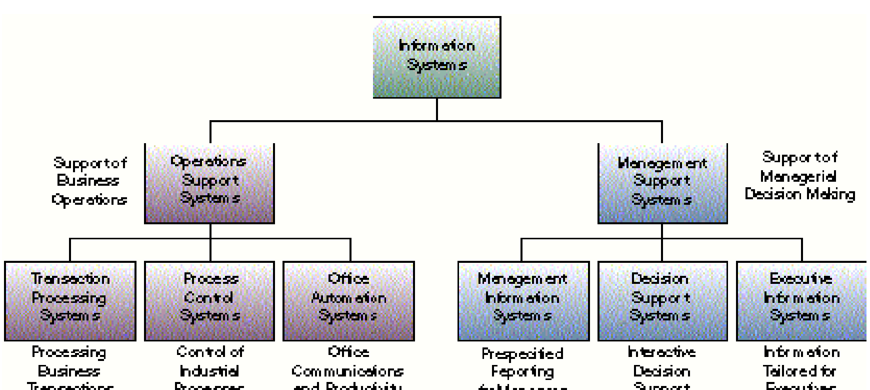


Figure 3: Operations and management classifications of information systems.

3.5 Relationship of Systems to One Another

Different types of systems exist in organisations. Not all organisations have all of the types of systems described here. Many organisations may not have knowledge work systems, executive support systems or decision support systems. But today most organisations make use of office automation systems and have a portfolio of information system applications based on TPS and MIS (marketing systems, manufacturing systems, human resources systems). Some organisations have hybrid information systems that contain some of the characteristics of different types of systems.

The field of information systems is moving so quickly that the features of one particular type of system are integrated to other types (e.g. MIS having many of the features of ESS). System characteristics evolve and new types of systems emerge. Yet the classification of information systems into these different types is useful because each type of system has certain features that are relevant in particular situations.

4.0 CONCLUSION

Though there are different perspectives to the definition of information system, there are basic elements associated with each definition. Also, even as the different types of information systems keep emerging as new technologies emerge by the day, information systems are still basically responsible for making business operations to be effective and productive.

5.0 SUMMARY

This unit is summarised as follows:

- The term Information System (IS) sometimes refers to a system of persons, data records and activities that process the data and information in an organisation, and it includes the organisation's manual and automated processes.
- Since the 1950's, the concept of information systems has evolved from electronic accounting machines, speeding paperwork and bureaucracy, to providing general information support in functional areas through the generation of routine performance reports in the 1960's (early Management Information Systems).
- The structure of an information system may be visualised as infrastructure plus applications. The applications have a conceptual structure based on the purpose or needs being met and the functions of the organisation that employ them.
- Information systems support different types of decisions at different levels of the organisational hierarchy. While operational managers mostly make structured decisions, senior managers deal with

unstructured decisions; middle managers are often faced with semi-structured decisions.

- Different types of systems exist in organisations. Not all organisations have all of the types of systems described here. Many organisations maynot have knowledge work systems, executive support systems or decision support systems.

6.0 TUTOR-MARKED ASSIGNMENT

1. Identify the basic components of information systems
2. Briefly discuss the database infrastructure of information systems

7.0REFERENCES/FURTHER READING

Angell, I.O. and Smithson S. (1991). Information Systems Management: *Opportunities and Risks.*

Beynon-Davies P. (2002). Information Systems: An Introduction to *Informatics in Organisations.Palgrave, Basingstoke, UK. ISBN:* 0-333-96390-3.

Federal Standard 1037C, MIL-STD-188, and National Information Systems Security Glossary.

ION 501; Introduction to Information Systems.

O'Brien J.A.; Introduction to Information Systems, 8th Edition, ISBN: 0-256-20937-5.

Rockart et. al (1996). Eight Imperatives for the New IT Organisation Sloan Management Review.

Sloan Career Cornerstone Center. (2008). Information Systems. Alfred P. Sloan Foundation.

Langefors, Börje (1973). Theoretical Analysis of Information Systems, Auerbach. ISBN 0-87769-151-7.

Ciborra, C. (2002) Labyrinths of Information. Oxford: Oxford University Press.

Master of Science in Information Systems at the University of Münster .

UNIT 2 OVERVIEW OF DECISION SUPPORT SYSTEM

CONTENTS

1.0 INTRODUCTION

Decision Support Systems constitute a class of computer-based information systems including knowledge-based systems that support decision-making activities.

Decision Support Systems (DSS) are a specific class of computerised information system that supports business and organisational decision-making activities. A properly-designed DSS is an interactive software-based system intended to help decision makers compile useful information from raw data, documents, personal knowledge, and/or business models to identify and solve problems and make decisions.

Typical information that a decision support application might gather and present would be:

- an inventory of all of your current information assets (including legacy and relational data sources, cubes, data warehouses, and data marts),
- comparative sales figures between one week and the next,
- projected revenue figures based on new product sales assumptions;
- the consequences of different decision alternatives, given past experience in a context that is described.

2.0 OBJECTIVES

At the end of this unit, you should be able to:

- define decision support system (DSS)
- outline the history and the development process of DSS
- explain what constitutes the architecture of a typical DSS
- classify DSS into various forms
- identify the fields in which DSS are used
- state the benefits associated with the use of DSS
- compare and contrast DSS and MIS

3.0 MAIN CONTENT

3.1 What is a Decision Support System?

A decision support system is a way to model data and make quality decisions based upon it. Making the right decision in business is usually based on the quality of your data and your ability to sift through and analyse the data to find trends in which you can create solutions and strategies for. DSS or decision support systems are usually computer applications along with a human component that can sift through large amounts of data and pick between the many choices.

While many people think of decision support systems as a specialised part of a business, most companies have actually integrated this system into their day to day operating activities. For instance, many companies constantly download and analyse sales data, budget sheets and forecasts and they update their strategy once they analyse and evaluate the current results. Decision support systems have a definite structure in businesses, but in reality, the data and decisions that are based on it are fluid and constantly changing.

The key to decision support systems is to collect data, analyse and shape the data that is collected and then try to make sound decisions or construct strategies from analysis. Whether computers, databases or people are involved usually doesn't matter, however it is this process of taking raw or unstructured data, containing and collecting it and then using it to aid decision making.

Although computers and artificial intelligence are atwork or at play with data, it is ultimately up to humans to executethese strategies or comprehend the data into a usable hypothesis.

It is important to note that the field of DSS does not have a universally accepted model, meaning that there are many theories vying for supremacy in this broad field. Because there are many working theories on the topic of DSS, there are many ways to classify DSS.

For instance, one of the DSS models available is with the relationship of the user in mind. This model takes into consideration passive, active and co-operative DSS models.

Decision support systems that just collect data and organise them effectively are usually called passive models, they do not suggest a specific decision, and they only reveal the data. An active decision support system actually processes data and explicitly shows solutions based upon that data. While there are many systems that are able to be active, many organisations would be hard pressed to put all their faith into a computer model without any human intervention.

A co-operative decision support system is when data is collected, analysed and then provided with a human component which helps the system revise or refine it. It means that both a human component and computer component work together to come up with the best solution.

While the above DSS model takes the relationship of the user in mind, another popular DSS model takes into consideration the mode of assistance as the underlying basis of the DSS model. This includes the Model Driven DSS, Communications Driven DSS, Data Driven DSS, Document Driven DSS, and Knowledge Driven DSS.

Model Driven DSS is when decision makers use statistical, simulations or financial models to come up with a solution or strategy. Keep in mind that these decisions are based on models; however they do not have to be overwhelmingly data intensive.

A Communications Driven DSS model is when many collaborators work together to come up with a series of decisions to set in motion a solution or strategy. This communications driven DSS model can be in an office environment or on the web.

A Data Driven DSS model puts its emphasis on collected data that is then manipulated to fit the decision maker's needs. This data can be internal, external and in a variety of formats. It is important that usually data is collected and categorised as a time series which is a collection of data that forms a sequence, such as daily sales, operating budgets from one quarter to the next, inventory levels over the previous year, etc.

A Document Driven DSS model uses documents in a variety of data types such as text documents, spreadsheets and database records to come up with decisions as well as further manipulate the information to refine strategies.

A Knowledge Driven DSS model uses special rules stored in a computer or used by a human to determine whether a decision should be made. For instance, for many days, traders stop loss limit can be seen as a knowledge driven DSS model. These rules or facts are used in order to make a decision.

You can also look at the scope in which decisions are made as a model of DSS. For instance, an organisational wide decision, department decision or single user decision, can be seen in the scope wide model.

3.2 A Brief History

In the absence of an all-inclusive definition, we focus on the history of DSS. According to Keen, the concept of decision support has evolved from two main areas of research: the theoretical studies of organisational decision making done at the Carnegie Institute of Technology during the late 1950s and early 1960s, and the technical work on interactive computer systems, mainly carried out at the Massachusetts Institute of Technology in the 1960s. It is considered that the concept of DSS became an area of research of its own in the middle of the 1970s, before gaining in intensity during the 1980s. In the middle and late 1980s, executive information systems (EIS), group decision support systems (GDSS), and organisational decision support systems (ODSS) evolved from the single user and model-oriented DSS. Beginning in about 1990, data warehousing and on-line analytical processing (OLAP) began broadening the realm of DSS. As the turn of the millennium approached, new Web-based analytical applications were introduced.

It is clear that DSS belong to an environment with multidisciplinary foundations, including (but not exclusively) database research, artificial intelligence, human-computer interaction, simulation methods, software engineering, and telecommunications.

DSS also have a weak connection to the user interface paradigm of hypertext. Both the University of Vermont PROMIS system (for medical decision making) and the Carnegie Mellon ZOG/KMS system (for military and business decision making) were decision support systems which also were major breakthroughs in user interface research. Furthermore, although hypertext researchers have generally been concerned with information overload, certain researchers, notably Douglas Engelbart, have been focused on decision makers in particular.

3.3 Taxonomies

As with the definition, there is no universally-accepted taxonomy of DSS either. Different authors propose different classifications. Using the relationship with the user as the criterion, Haettenschwiler differentiates *passive, active, and co-operative DSS. A passive DSS is a system that* aids the process of decision making, but that cannot bring out explicit decision suggestions or solutions. An active DSS can bring out such decision suggestions or solutions. A co-operative DSS allows the decision maker (or its advisor) to modify, complete, or refine the decision suggestions provided by the system, before sending them back to the system for validation. The system again improves, completes, and refines the suggestions of the decision maker and sends them back to her for validation. The whole process then starts again, until a consolidated solution is generated.

Using the mode of assistance as the criterion, Power differentiates *communication-driven DSS, data-driven DSS, document-driven DSS, knowledge-driven DSS, and model-driven DSS.*

- A model-driven DSS emphasises access to and manipulation of a statistical, financial, optimisation, or simulation model. Model-driven DSS use data and parameters provided by users to assist decision makers in analysing a situation; they are not necessarily data-intensive. Dicodess is an example of an open source model-driven DSS generator.
- A communication-driven DSS supports more than one person working on a shared task; examples include integrated tools like Microsoft's NetMeeting or Groove.
- A data-driven DSS or data-oriented DSS emphasises access to and manipulation of a time series of internal company data and, sometimes, external data.
- A document-driven DSS manages, retrieves, and manipulates unstructured information in a variety of electronic formats.
- A knowledge-driven DSS provides specialised problem-solving expertise stored as facts, rules, procedures, or in similar structures.

Using scope as the criterion, Power differentiates enterprise-wide DSS and desktop DSS. An enterprise-wide DSS is linked to large data warehouses and serves many managers in the company. A desktop, *single-user DSS is a small system that runs on an individual manager's* PC.

3.4 Architectures

Once again, different authors identify different components in a DSS. For example, Sprague and Carlson identify three fundamental components of DSS: (a) the database management system (DBMS), (b) the model-base management system (MBMS), and (c) the dialogue generation and management system (DGMS).

- Haag et al. describe these three components in more detail:

The Data Management Component stores information (which can be further subdivided into that derived from an organisation's traditional data repositories, from external sources such as the Internet, or from the personal insights and experiencesof individual users); the Model Management Component handles representations of events, facts, or situations (using various kinds of models, two examples being optimisation models and goal-seeking models); and the User Interface Management Component is, of course, the component that allows a user to interact with the system.

- According to Power , academics and practitioners have discussed building DSS in terms of four major components: (a) the user interface, (b) the database, (c) the model and analytical tools, and (d) the DSS architecture and network.
- Hättenschwiler identifies five components of DSS:

(a) users with different roles or functions in the decision making process (decision maker, advisors, domain experts, system experts, data collectors),
(b) a specific and definable decision context,
(c) a target system describing the majority of the preferences,
(d) a knowledge base made of external data sources, knowledge databases, working databases, data warehouses and meta-databases, mathematical models and methods, procedures, inference and search engines, administrative programs, and reporting systems, and
(e) a working environment for the preparation, analysis, and documentation of decision alternatives.

- Arakas proposes a generalised architecture made of five distinct parts:

(a) the data management system,
(b) the model management system,
(c) the knowledge engine,
(d) the user interface, and
(e) the user(s).

3.5 Development Frameworks

DSS systems are not entirely different from other systems and require a structured approach. A framework was provided by Sprague and Watson (1993). The framework has three main levels:

1). Technology levels
2). People involved

3). The developmental approach

1. Technology Levels

Sprague has suggested that there are three levels of hardware and software that has been proposed for DSS.

a) Level 1 – Specific DSS

This is the actual application that will be used by the user. This is the part of the application that allows the decision maker to make decisions in a particular problem area. The user can act upon that particular problem.

b) Level 2 – DSS Generator

This level contains hardware/software environment that allows people to easily develop specific DSS applications. This level makes use of case tools or systems such as Crystal, AIMMS, iThink and Clementine.

c) Level 3 – DSS Tools

Contains lower level hardware/software. DSS generators including special languages, function libraries and linking modules .

2. People Involved

Sprague suggests that there are 5 roles involved in a typical DSS development cycle.

a) The end user
b) An intermediary
c) DSS developer
d) Technical supporter
e) Systems expert

3. The Developmental Approach

The developmentalapproach for a DSS system should be strongly interactive. This will allow for the application to be changed and redesigned at various intervals. The initial problem is used to design the system and then tested and revised to ensure the desired outcome is achieved.

3.6 Classifying DSS

There are several ways to classify DSS applications. Not every DSS fits neatly into one category, but a mix of two or more architecture in one.

Holsapple and Whinston classify DSS into the following six frameworks: Text-oriented DSS, Database-oriented DSS, Spreadsheet-oriented DSS, Solver-oriented DSS, Rule-oriented DSS, and Compound DSS.

A compound DSS is the most popular classification for a DSS. It is a hybrid system that includes two or more of the five basic structures described by Holsapple and Whinston .
The support given by DSS can be separated into three distinct interrelated categories:

Personal Support,
Group Support, and
Organisational Support

Additionally, the build up of a DSS is also classified into a few characteristics.

1) **Inputs: This is used so the DSS can have factors, numbers, and** characteristics to analyse.
2) **User knowledge and Expertise: This allows the system to** decide how much it is relied on, and exactly what inputs must be analysed with or without the user.
3) **Outputs: This is used so the user of the system can analyse the** decisions that may be made and then potentially.
4) **Make a Decision: This decision making is made by the DSS,** however, it is ultimately made by the user in order to decide on which criteria it should use.

DSS which perform selected cognitive decision-making functions and are based on artificial intelligence or intelligent agents' technologies are called Intelligent Decision Support Systems (IDSS)

3.7 Applications

As mentioned above, there are theoretical possibilities of building such systems in any knowledge domain.

One example is the Clinical decisionsupport systemfor medical diagnosis. Other examples include a bank loan officer verifying the credit of a loan applicant or an engineering firm that has bids on several projects and wants to know if they can be competitive with their costs.

DSS is extensively used in business and management. Executive dashboard and other business performance software allow faster decision making, identification of negative trends, and better allocation of business resources.

A growing area of DSS application, concepts, principles, and techniques is in agricultural production, marketing for sustainable development. For example, the DSSAT4 package, developed through financial support of USAID during the 80's and 90's, has allowed rapid assessment of several agricultural production systems around the world to facilitate decision-making at the farm and policy levels. There are, however, many constraints to the successful adoption on DSS in agriculture.

A specific example concerns the Canadian National Railway system, which tests its equipment on a regular basis using a decision support system. A problem faced by any railroad is worn-out or defective rails, which can result in hundreds of derailments per year. Under a DSS, CN managed to decrease the incidence of derailments at the same time other companies were experiencing an increase.

DSS has many applications that have already been spoken about. However, it can be used in any field where organisation is necessary. Additionally, a DSS can be designed to help make decisions on the stock market, or deciding which area or segment to market aproduct toward.

3.8 Benefits of DSS

1. Improves personal efficiency
2. Expedites problem solving
3. Facilitates interpersonal communication
4. Promotes learning or training
5. Increases organisational control
6. Generates new evidence in support of a decision
7. Creates a competitive advantage over competition
8. Encourages exploration and discovery on the part of the decision maker
9. Reveals new approaches to thinking about the problem space

3.9 Decision Support System versus Management

Information System

Sales managers typically rely on management information systems (MIS) to produce sales analytic reports. These reports contain sales performance figure by products line, sales person, sales region, and so on. A decision support system (DSS), on the other hand, would also interactively show a sales manager the effects on sales performance of changes in a variety of factors (such as promotion expenses and sales compensation). The DSS could then use several criteria (such as expected gross margin and market share) to evaluate and rank several alternative combinations of sales performance factors.

	MIS	*DSS*
Decision Support Provided	Provide information about the performance of the organisation	Provide information and decision support techniques to analyse specific problem or opportunities
Information form and Frequency	Periodic, exception, demand, and push reports and responses	Interactive enquiries and responses
Information Format	Pre-specified, fixed format	Ad hoc, flexible and adaptable format
Information Processing Methodology	Information provided by extraction and manipulation of business data.	Information produced by analytic and modeling of business data

Figure 1: Comparing DSS and MIS

Using DSS involves an interactive modeling process. For example, using a DSS software package for decision support may result in a series of displays in response to alternative what-if changes entered by a manager. This differs from the demand response of MIS, since decision makers are not demanding pre-specified information. Rather, they are exploring possible alternatives. Thus, they do not have to specify their information need in advance. Instead they use a DSS to find the information they need to help make a decision. That is the essence of DSS concept.

4.0 CONCLUSION

Decision support systems are a form of management information systems used by management to aid decision making process. The complex problems that DSS are out to solve are often very poorly defined with a high level of uncertainty about the true nature of the problem, the various responses which the management could undertake or the likely impact of those actions. Many decision support systems have been developed so that they support implementation on a wide range of hardware.

5.0 SUMMARY

- Decision Support Systems constitute a class of computer-based information systems including knowledge-based systems that support decision-making activities.
- A decision support system is a way to model data and make quality decisions based upon it. Making the right decision in business is usually based on the quality of your data and your ability to sift through and analyse the data to find trends in which you can create solutions and strategies for growth.
- According to Keen, the concept of decision support has evolved from two main areas of research: the theoretical studies of organisational decision makingdone at the Carnegie Institute of Technology during the late 1950s and early 1960s, and the technical work on interactive computer systems, mainly carried out at the Massachusetts Institute of Technology in the 1960s.
- As with the definition, there is no universally-accepted taxonomy of DSS either. Different authors propose different classifications.
- Also, different authors identify different components in a DSS.
- DSS systems are not entirely different from other systems and require a structured approach.
- There are several ways to classify DSS applications. Not every DSS fits neatly into one category, but a mix of two or more architecture in one.
- DSS is extensively used in business and management. Executive dashboard and other business performance software allow faster decision making, identification of negative trends, and better allocation of business resources.
- Sales managers typically rely on management information systems (MIS) to produce sales analytic reports.

6.0 TUTOR-MARKED ASSIGNMENT

1. By taxonomy, list the models of decision support system

2. Mention 5 benefits of decision support systems

7.0 REFERENCES/FURTHER READING

Power, D. J. A May 31, (2003). Brief History of Decision Support *Systems, DSSResources.COM, World Wide Web, version 2.8.*

Keen, P. G. W. (1978). Decision Support Systems: An Organisational *Perspective. Reading, Mass.: Addison-Wesley Pub. Co. ISBN* 0-201-03667-3.

Haettenschwiler, P. (1999). Neues anwenderfreundliches Konzept der *Entscheidungsunterstützung. Gutes Entscheiden in Wirtschaft, Politik und Gesellschaft. Zurich, vdf Hochschulverlag AG:* 189-208.

Power, D. J. (2002). Decision Support Systems: Concepts and *Resources for Managers. Westport, Conn.: Quorum Books.*

Gachet, A. (2004). Building Model-Driven Decision Support Systems *with Dicodess, Zurich, VDF.*

Sprague, R. H. and E. D. Carlson (1982). Building Effective Decision *Support Systems. Englewood Cliffs: N.J., Prentice-Hall. ISBN* 0-130-86215-0.

Haag, Cummings, McCubbrey, Pinsonneault, Donovan (2000). *Management Information Systems for the Information Age.* McGraw-Hill Ryerson Limited: 136-140. ISBN 0-072-81947-2.

Marakas, G. M. (1999). Decision Support Systems In The Twenty-First *Century. Upper Saddle River, N.J.: Prentice Hall.*

Holsapple, C.W.; A. B. Whinston (1996). Decision Support Systems: A *Knowledge-Based Approach. St. Paul: West Publishing. ISBN:* 0-324-03578-0.

Hackathorn, R. D.; P. G. W. Keen (1981, September). "Organisational *Strategies for Personal Computing in Decision Support Systems,"* MIS Quarterly, Vol. 5, No. 3.

Gadomski A. M. et al. (1998). Integrated Parallel Bottom-up and Top- *Down Approach to the Development of Agent-Based Intelligent DSSs for Emergency Management, TIEMS98. Washington:* CiteSeerx – alfa.

DSSAT4 (pdf) The Decision Support System for Agrotechnology *Transfer.*

Stephens, W. and Middleton, T. (2002). Why has the uptake of Decision *Support Systems been so poor? In: Crop-Soil Simulation Models in Developing Countries.*

UNIT 3 DECISION THEORY AND DECISION MAKING

CONTENTS

1.0 INTRODUCTION

Decision Theory in mathematics and statistics is concerned with identifying the values, uncertainties and other issues relevant in a given decision and the resulting optimal decision.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define decision support theory
- identify different types of decision theory
- describe and differentiate normative and descriptive decision theory
- compare and contrast the two major forms of decision theory
- answer the question of what type of decision needs a theory to back it up
- categorise styles and methods of decision making
- explain cognitive and personal biases in decision making.

3.0 MAIN CONTENT

3.1 Normative and Descriptive Decision Theory

Most of decision theory is normative or prescriptive, i.e. it is concerned with identifying the best decision to take, assuming an ideal decision maker who is fully informed, able to compute with perfect accuracy, and fully rational. The practical application of this prescriptive approach (how people should make decisions) is called decision analysis, and aimed at finding tools, methodologies and software to help people make better decisions. The most systematic and comprehensive software tools developed in this way are called decision support systems.

Since it is obvious that people do not typically behave in optimal ways, there is also a related area of study, which is a positive or descriptive discipline, attempting to describe what people will actually do. Since the normative, optimal decision often creates hypotheses for testing against actual behaviour, the two fields are closely linked. Furthermore it is possible to relax the assumptions of perfect information, rationality and so forth in various ways, and produce a series of different prescriptions or predictions about behaviour, allowing for further tests of the kind of decision-making that occurs in practice.

3.2 What Kinds of Decisions Need a Theory?

1. Choice between incommensurable commodities
2. Choice under uncertainty

This area represents the heart of decision theory. The procedure now referred to as expected value was known from the 17th century. Blaise Pascal invoked it in his famous wager (see below), which is contained in his Pensées, published in 1670. The idea of expected value is that, when faced with a number of actions, each of which could give rise to more than one possible outcome with different probabilities, the rational procedure is to identify all possible outcomes, determine their values (positive or negative) and the probabilities that will result from each course of action, and multiply the two to give an expected value. The action to be chosen should be the one that gives rise to the highest total expected value. In 1738, Daniel Bernoulli published an influential paper entitled Exposition of a New Theory on the Measurement of Risk, in which he uses the St. Petersburg paradox to show that expected value theory must be normatively wrong. He also gives an example in which a Dutch merchant is trying to decide whether to insure a cargo being sent from Amsterdam to St Petersburg in winter, when it is known that there is a 5% chance that the ship and cargo will be lost. In his solution, he defines a utility function and computes expected utility rather than expected financial value.

In the 20th century, interest was reignited by Abraham Wald's 1939 paper pointing out that the two central concerns of orthodox statistical theory at that time, namely statistical hypothesis testing and statistical estimation theory, could both be regarded as particular special cases of the more general decision problem. This paper introduced much of the mental landscape of modern decision theory, including loss functions, risk functions, admissible decision rules, a priori distributions, Bayes decision rules, and minimax decision rules. The phrase "decision theory" itself was first used in 1950 by E. L. Lehmann.

The rise of subjective probability theory, from the work of Frank Ramsey, Bruno de Finetti, Leonard Savage and others, extended the scope of expected utility theory to situations where only subjective probabilities are available. At this time it was generally assumed in economics that people behave as rational agents and thus expected utility theory also provided a theory of actual human decision-making behaviour under risk. The work of Maurice Allais and Daniel Ellsberg showed that this was clearly not so. The prospect theory of Daniel Kahneman and Amos Tversky placed behavioural economics on a more evidence-based footing. It emphasised that in actual human (as opposed to normatively correct) decision-making "losses loom larger than gains", people are more focused on changes in their utility states than the states themselves and estimation of subjective probabilities is severely biased by anchoring.

Castagnoli and LiCalzi (1996), Bordley and LiCalzi (2000) recently showed that maximising expected utility is mathematically equivalent to maximising the probability that the uncertain consequences of a decision are preferable to an uncertain benchmark (e.g., the probability that a mutual fund strategy outperforms the S&P 500 or that a firm outperforms the uncertain future performance of a major competitor.). This reinterpretation relates to psychological work suggesting that individuals have fuzzy aspiration levels (Lopes & Oden), which may vary from choice context to choice context. Hence it shifts the focus from utility to the individual's uncertain reference point.

Pascal's Wager is a classic example of a choice under uncertainty. The uncertainty, according to Pascal, is whether or not God exists. Belief or non-belief in God is the choice to be made. However, the reward for belief in God if God actually does exist is infinite. Therefore, however small the probability of God's existence, the expected value of belief exceeds that of non-belief, so it is better to believe in God. (There are several criticisms of the argument.)

Inter-Temporal Choice

This area is concerned with the kind of choice where different actions lead to outcomes that are realised at different points in time. If someone received a windfall of several thousand dollars, they could spend it on an expensive holiday, giving them immediate pleasure, or they could invest it in a pension scheme, giving them an income at some time in the future. What is the optimal thing to do? The answer depends partly on factors such as the expected rates of interest and inflation, the person's life expectancy, and their confidence in the pensions industry. However even with all those factors taken into account, human behaviour again deviates greatly from the predictions of prescriptive decision theory, leading to alternative models in which, for example, objective interest rates are replaced by subjective discount rates.

Competing Decision Makers

Some decisions are difficult because of the need to take into account how other people in the situation will respond to the decision that is taken. The analysis of such social decisions is the business of game theory, and is not normally considered part of decision theory, though it is closely related. In the emerging socio-cognitive engineering the research is especially focused on the different types of distributed decision-making in human organisations, in normal and abnormal/emergency/crisis situations. The signal detection theory is based on the Decision theory.

Complex Decisions

Other areas of decision theory are concerned with decisions that are difficult simply because of their complexity, or the complexity of the organisation that has to make them. In such cases the issue is not the deviation between real and optimal behaviour, but the difficulty of determining the optimal behaviour in the first place. The Club of Rome, for example, developed a model of economic growth and resource usage that helps politicians make real-life decisions in complex situations.

3.3 Paradox of Choice

Observed in many cases is the paradox that more choices may lead to a poorer decision or a failure to make a decision at all. It is sometimes theorised to be caused by analysis paralysis, real or perceived, or perhaps from rational ignorance. A number of researchers including Sheena S. Iyengar andMark R. Lepper have published studies on this phenomenon. A popularisation of this analysis was done by Barry Schwartz in his 2004 book, The Paradox of Choice.

3.4 Statistical Decision Theory

Several statistical tools and methods are available to organise evidence, evaluate risks, and aid in decision making. The risks of type I and type II errors can be quantified (estimated probability, cost, expected value, etc) and rational decision making is improved.

One example shows a structure for deciding guilt in a criminal trial:

		Actual condition	
		Guilty	**Not guilty**
Decision	**Verdict of 'guilty'**	**True Positive**	False Positive (i.e. guilt reported unfairly) **Type I error**
	Verdict of 'not guilty'	False Negative (i.e. guilt not detected) **Type II error**	True Negative

3.5 Alternatives to Probability Theory

A highly controversial issue is whether one can replace the use of probability in decision theory by other alternatives. The proponents of fuzzy logic, possibility theory, Dempster-Shafer theory and info-gap decision theory maintain that probability is only one of many alternatives and point to many examples where non-standard alternatives have been implemented with apparent success. Work by Yousef and others advocate exotic probability theories using complex-valued functions based on the probability amplitudes developed and validated by Birkhoff and Von Neumann in quantum physics.

Advocates of probability theory point to:

- the work of Richard Threlkeld Cox for justification of the probability axioms,
- the Dutch book paradoxes of Bruno de Finetti as illustrative of the theoretical difficulties that can arise from departures from the probability axioms, and
- the complete class theorems which show that all admissible decision rules are equivalent to a Bayesian decision rule with some prior distribution (possibly improper) and some utility function. Thus, for

any decision rule generated by non-probabilistic methods, either there is an equivalent rule derivable by Bayesian means, or there is a rule derivable by Bayesian means which is never worse and (at least) sometimes better.

3.6 Decision Making

Decision Making can be regarded as an outcome of mental processes (cognitive process) leading to the selection of a course of action among several alternatives. Every decision making process produces a final choice. The output can be an action or an opinion of choice.

3.7 Overview of Decision Making

Human performance in decision making terms has been subject of active research from several perspectives. From a psychological perspective, it is necessary to examine individual decisions in the context of a set of needs, preferences an individual has and values he/she seeks. From a cognitive perspective, the decision making process must be regarded as a continuous process integrated in the interaction with the environment. From a normative perspective, the analysis of individual decisions is concerned with the logic of decision making and rationality and the invariant choice it leads to.

Yet, at another level, it might be regarded as a problem solving activity which is terminated when a satisfactory solution is found. Therefore, decision making is a reasoning or emotional process which can be rational or irrational, canbe based on explicit assumptions or tacit assumptions.

Logical decision making is an important part of all science-based professions, where specialists apply their knowledge in a given area to making informeddecisions. For example, medical decision making often involves making a diagnosis and selecting an appropriate treatment. Some research using naturalistic methods shows, however, that in situations with higher time pressure, higher stakes, or increased ambiguities, experts use intuitive decision making rather than structured approaches, following a recognition primed decision approach to fit a set of indicators into the expert's experience and immediately arrive at a satisfactory course of action without weighing alternatives. Also, recent robust decision efforts have formally integrated uncertainty into the decision making process.

3.8 Decision Making Processes Topics

According to behaviouralist Isabel Briggs Myers, a person's decision making process depends on a significant degree on their cognitive style. Myers developed a set of four bi-polar dimensions, called the Myers-Briggs Type Indicator (MBTI). The terminal points on these dimensions are: thinking and feeling; extroversion and introversion; judgment and *perception; and sensing and intuition. She claimed that a person's* decision making style is based largely on how they score on these four dimensions. For example, someone who scored near the thinking, extroversion, sensing, and judgment ends of the dimensions would tend to have a logical, analytical, objective, critical, and empirical decision making style.

Other studies suggest that national or cross-cultural differences exist across entire societies. For example, Maris Martinsons has found that American, Japanese and Chinese business leaders each exhibit a distinctive national style of decision making.

3.9 Cognitive and Personal Biases

Some of the decision making techniques that we use in everyday life include:

- listing the advantages and disadvantages of each option, popularised by Plato and Benjamin Franklin
- flipping a coin, cutting a deck of playing cards, and other random or coincidence methods
- accepting the first option that seems like it might achieve the desired result
- prayer, tarot cards, astrology, augurs, revelation, or other forms of divination
- acquiesce to a person in authority or an "expert"
- calculating the expected value or utility for each option.

For example, a person is considering two jobs. At the first job option the person has a 60% chance of getting a 30% raise in the first year. And at the second job option the person has an 80% chance of getting a 10% raise in the first year. The decision maker would calculate the expected value of each option, calculating the probability multiplied by the increase of value. (0.60*0.30=0.18 [option a] 0.80*0.10=0.08 [option b]). The person deciding on the job would choose the option with the highest expected value, in this example option number one. An alternative may be to apply one of the processes described below, in particular in the Business and Management section.

Biases can creep into our decision making processes. Many different people have made a decision about the same question (e.g. "Should I have a doctor look at this troubling breast cancer symptom I've discovered?" "Why did I ignore the evidence that the project was going over budget?") and then craft potential cognitive interventions aimed at improving decision making outcomes.

Below is a list of some of the more commonly debated cognitive biases.

- Selective search for evidence (a.k.a. Confirmation bias in psychology) (Scott Plous, 1993) - We tend to be willing to gather facts that support certain conclusions but disregard other facts that support different conclusions.
- Premature termination of search for evidence - We tend to accept the first alternative that looks like it might work.
- Inertia - Unwillingness to change thought patterns that we have used in the past in the face of new circumstances.
- Selective perception - We actively screen-out information that we do not think is salient.
- Wishful thinking or optimism bias - We tend to want to see things in a positive light and this can distort our perception and thinking.
- Choice-supportive bias occurs when we distort our memories of chosen and rejected options to make the chosen options seem relatively more attractive.
- Recency - We tend to place more attention on more recent information and either ignore or forget more distant information. The opposite effect in the first set of data or other information is termed Primacy effect (Plous, 1993).
- Repetition bias - A willingness to believe what we have been told most often and by the greatest number of different sources.
- Anchoring and adjustment -Decisions are unduly influenced by initial information that shapes our view of subsequent information.
- Group thinks - Peer pressure to conform to the opinions held by the group.
- Source credibility bias - We reject something if we have a bias against the person, organisation, or group to which the person belongs: We are inclined to accept a statement by someone we like.
- Incremental decision making and escalating commitment - We look at a decision as a small step in a process and this tends to perpetuate a series of similar decisions. This can be contrasted with zero-based **decision making.**
- Attribution asymmetry - We tend to attribute our success to our abilities and talents, but we attribute our failures to bad luck and external factors. We attribute other's success to good luck, and their failures to their mistakes.

- Role fulfillment (Self Fulfilling Prophecy) - We conform to the decision making expectations that others have of someone in our position.
- Underestimating uncertainty and the illusion of control - We tend to underestimate future uncertainty because we tend to believe we have more control over events than we really do. We believe we have control to minimise potential problems in our decisions.

3.10 Neuroscience Perspective

The anterior cingulate cortex (ACC) and orbitofrontal cortex are brain regions involved in decision making processes. A recent neuro-imaging study, Interactions between decision making and performance *monitoring within prefrontal cortex, found distinctive patterns of neural* activation in these regions depending on whether decisions were made on the basis of personal volition or following directions from someone else.

Another recent study by Kennerly, et al. (2006) found that lesions to the ACC in the macaque resulted in impaired decision making in the long run of reinforcement guided tasks suggesting that the ACC is responsible for evaluating past reinforcement information and guiding future action.

Emotion appears to aid the decision making process:

- Decision making often occurs in the face of uncertainty about whether one's choices will lead to benefit or harm (see also Risk). The somatic-marker hypothesis is a neurobiological theory of how decisions are made in the face of uncertain outcome. This theory holds that such decisions are aided by emotions, in the form of bodily states, that are elicited during the deliberation of consequences and that mark different options for behaviour as being advantageous or disadvantageous. This process involves an interplay between neural systems that elicit emotional/bodily states and neural systems that map these emotional/bodily states.

3.11 Styles and Methods of Decision Making

Styles and methods of decision making were elaborated by the founder of Predispositioning Theory, Aron Katsenelinboigen. In his analysis on styles and methods Katsenelinboigen referred to the game of chess, saying that “chess does disclose various methods of operation, notably the creation of predisposition—methods which may be applicable to other, more complex systems.”[

In his book Katsenelinboigen states that apart from the methods (reactive and selective) and sub-methods (randomisation, predispositioning, programming), there are two major styles – positional and combinational. Both styles are utilised in the game of chess. According to Katsenelinboigen, the two styles reflect two basic approaches to the uncertainty: deterministic (combinational style) and indeterministic (positional style). Katsenelinboigen's definition of the two styles are the following.

The combinational style is charaterised by:

- a very narrow, clearly defined, primarily material goal, and
- a program that links the initial position with the final outcome.

In defining the combinational style in chess, Katsenelinboigen writes:

> *The combinational style features a clearly formulated limited objective, namely the capture of material (the main constituent element of achess position). The objective is implemented via a well defined and in some cases in a unique sequence of moves aimed at reaching the set goal. As a rule, this sequence leaves no options for the opponent. Finding a combinational objective allows the player to focus all his energies on efficient execution, that is, the player's analysis may be limited to the pieces directly partaking in the combination. This approach is the crux of the combination and the combinational style of play.*

The positional style is distinguished by:

- a positional goal and
- a formation of semi-complete linkages between the initial step and final outcome.

> *Unlike the combinational player, the positional player is occupied, first and foremost, with the elaboration of the position that willallow him to develop in the unknown future. In playing the positional style, the player must evaluate relational and material parameters as independent variables. (…) The positional style gives the player the opportunity to develop a position until it becomes pregnant with a combination. However, the combination is not the final goal of the positional player —it helps him to achieve the desirable, keeping in mind a predisposition for the future development. The Pyrrhic*

victory is the best example of one's inability to think positionally.

The positional style serves to:

a) create a predisposition to the future development of the position;
b) induce the environment in a certain way;
c) absorb an unexpected outcome in one's favor;
d) avoid the negative aspects of unexpected outcomes.

The positional style gives the player the opportunity to develop a position until it becomes pregnant with a combination. Katsenelinboigen writes:

> *As the game progressed and defense became more sophisticated the combinational style of play declined. . . .*
> *The positional style of chess does not eliminate the combinational one with its attempt to see the entire program of action in advance. The positional style merely prepares the transformation to a combination when the latter becomes feasible.*

4.0 CONCLUSION

Decision support systems do not make decisions. The objective is to allow the manager to consider a number of alternatives and evaluate them under a variety of potential conditions. And there are several theories in decision making. Though these theories vary in approach to decision making, they all have and share the same basic objective of assisting managers to make more effective decision.

5.0 SUMMARY

- Decision Theory in mathematics and statistics is concerned with identifying the values, uncertainties and other issues relevant in a given decision and the resulting optimal decision.
- Most of decision theory is normative or prescriptive, i.e. it is concerned with identifying the best decision to take, assuming an ideal decision maker who is fully informed, able to compute with perfect accuracy, and fully rational.
- Several statistical tools and methods are available to organise evidence, evaluate risks, and aid in decision making.
- A highly controversial issue is whether one can replace the use of probability in decision theory by other alternatives.

- Decision Making can be regarded as an outcome of mental processes (cognitive process) leading to the selection of a course of action among several alternatives.
- Human performance in decision making terms has been subject of active research from several perspectives. From a psychological perspective, it is necessary to examine individual decisions in the context of a set of needs, preferences an individual has and values he/she seeks.
- According to behaviouralist Isabel Briggs Myers, a person's decision making process depends on a significant degree on their cognitive style.
- The anterior cingulate cortex (ACC) and orbitofrontal cortex are brain regions involved in decision making processes.
- Styles and methods of decision making were elaborated by the founder of Predispositioning Theory, Aron Katsenelinboigen.

8.0 TUTOR-MARKED ASSIGNMENT

1. Briefly discuss the Inter-temporal Choice as a kind of decision that needs theory
2. Mention 5 decision-making techniques used in every day life

9.0 REFERENCES/FURTHER READING

James Reason (1990). Human Error. Ashgate. ISBN 1840141042.

Daniel Kahneman, Amos Tversky (2000). Choice, Values, Frames. The Cambridge University Press. ISBN:0521621720.

Isabel Briggs Myers|Myers, I. (1962). Introduction to Type: A *Description of The Theory and Applications of the Myers-Briggs Type Indicator, Palo Alto Ca: Consulting Psychologists Press.*

Martinsons, Maris G. (August 2001). Comparing the Decision Styles of *American, Chinese and Japanese Business Leaders. Best Paper Proceedings of Academy of Management Meetings, Washington,* DC.

Katsenelinboigen, Aron. (1997). The Concept of Indeterminism and Its *Applications: Economics, Social Systems, Ethics, Artificial Intelligence, and Aesthetics. Praeger: Westport, Connecticut, p.6*

Ulea, V. (2002). The Concept of Dramatic Genre and the Comedy of a *New Type. Chess, Literature, and Film. Southern Illinois* University Press, p.p.17-18].

Selected Topics in Indeterministic Systems Intersystems Publications: California, 1989, p. 21.

A New Formula for the Index of Cost of Living, 1939, in Econometrica.

Goode, Erica. (2001). In Weird Math of Choices, 6 Choices Can Beat *600. The New York Times.*

UNIT 4 TYPES OF DSS

CONTENTS

1.0 INTRODUCTION

There are several types of decision support systems and many more are emerging. This goes to show the dynamic nature of the concept of decision support systems. In this unit we shall examine the various types of decision support systems and their applications.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- identify and define the major types of decision support system (DSS)
- understand the applications of the different types of DSS
- compare and contrast the forms of DSS
- understand the concept behind each type of DSS
- appreciate the workings of some of the DSS such as the spatial Decision Support System
- trace the history and development processes of some of the DSS.

3.0 MAIN CONTENT

3.1 Group Decision Support System

This is an interactive computer-based system that facilitates users finding (partial) solutions to semi-structured problems. A GDSS uses a set of decision makers working together as a group.

Technical developments in electronic communication, computing, and decision support, coupled with new interest on the part of organisations to improve meeting effectiveness, are spurring research in the area of group decision support systems (GDSS). A GDSS combines communication, computing, and decision support technologies to facilitate formulation and solution of unstructured problems by a group of people. This unit presents a conceptual overview of GDSS based on an information-exchange perspective of decision making. Three levels of systems are described, representing varying degrees of intervention into the decision process. Research on GDSS is conceived as evolving over time from the study of simple "shell" systems, consisting of menus of features, available for selection by a group, to consideration of sophisticated rule-based systems that enable a group to pursue highly structured and novel decision paths. A multi-dimensional taxonomy of systems is proposed as an organizing frame- work for research in the area. Three environmental contingencies are identified as critical to GDSS design: group size, member proximity, and the task confronting the group. Potential impacts of GDSS on group processes and outcomes are discussed, and important constructs in need of study are identified.

Group Decision Support System Application

The first example or application for a groupware is GDSS, which stands for Group Decision Support System. There is such a GDSS system supported by the Faculty of Management at the University of Arizona. This kind of same-time-same-place conference is oriented toward business meetings and decision making.

The GDSS started originally from the Management Information System at University of Arizona. Some kind of problems has always been observed that are associated more with large meetings than with small meetings. By large meetings we mean meetings with generally more than 15 participants, but can go much beyond that, e.g. 40 or even 50.
Some of the identified problems are:

- time consuming;
- dominance over the meeting; and
- honesty and participation.

However, it is important to note that we are not trying to say that small meetings do not have the above problems; these problems mentioned exist in every kind of meetings, but we are just trying to stress that they are more commonly found in large meetings. Small meetings tend to be more easily controlled than large meetings.

In a GDSS environment, there is usually a big room with something like 40 seats, which means that 40 people can be at the meeting at any one time. There are not only 40 seats but also 40 microcomputers. This enables every participant to have the use of one microcomputer during the course of the meeting. The reason why each participant needs a microcomputer depends on how GDSS works.

In the GDSS, with special computer software, the facilitator of each meeting will first make the agenda of the meeting, which will be projected on to a big screen that everyone can see. The participants will type simultaneously their ideas of the topic of discussion on the individual microcomputers next to them. The computer will then sort the ideas, and the participants will then vote or comment on which ideas they like or dislike. In the course of the whole meeting, GDSS stores, categorises and prints out all the ideas, comments and vote tallies, so that each of the meeting participants will get a summary of the meeting when it ends.

What is so special about GDSS is that it enables meeting participants to simultaneously "talk", when the computer sorts and sends ideas to each of the terminal, all at the same time. That saves a tremendous amount of time, because all these are done electronically instead of manually, and the time saved will enable participants to spend more time manipulating and expressing their ideas. This can consequently increase the productivity and efficiency of the group. The time-saving benefit also has an added bonus: when productivity and efficiency in meetings increase, it is likely that the team spirit can be consolidated, resulting in an increase of the strength of binding among team members.

Besides, under this GDSS, noone can dominate the meeting. This is because of another feature of GDSS. GDSS provides an anonymous scheme, so that whatever you type in the terminal (i.e. your opinion) will be protected. Under this circumstance, no one really knows who is typing what. Because of this, not a single person can dominate the meetings. In the worst case, we might say "some ideas" are dominating the meeting, but this is perfectly fine because this is as a matter of fact an aim of the GDSS: to help meeting participants voice their opinions from an idea-oriented mindset. For example, simply because you have a prejudice against person A does not mean that you are going to reject the idea being proposed in the meeting, because you do not know who is proposing that idea!!

Besides, this anonymity scheme will also help those team members who are shy to voice opinions. And with the anonymity, people are likely to be more honest, just as you will say more, and more honestly on the professor's evaluation form if you know whateveryou write will not affect your final grade on the course. This, of course, is because you know you do not have to worry about the consequences.

However, whether this anonymity is good or not can be very controversial. The success of meetings supported by GDSS depends largely on the conduct of the participants. If people are taking advantage of the anonymity system by typing obscene words or foul languages, this system may be banned for the good of the organisation.

3.2 Business Intelligence

Business Intelligence (BI) refers to technologies, applications and practices for the collection, integration, analysis, and presentation of business information and sometimes to the information itself. The purpose of business intelligence--a term that dates at least to 1958--is to support better business decision making. Thus, BI is also described as a decision support system (DSS).

BI is sometimes used interchangeably with briefing books, report and query tools and executive information systems. In general, business intelligence systems are data-driven DSS.

BI systems provide historical, current, and predictive views of business operations, most often using data that has been gathered into a data warehouse or a data mart and occasionally working from operational data. Software elements support the use of this information by assisting in the extraction, analysis, and reporting of information. Applications tackle sales, production, finance, and many other sources of business data for purposes that include, notably, business performance management. Information may be gathered on comparable companies to produce benchmarks.

History

Prior to the start of the Information Age in the late 20th century, businesses had to collect data from non-automated sources. Businesses then lacked the computing resources necessary to properly analyse the data, and as a result, companies often made business decisions primarily on the basis of intuition.

As businesses automated systems the amount of data increased but its collection remained difficult due to the inability of information to be moved between or within systems. Analysis of information informed for long-term decision making, but was slow and often required the use of instinct or expertise to make short-term decisions. Business intelligence was defined in 1958 by Hans Peter Luhn who wrote that, “business is a *collection of activities carried on for whatever purpose, be it science, technology, commerce, industry, law, government, defense, et cetera".* The communication facility serving the conduct of a business (in the broad sense) may be referred to as an intelligence system. The notion of intelligence is also defined here, in a more general sense, as "the ability *to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal."*

In 1989 Howard Dresner, later a Gartner Group analyst, popularised BI as an umbrella term to describe "concepts and methods to improve *business decision makingby using fact-based support systems." In* modern businesses the use of standards, automation and specialised software, including analytical tools, allows large volumes of data to be extracted, transformed, loaded and warehoused to greatly increase the speed at which information becomes available for decision-making.

Key Intelligence Topics

Business intelligence often uses key performance indicators (KPIs) to assess the present state of business and to prescribe a course of action. Examples of KPIs are things such as lead conversion rate (in sales) and inventory turnover (in inventory management). Prior to the widespread adoption of computer and web applications, when information had to be manually inputed and calculated, performance data was often not available for weeks or months. Recently, banks have tried to make data available at shorter intervals and have reduced delays. The KPI methodology was further expanded with the Chief Performance Officer methodology which incorporated KPIs and root cause analysis into a single methodology.

Businesses that face higher operational/credit risk loading, such as credit card companies and "wealth management" services, often make KPI-related data available weekly. In some cases, companies may even offer a daily analysis of data. This fast pace requires analysts to use IT systems to process this large volume of data.

3.3 Land Allocation Decision Support System

LADSS or Land Allocation Decision Support System is an agricultural land use planning tool being developed at The Macaulay Institute.

LADSS is implemented using the programming language G2 from Gensym alongside a Smallworld GIS application using the Magik programming language and an Oracle database. LADSS models crops using the CropSyst simulation model. LADSS also contains a livestock model plus social, environmental and economic impact assessments.

LADSS has recently been used to address climate change issues affecting agriculture in Scotland and Italy. Part of this work has involved the use of General Circulation Models (also known as Global climate models) to predict future climate scenarios. Other recent work has been a study into how Common Agricultural Policy reform will affect the uplands of Scotland, an assessment of agricultural sustainability and for rural development research within the AGRIGRID project.

3.4 Spatial Decision Support System

Spatial Decision Support Systems (SDSS) developed in parallel with the concept of Decision Support Systems (DSS).

An SDSS is an interactive, computer-based system designed to support a user or group of users in achieving a higher effectiveness of decision making while solving a semi-structured spatial problem. It is designed to assist the spatial planner with guidance in making land use decisions. For example, when deciding where to build a new airport many contrasting criteria, such as noise pollution vs employment prospects or the knock on effect on transportation links, which make the decision difficult. A system which models decisions could be used to help identify the most effective decision path.

An SDSS is sometimes referred to as a Policy Support System. A spatial decision support system typically consists of the following components.

1. A database management system - This system holds and handles the geographical data. A standalone system for this is called a Geographical Information System, (GIS).
2. A library of potential models that can be used to forecast the possible outcomes of decisions.
3. An interface to aid the users' interaction with the computer system and to assist in analysis of outcomes.

This concept fits dialog, data and modelling concepts outlined by Sprague and Watson as the DDM paradigm.

How does an SDSS work?

An SDSS usually exists in the form of a computer model or collection of interlinked computer models, including a land use model. Although various techniques are available to simulate land use dynamics, two types are particularly suitable for SDSS. These are Cellular Automata (CA) based models and Agent Based Models (ABM).

An SDSS typically uses a variety of spatial and non-spatial information, like data on land use, transportation, water management, demographics, agriculture, climate or employment. By using two (or, better, more) known points in history the models can be calibrated and then projections into the future can be made to analyse different spatial policy options. Using these techniques spatial planners can investigate the effects of different scenarios, and provide information to make informed decisions. To allow the user to easily adapt the system to deal with possible intervention possibilities an interface allows for simple modification to be made.

Examples where an SDSS has been used

Community Viz

Community Viz is a land-use planning SDSS that works as an extension to ArcGIS Geographic Information System software produced by ESRI. It uses a scenario planning approach and calculates economic, environmental, social and visual impacts and indicators dynamically as users explore alternatives. Interactive 3D models and various tools for public participation and collaboration are also included. It has been commercially available since 2001.

Environment Explorer

The Environment Explorer (LOV) is a spatial, dynamic model, in which land use and the effects on social, economicand ecological indicators are modelled in an integrated way. Its primary goal is to explore future developments, combining autonomous developments with alternative policy options, in relation to the quality of the environment in which inhabitants of the Netherlands live, work and recreate. Various policy options from governmental departments are translated into a spatial, dynamic image of the Netherlands future with respect to issues such as: economic activity, employment, social well-being, transportation and accessibility, and the natural environment. The model covers the whole of The Netherlands.

LUMOCAP

LUMOCAP aims at delivering an operational tool for assessing land use changes and their impact on the rural landscape according to a Common Agricultural Policy (CAP) orientation. It focuses on the relations between the CAP and landscape changes and emphasises the spatio-temporal dimension of the former. The core of the tool is a dynamic Cellular Automata based land use model. Current usage areas - Poland (2 areas), Germany / The Netherlands (1 cross border area)

MOLAND

The aim of MOLAND is to provide a spatial planning tool that can be used for assessing, monitoring and modelling the development of urban and regional environments. The project was initiated in 1998 (under the name of MURBANDY – Monitoring Urban Dynamics) with the objective to monitor the developments of urban areas and identify trends at the European scale. The work includes the computation of indicators and the assessment of the impact of anthropogenic stress factors (with a focus on expanding settlements, transport and tourism) in and around urban areas, and along development corridors. Models now covering 23 European cities (map)

MURBANDY

The overall objective of MURBANDY is to provide datasets to study past and current land uses, to develop an Earth Observation based procedure to monitor the dynamics of European cities; to develop a number of "urban" and "environmental" indicators that allow to understand these dynamics and the impact these cities have on the environment, and finally to elaborate scenarios of urban growth. Initially this project covered 5 European cities, but the project has expanded into the MOLAND project.

Zer0-M

Zer0-M aims at concepts and technologies to achieve optimized close-loop usage of all water flows in small municipalities or settlements (e.g. tourism facilities) not connected to a central wastewater treatment - the Zero Outflow Municipality (Zer0-M).

3.5 Self Service Software

Self Service Software is a subset within the Knowledge Management software category and which contains a range of software that specialises in the way information, process rules and logic are collected, framed within an organised taxonomy, and accessed through decision support interviews. Self-service software allows people to secure answers to their inquiries and/or needs through an automated interview fashion instead of traditional search approaches.

Self Service Software Functionality

Self service software allows authors (typically subject matter experts) to readily automate the deployment of, the timeliness of, and compliance around a variety of processes with which they are involved in communicating without havingto physically address the questions, needs, and solicitations of end users who are inquiring about the particular process being automated.

Self service software primarily addresses closed-loop inquiries whereby the author emulates a variety of known (finite) questions and related (known) responses on hand or required steps that must be addressed to derive and deliver a final answer or directive. Often the author using such software codifies such known processes and steps then generates (publishes) end-user facing applications which can encompass a variety of code bases and platforms.

Self service software is sometimes referred to as decision support software and even expert systems. It is typically categorised as a sub-topic within the knowledge management software category. Self service software allows individuals and companies alike to tailor and address customer support, technical support and employee support inquiries and needs in an on-demand fashion where the person with a question (need) can interface with the author's generated application via a computer, a handheld device, a kiosk, register, or other machine type to secure their answers as if they were directly interacting (talking to) with the author.

3.6 Dashboards (Management Information Systems)

In management information systems, a dashboard is an executive information system user interface that (similar to an automobile's dashboard) is designed to be easy to read. For example, a product might obtain information from the local operating system in a computer, from one or more applications that may be running, and from one or more remote sites on the Web and present it as though it all came from the same source.

Types of Dashboards

Dashboard of Sustainability screen shot illustrating example dashboard layout. (Sentence is hanging)

Digital dashboards may be laid out to track the flows inherent in the business processes that they monitor. Graphically, users may see the high-level processes and then drill down into low level data. This level of detail is often buried deep within the corporate enterprise and otherwise unavailable to the senior executives.

Three main types of digital dashboard dominate the market today: stand alone software applications, web-browser based applications, and desktop applications also known as desktop widgets. These latter would be driven by a Widget engine.

Specialised dashboards may track all corporate functions. Examples include human resources, recruiting, sales, operations, security, information technology, project management, customer relationship management and many more departmental dashboards.

Digital dashboard projects involve business units as the driver and the information technology department as the enabler. The success of digital dashboard projects often relies on the correct selection of metrics to monitor. Key performance indicators, balanced scorecards, sales performance figures — these are just some of the content appropriate on business dashboards.

Interface Design Styles

To some extent, most graphical user interfaces (GUIs) resemble an automobile dashboard. Although a computer dashboard is more likely to be interactive than an automobile dashboard, some product developers consciously employ this metaphor (and sometimes the term) in the interface design so that the user instantly recognises the similarity. Some products that aim to integrate information from multiple components into a unified display refer to themselves as dashboards. Based on the metaphor of the instrument panel in a car, the computer, or "digital" version of a dashboard provides a business manager with the input necessary to "drive" the business. Highlights with colors similar to traffic lights, alerts, drill-downs, summaries, and graphics such as bar charts, pie charts, bullet graphs, sparklines and gauges are usually set in a portal-like environment that is often role-driven and customisable.

History

Historically, the idea of digital dashboards follows the work in the 1970s with the study of decision support systems. In the late 1990s with the surge of the web, digital dashboards as we know them today began appearing. Many systems were developed in-house by organisations to consolidate and display data alreadybeing gathered in various information systems throughout the organisation. Today, digital dashboard technology is available "out-of-the-box" with many software providers on the scene. Certain companies however still continue to do in-house development and maintenance of dashboard applications. For example, GE Aviation has developed a proprietary software/portal called "Digital Cockpit" to monitor the trends in aircraft spare parts business.

Benefits of Digital Dashboards

Most organisations have various departments all contributing to its overall success and thus it is important to be able to assess the progress of each department. Digital dashboards, which are a type of executive information system, allow managers to do just that. To gauge exactly how well an organisation is performing overall, digital dashboards allow you to capture and report specific data points from each department within the organisation, thus providing a "snapshot" of performance.

Some benefits of using digital dashboards include:

- visual presentation of performance measures
- ability to identify and correct negative trends.
- measures efficiencies/inefficiencies.
- ability to generate detailed reports showing new trends.
- increases productivity
- ability to make more informed decisions based on collected BI (business intelligence)
- aligns strategies and organisational goals.
- saves time over running multiple reports
-

4.0 CONCLUSION

As there are several field in human endeavour so are there likely to be models and types of decision making systems. Aside from generalised decision support systems applicable to all, there are very specific decision support systems that address a particular subject or discipline. These types and models commonly have their strengths and weaknesses, but they all have the same concept of presenting a decision maker with variety of decisions to choose from based on the individual need of the managers.

5.0 SUMMARY

- A GDSS is an interactive computer-based system that facilitates users finding (partial) solutions to semi-structured problems. A GDSS uses a set of decision makers working together as a group.
- Business Intelligence (BI) refers to technologies, applications and practices for the collection, integration, analysis, and presentation of business information and sometimes to the information itself.
- LADSS or Land Allocation Decision Support System is an agricultural land use planning tool being developed at The Macaulay Institute.
- Spatial Decision Support Systems (SDSS) is developed in parallel with the concept of Decision Support Systems (DSS).An SDSS is an interactive computer.
- Self Service Software (SSS) is a subset within the Knowledge Management software category and which contains a range of software that specialises in the way information, process rules and logic are collected, framed within an organised taxonomy, and accessed through decision support interviews.
- In management information systems, a dashboard is an executive information system user interface that (similar to an automobile's dashboard) is designed to be easy to read.

6.0 TUTOR-MARKED ASSIGNMENT

Mention 5 types of decision support systems.

7.0 REFERENCES/FURTHER READING

Bontis, N., Dragonetti, N., Jacobsen, K. and G. Roos. (1999). The *Knowledge Toolbox: A Review Of The Tools Available To Measure And Manage Intangible Resources, European* Management Journal, 17, 4, 391-402. 1.

D. J. Power (2007-03-10). A Brief History of Decision Support Systems, version 4.0. DSSResources.COM.

H. P. Luhn (October 1958). A Business Intelligence System (PDF). IBM *Journal.*

Dictionary definition of dashboards, contributed by Matt Sorenson, SearchCIO.com, 25 February 2002. Retrieved: 9 February 2008.

Few, Stephen (2006). Information Dashboard Design. O'Reilly. ISBN 978-0-596-10016-2.

UNIT 5 DECISION SUPPORT IN E-BUSINESS

CONTENTS

1.0 INTRODUCTION

To succeed in e-business and e-commerce, companies need information systems that can support the diverse information and decision-making needs of the mangers and business professionals.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- link the concept of e-business and decision support systems
- identify the major e-business decision support applications that are being customised, personalised, and web-enabled for use in e-business and e-commerce
- use a case study to illustrate the application of decision support system on e-business
- identify the major structure and corresponding characteristic of information
- use decision support system in e-business
- describe how to create an effective decision support system for e-business
- explain the relationship between information, decision and management.

3.0 MAIN CONTENT

3.1 E-Business Decision Support Trend

Using information systems to support business decision-making has been one of the primary thrusts of the business use of information technology. However, the e-commercerevolution spawned by the Internet and the World Wide Web is expanding the information and decision-support uses, and expectations of many companies's employees, managers, customers, suppliers and other business partners.
But these changes were noticed even earlier, as both academic researchers and business practitioners began reporting that the traditional managerial focus originating in classic MIS (1950s), DSS (1970s) and EIS (1980s) was expanding. The fast pace of new information technologies like personal computer hardware and software suites, client/server networks and networked PC versions of DSS/EIS software made decision support available to lower levels of management as well as non-managerial individuals and self-directed teams of business professionals.

This trend has accelerated with Internet and e-commerce revolutions, and the dramatic growth in Intranets and Extranets that internetworked e-business enterprises and their stakeholders. Figure 1 illustrates that all e-commerce participants expect easy and instant access to information and Web-enabled self-service data analysis. Internetworked e-business enterprises are responding with a variety of personalised and proactive Web-based analytic technologies to support the decision-making requirement of all their stakeholders.

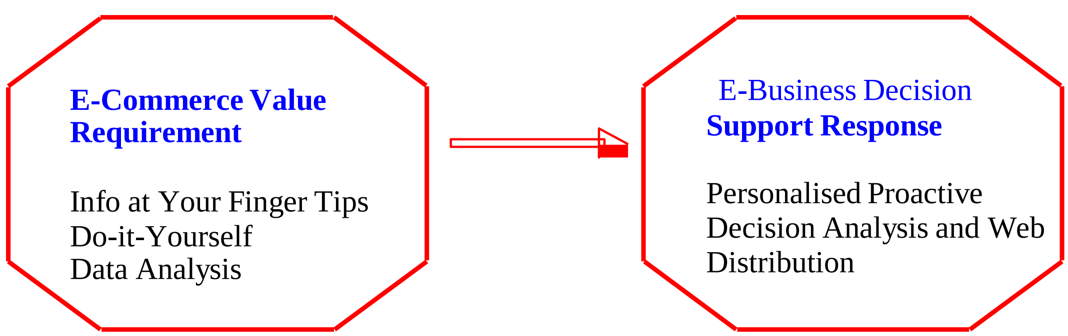


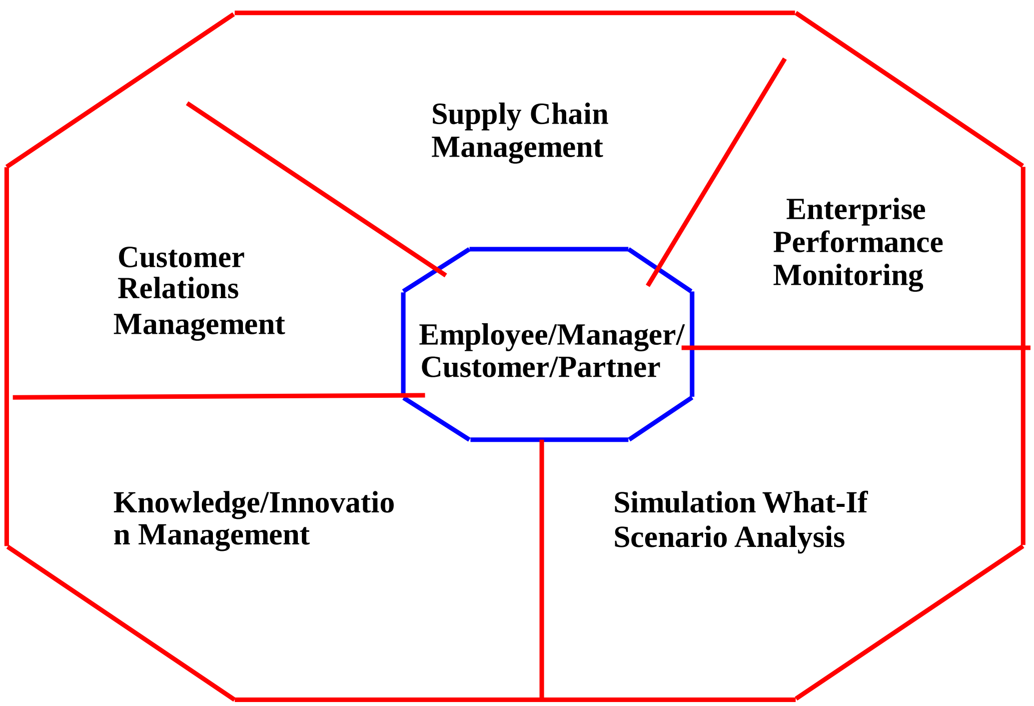
Figure 1: An E-Business Meeting the information and Data Needs
Source: Adapted from Ravi Kalakota and Marcia Robinson, E-Business Roadmap for Success.

> ***Note: An e-business enterprise must meet the information and data analysis*** requirements of customers and companies in e-commerce with more personalised and proactive Web-based decision support.

Thus, the growth of corporate intranets, extranets, as well as the Web has accelerated development and use of "economic class" information delivery and decision support software tools by lower level of management and by individuals and teams of business professionals. In addition, the dramatic expansion of e-commerce has opened the door to the use of e-business DSS tools by thesuppliers, customers, and other business stakeholders of many companies from customer relationship management, supply and chain management and other e-business applications.

Figure 2 highlights several of the major e-business Decision Support applications that are being customised, personalised, and Web-enabled for use in e-business and e-commerce

Figure 2: Examples of E-Business Decision Support Applications available to managers, employees, customers, suppliers, and other business partners of e-business enterprise.



Source: Adapted from Ravi Kalakota and Marcia Robinson, E-Business Roadmap for Success.

3.2 Case Study: Target Corporation E-Business Decision Support System

Target Corporation's Decision Support System (DSS) is composed of several applications known as Decision Maker's Workbench, which use Decision Suite and WebOLAP Software from Information Advantage. The DSS and Target's corporate intranet supports more than 1,700 active users creating more than 60,000 ad-hoc online processing (OLAP) reports each month. During the Christmas seasons, more than 20,000 analytic OLAP reports are produced each day by integrating the Web with its corporate data warehouse to monitor the sales and performance of their own products via secure extranet links across the Internet.

With the Target Store systems complete, the corporation has standardised it as a model for the entire company. Already the standardised warehouse has enabled Target Corporation to obtain more accurate data on how items are performing across divisions, across the company. This has improved vendor negotiations considerably by enabling the different divisions to consolidate orders and receive a better price. The standardised DSS applications also allowfor cross referencing offashion trend across divisions, and they have helped validate merchandizing hunches through analysis of cross-company data.

3.3 Information, Decisions and Management

Figure 3 emphasises that the type of information required by decision makers in a company is directly related to the level of management **decision-making and the amount of structure in the decision situation** they face. You should realise that the framework of the classic **managerial pyramid shown in figure 3 applies in even today's downsized organisations and flattened or non-hierarchical** organisational structures. Levels of management decision making still exists, but their sizes, shape and participants continue to change as today's fluid organisational structure evolves. Thus the levels of managerial decision making development that must be supported by information technology in a successful organisation are:

- Strategic Management: Typically, a board of directors and an executive committee of the CEO and top executives develop overall original goals, strategies, policies and objectives as part of a strategic planning process. They also monitor the strategic performance of the organisation and its overall direction in the political, economic, and competitive business environment.

- Tactical Management: Increasingly, business professionals in self-directed teams, as well as business unit's managers develop their short-based and medium-range plans, schedules, and budgetsand specify the policies, procedures, and business objectives for their subunits of the company. They also allocate resources and monitor the performance of their organisational sub units, including departments, divisions, process teams, project teams, and other workgroups.

- Operational Management: The members of self-directed teams or operating managers develop short-range plans such as weekly production schedules. They direct the use of resources and the performance of the tasks according to the procedures and within budgets and schedules theyestablished for the team and other workgroups of the organisation.

3.4 Decision Structure

Decisions made at the operational management level tend to more **structure, those at tactical level more semi-structured, and those at the** strategic management level more unstructured. Structured decisions involve situations where the procedures to follow when a decision is needed can be specified in advance. The inventory reorder decisions faced by most business are a typical example. Unstructured decisions involve decisions where it is not possible to specify in advance most of the decisions and procedures to follow. At most, many decision situations are semi-structured. That is, some decision procedures can be pre-specified, but not enough to lead to a definite recommendation decision. For example, decisions involved in the starting of a new line of e-commerce services or making major changes to employee benefits would probably range from unstructured to semi-structured. Figure 4 provides a variety of examples of business decisions by type of decision structured and level of management.

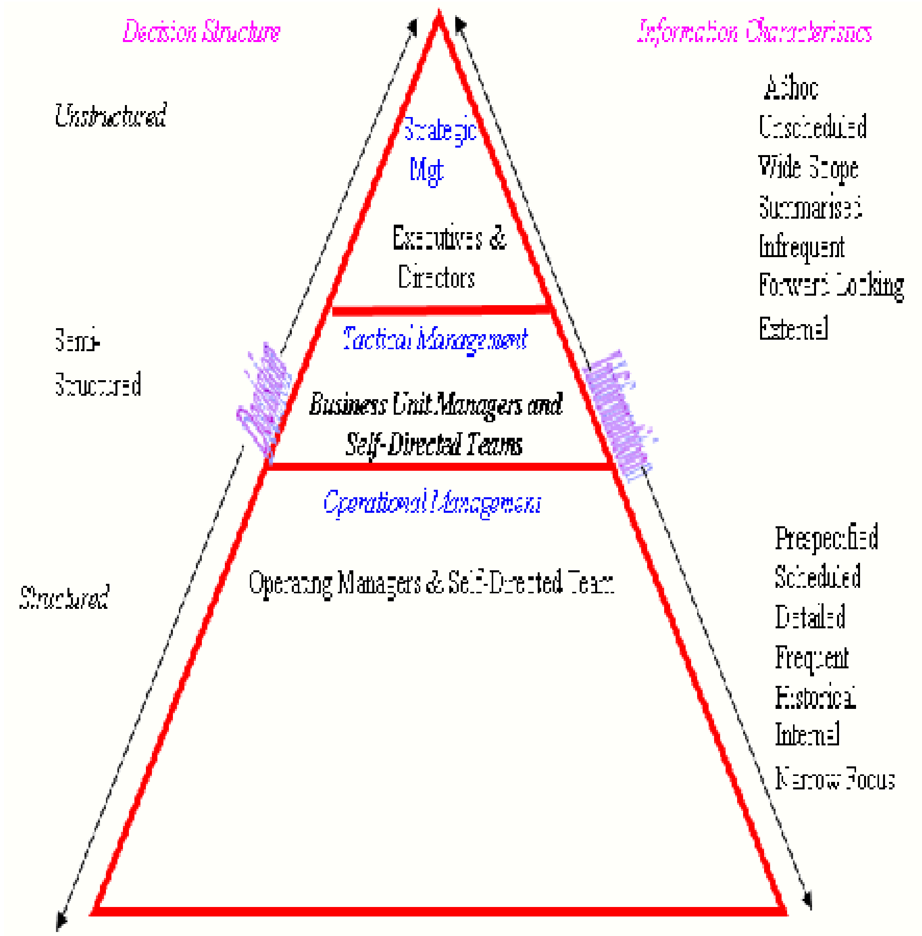


Figure 3: Information Requirement for Decision Makers

Therefore, information systems must be designed to produce a variety of products to meet the changing needs of decision makers throughout an organisation. For example, decision makers at the strategic management levels require more summarised adhoc, unscheduled reports, forecasts, and external intelligence to support their more unstructured planning and policy-making responsibilities. Decision makers at the operational management level, on the other hand, may require more pre-specified internal reports emphasising detailed current and historical data comparison that support their more structured responsibilities on day-to-day operations.

Decision Structure	Operational Management	Tactical Management	Strategic Management
Unstructured	• Cash Management	• Business Process reengineering • Workgroup performance analysis	• New e-commerce initiative • Company organisation
Semi-Structured	• Credit Management • Production Schedule • Daily Work Assignment	• Employee performance appraisal • Capital Budgeting • Progra m Budgeting	• Product Planning • Merger & Acquisition • Site Location
Structured	• Inventory Control	• Progra m Control	

Figure 4 Examples of Business Decisions by Type of Decision

3.5 Using Decision Support System

Typically, the four major options used in decision support systems are:

1. Analytic models
2. Specialised Databases like data warehouse
3. Decision-Makers own insights and judgments
4. An Interactive Computer-based modelling process to support the making of semi-structured and unstructured decisions.

Of the four major options above, using one of the option, that is, the analytic modelling, also involves four basic activities:

- What-If Analysis
- Sensitivity Analysis
- Goal Seeking Analysis
- Optimisation Analysis

What-If Analysis

In What-if analysis, an end user makes changes to variables or relationships among variables, and observes the resulting changes in the value of other variables. For example, if you are using a spreadsheet, you might change a revenue amount (variable) or a tax rate formula (a relationship among variables) in a simple financial spreadsheet model. Then you could command the spreadsheet program to instantly recalculate all affected variables in the spreadsheet. A managerial user would be very interested in observing and evaluating changes that occurred to the value of the spreadsheet, especially to a variable such as net profit after taxes. To many managers, net after tax is an example of the bottom line, that is, a key factor in making many types of decisions. This type of analysis would be repeated until the manager was satisfied with what the results revealed about the effects of various possible decisions.

Sensitivity Analysis

This is a special case of What-If Analysis. Typically the value of only one variable is changed repeatedly, and the resulting changes on other variables are observed. So sensitivity analysis is really a case of What-If Analysis involving repeated changes to only one variable at a time. Some DSS packages automatically make repeated changes to a variable when asked to perform sensitivity analysis. Typically sensitivity analysis is used when decision makers are uncertain about the assumptions made in estimating the value of certain key variables.

Goal-Seeking Analysis

This reverses the direction of the analysis done in What-If and sensitivity analysis. Instead of observing how changes in the variable affect other variables, goal-seeking analysis (also called How Can analysis) sets a target value (a goal) for a variable and then repeatedly changes other variables until the target value is achieved. For example, you could specify a target value (goal) of $2 million for net profit after taxes from a business venture. Then you could repeatedly change the value of revenue or expenses in a spreadsheet model until a result of $2 million is achieved. Thus, you would discover what amount of revenue level or level of expenses the business venture needs to achieve in order to reach the goal of $2 million in after tax profits. This form of analytical modeling would help answer the questions: "How can we achieve $2million in net profit after taxes?" instead of the question "What happens if we change revenue or expenses?". Thus goal-setting analysis is another important method of decision support.

Optimisation Analysis

This is a more complex extension of goal seeking analysis instead of setting a specific target value for a variable, the goal is to find the optimisation value for one or more target variables, given certain constraints. Then one or more other variables are changed repeatedly, subject to the specific constraints, until the best value for the target variables are discovered. For example, you could try to determine the highest possible level of profit that could be achieved by varying the values for selected revenue resources and expense categories. Changes to such variables could be subject to constraints such as limited capacity of the production process or limits to variable financing. Optimisation typically is accomplished by special-purpose software packages for optimisationtechniques such as linear programming or by advance decision support system generators.

3.6 Creating Effective Decision Support System

The concept of an interactive computer-based system that helps companies make better business decisions has been around since computers came into widespread use.

The vision is deceptively simple. Companies take advantage of in-depth reporting tools and predictive models to analyse data and learn what happened in their business, why it happened and, eventually, what will happen. This yields a deep, fact-based understanding that complements experience and intuition and leads to exemplary decision-making and dramatic competitive advantage.

Types of Analytic Modeling	Examples and Activities
What-If-Analysis	*Observing how changes to selected variables* affect other variables Example: What if we cut advertising by 10%? What would happen to sales
Sensitivity Analysis	*Observing how repeated changes to a single* variable affect other variables Example: Let's cut advertisement by $ 100 repeatedly so we can see its relationship to sales
Goal-Seeking Analysis	*Making repeated changes to selected variables* until a chosen variable reaches a target value Example: Let's try increases in advertisement until sales reaches $1 million
Optimisation Analysis	*Finding an optimum value fro selected variables,* given certain constraints. Example: What's the best amount of advertisement to have, given our budget and choice of media.

Figure 5: Activities and Examples of the Major Types of Analytic Modeling

While this sounds simple, in reality there are a number of pitfalls that can complicate the implementation of a decision support system (DSS). Nevertheless, a number of innovative companies have avoided the pitfalls by understanding that creation of a successful data-driven DSS requires three essential elements: first, a process that carefully gauges the short- and long-term decision support needs of the business and what is required to meet those needs; second, a flexible, step-by-step plan for growth; and, third, a centralised data warehouse that can deliver a single, comprehensive, and up-to-date picture of the enterprise that all levels of the organisation can access as needed.

Defining the Business Needs

The process begins with an analysis aimed at precisely defining the business challenge(s) that your DSS will address both in the short- and long-term.

After identifying all the areas that might benefit from improved decision support, you must then prioritise. For example, you may recognise that your marketing department has delivered solid results for its direct marketing programmes. How can you maintain or boost the direct marketing programme's success while lowering the cost? The data-driven DSS enables you to cost effectively target customers that are most likely to respond. It helps you more clearly understand the customers you have been targeting, the customers you need to target (perhaps broken down by geography and demographics), and what you can change to improve the targeted mailings.

In another part of the company your customer service department has been struggling to get accurate, up-to-date information on computer screens at the call-in centre. Here, you are hoping the data-driven DSS can help improve customer service and create opportunities for your agents to offer additional products and services.

Both situations present legitimate cases for implementing a DSS, but to decide how to proceed you need to consider such things as: Which business unit can deliver the most return on investment (ROI)? Which project can deliver the most substantial ROI? How quickly can that ROI be delivered?

The bottom line here is that companies that successfully implement a DSS clearly define and prioritise the problems they are trying to solve, understand what they expect the DSS to achieve, and establish a way to measure its success. Once you have prioritised and implemented - your measurable results can justify extending the data-driven DSS to other business areas and applications.

Defining User Needs

Once youhave defined the business goal, it is time to define the targeted user group and its particular needs. Two of the most important questions to explore are: How current does the information need to be? And in what form do the users need it?

Consider three potential user groups. Executives tend to need a global, high-level look with much of the analysis done behind the scenes. Rarely do they need information in real-time. In contrast, the front-line employee does need up-to-date information in real-time, but does not typically need global or in-depth types of analysis. And your power users - mid-tier decision-makers such as buyers in retail stores or operations managers who do scheduling for airlines - need both real time and historical detail.

It is helpful to create a matrix or table that details the system's users by job positions or by department and document how many individuals comprise each category. From there, you can define the general data requirements of each audience. Include such things as output formats, the type of data, the audience's need to analyse the data, the required frequency of the reports, and how the data will be used.

It is also crucial at this point to consider what the needs of each user group imply for privacy and security concerns. For example, at a health insurer, some people may need to see only the number of emergency room admissions over the last six months, while others may need to see individual diagnoses and treatments. The system must be designed in such a way that it clearly identifies and provides access to those who need the selected information and those who should be denied access. This too should become part of your matrix.

Defining the Data Needs

The third step is to evaluate what data you will need for the system, which involves creating a model that will describe all the data you will need to address the business problem you are trying to solve.

Data Sources: Let's assume you are trying to create a direct marketing campaign for a new life insurance product. You know that at the very least you want to be able to see your prospective clients broken down by age, income, and their current life stage (for example, whether or not they have recently had children). Once you have identified the data you will need, you can map that data against the data elements that currently exist in your environment - and identify where those elements exist. You may find that you are not collecting some data that you need and so will have to start collecting it.

Then you can determine contributing factors such as: How many sources are involved? What type of data is contained in each source and in what formats? How much data does each database contain (the number of tables and columns available for query purposes)? Are these databases snapshots at a point in time or are they online in real time?

You will need all this information to determine your strategy for loading data into the DSS. Which data gets loaded first? Does all the data need to be refreshed simultaneously? You need to look holistically at all the source information to create your data loading strategy.

It is also important to remember that as you identify common data from multiple sources, you must plan to integrateit into the singledata warehouse so that you have a consistentview of the same data. This may require data cleansing or data transformations.

This consistent view of the data is crucial because so often a data-driven DSS project can be tripped up by bad or inconsistent data. When the entire user group is working from the same, reliable and up-to-date information, they make better, more aligned decisions across the board and get the most out of the information you possess.

Data Access: Getting the maximum return on information, of course, is the ultimate goal of data-driven DSS. It is encouraging that the potential to achieve that maximum return has grown with the emergence of powerful analytical tools and applications that provide access to data in a number of ways - and which deliver important new business insights.

To get the most out of these tools they must connect to a database that is optimally configured to handle multiple, concurrent queries and to keep the data accurate, up-to-date and consistent for all of the system's users. An enterprise data warehouse provides those qualities in the most efficient and cost-effective way.

Gap Analysis to Action Plan

Having identified the business needs, the user needs, and the data needs, you can now compare your findings to the existing environment and determine what gaps you will need to fill.

The gap analysis should lead to a step-by-step action plan that addresses both short- and long-term needs. You need to identify priorities and develop strategies that include a description of initial investments as well as pilot projects upon which the entire DSS can build.

One common problem in the planning phase is that IT projects are not well-defined. This leads to unmet expectations, cost overruns and a perception of limited success which translates into not getting an adequate return on investment. It is critical to spend sufficient time planning to avoid these pitfalls.

Ultimately, building a flexible, scalable system that keeps business intelligence flowing and that can respond to ever-changing business needs is the mark of an effective data-driven DSS. A careful, step-by-step development process tailored to the needs of your business - with an enterprise data warehouse at the heart of the solution - can deliver ongoing competitive advantage.

4.0 CONCLUSION

The effect or importance of decision support systems are felt more in business organisations than probably in other areas of human endeavour. Moreover now that business is tending towards being totally electronic, decision support systems now become much more important in the business world. Decision support systems go a long way in assisting businesses to be efficient and productive.

5.0 SUMMARY

- Using information systems to support business decision-making has been one of the primary thrusts of the business use of information technology.
- Target Corporation's decision support system (DSS) is composed of several applications known as Decision Maker's Workbench, which use Decision Suite and WebOLAP Software from Information Advantage.
- The type of information required by decision makers in a company is directly related to the level of management decision-making and the amount of structure in the decision situation they face.
- Decisions made at the operational management level tend to more **structure, those at tactical level more semi-structured, and those at** the strategic management level more unstructured.
- In What-if analysis, an end user makes changes to variables, or relationships among variables, and observes the resulting changes in the value of other variables.
- The concept of an interactive computer-based system that helps companies make better business decisions has been around since computers came into widespread use.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention the characteristics of unstructured decisions
2. Briefly discuss Goal-Seeking Analysis

7.0 REFERENCES/FURTHER READING

Jessani, R., Creating an Effective Data-Driven Decision Support *System.*

DSSResources.COM, 12/05/2003.

MODULE 2

UNIT 1 ONLINE ANALYTIC PROCESSING, ONLINE TRANSACTION PROCESSING AND BUSINESS ANALYTICS

CONTENTS

1.0 INTRODUCTION

Online Analytical Processing, or OLAP (IPA), is an approach to quickly provide answers to analytical queries that are multi-dimensional in nature. OLAP is part of the broader category business intelligence, which also encompasses relational reporting and data mining. The typical applications of OLAP are in business reporting for sales, marketing, management reporting, business process management (BPM), budgeting and forecasting, financial reporting and similar areas.
The term OLAP was created as a slight modification of the traditional database term OLTP (Online Transaction Processing).

Databases configured for OLAP employ a multidimensional data model, allowing for complex analytical and ad-hoc queries with a rapid execution time. They borrow aspects of navigational databases and hierarchical databases that are speedier than their relational kin.

Nigel Pendse has suggested that an alternative and perhaps more descriptive term to describe the concept of OLAP is Fast Analysis of **Shared Multidimensional Information (FASMI)**

The output of an OLAP query is typically displayed in a matrix (or pivot) format. The dimensions form the row and column of the matrix; the measures, the values.

OLAP enables managers and analysts to interactively examine and manipulate large amounts of detailed and consolidated data from many perspectives. OLAP involves analysing complex relationships among thousands or even millions of data items stored in multi-dimensional databases to discover patterns, trends, and exception conditions An OLAP session takes place online in real-time, with rapid responses to a managers' or analyst's queries so that their analytical or decision-making process is undisturbed. (Figure 1).

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define online analytic processing (OLAP)
- trace the development and history of OLAP
- outline the functions of OLAP
- identify the different types of OLAP
- describe the market structure of OLAP
- define online transaction processing and its requirement, benefits and disadvantages
- define business analytic, its history, applications and challenges.

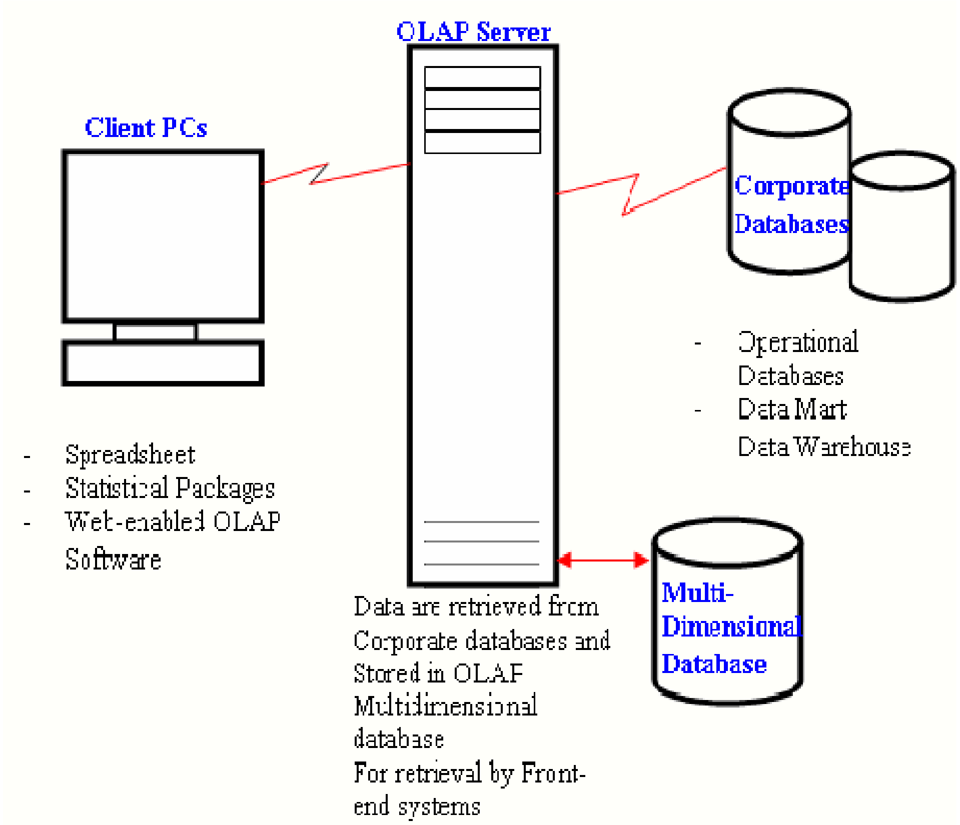


Figure 1: Online Analytic Processing may involve the use of Special Servers and Multidimensional Databases

3.0 MAIN CONTENT

3.1 History

The first product that performed OLAP queries was Express, which was released in 1970 (and acquired by Oracle in 1995 from Information Resources). However, the term did not appear until 1993 when it was coined by Ted Codd, whohas been described as "the father of the relational database". Codd's paper resulted from a short consulting assignment which Codd undertook for former Arbor Software (later Hyperion Solutions, and in 2007 acquired by Oracle), as a sort of marketing coup. The company had released its own OLAP product, *Essbase, a year earlier. As a result Codd's "twelve laws of online* analytical processing" were explicit in their reference to Essbase. There was some ensuing controversy and when Computerworld learned that Codd was paid by Arbor, it retracted the article. OLAP market experienced strong growth in the late 90s with dozens of commercial products going into the market. In 1998, Microsoft released its first OLAP Server - Microsoft Analysis Services,which drove wider adoption of OLAP technology and moved it into mainstream.

3.2 Functionality

In the core of any OLAP system is a concept of an OLAP cube (also called a multidimensional cube or a hypercube). It consists of numeric facts called measures which are categorised by dimensions. The cube metadata is typically created from a star schema or snowflake schema of tables in a relational database. Measures are derived from the records in the fact table and dimensions are derived from the dimension tables.

3.3 Aggregations

It has been claimed that for complex queries OLAP cubes can produce an answer in around 0.1% of the time for the same query on OLTP relational data. The single most important mechanism in OLAP which allows it to achieve such performance is the use of aggregations. Aggregations are built from the fact table by changing the granularity on specific dimensions and aggregating up data along these dimensions. The number of possible aggregations is determined by every possible combination of dimension granularities.

The combination of all possible aggregations and the base data contains the answers to every query which can be answered from the data. Due to the potentially large number of aggregations to be calculated, often only a predetermined number are fully calculated while the remainder are solved on demand. The problem of deciding which aggregations (a.k.a. views) to calculate is known as the view selection problem. View selection can be constrained by the total size of the selected set of aggregations, the time to update them from changes in the base data, or both. The objective of view selection is typically to minimise the average time to answer OLAP queries, although some studies also minimise the update time as well. Many different approaches have been taken to view selection (which is NP-Complete), including greedy algorithms, randomised search, genetic algorithms and A* search algorithm.

A very effective way to support aggregation and other common OLAP operations is the use of bitmap indexes.

3.4 Types

OLAP systems have been traditionally categorised using the following taxonomy.

Multidimensional

MOLAP is the 'classic' form of OLAP and is sometimes referred to as just OLAP. MOLAP uses database structures that are generally optimal for attributes such as time period, location, product or account code. The way that each dimension will be aggregated is defined in advance by one or more hierarchies.

Relational

ROLAP works directly with relational databases. The base data and the dimension tables are stored as relational tables and new tables are created to hold the aggregated information. It depends on a specialised schema design.

Hybrid

There is no clear agreement across the industry as to what constitutes "Hybrid OLAP", except that a database will divide data between relational and specialised storage. For example, for some vendors, a HOLAP database will use relational tables to hold the larger quantities of detailed data, and use specialised storage for at least some aspects of the smaller quantities of more-aggregate or less-detailed data.

Comparison

Each type has certain benefits, although there is disagreement about the specifics of the benefits between providers.

Some MOLAP implementations are prone to database explosion. Database explosion is a phenomenon causing vast amounts of storage space to be used by MOLAP databases when certain common conditions are met: high number of dimensions, pre-calculated results and sparse multidimensional data. The typical mitigation technique for database explosion is not to materialise all the possible aggregation, but only the optimal subset of aggregations based on the desired performance vs. storage trade off.

MOLAP generally delivers better performance due to specialised indexing and storage optimisations. MOLAP also needs less storage space compared to ROLAP because the specialised storage typically includes compression techniques.

ROLAP is generally more scalable. However, large volume pre-processing is difficult to implement efficiently so it is frequently skipped. ROLAP query performance can therefore suffer.

Since ROLAP relies more on the database to perform calculations, it has more limitations in the specialised functions it can use.

HOLAP encompasses a range of solutions that attempt to mix the best of ROLAP and MOLAP. It can generally pre-process quickly, scale well, and offer good function support.

Other types

The following acronyms are also sometimes used, although they are not as widespread as the ones above:

•WOLAP - Web-based OLAP
•DOLAP - Desktop OLAP
•RTOLAP - Real-Time OLAP

3.5 APIs and Query Languages

Unlike relational databases, which had SQL as the standard query language, and widespread APIs such as ODBC, JDBC and OLEDB, there was no such unification in the OLAP world for a long time. The first real standard API was OLE DB for OLAP (ODBO) specification from Microsoft which appeared in 1997 and introduced the MDX query language. Several OLAP vendors - both server and client - adopted it. In 2001 Microsoft and Hyperion announced the XML for Analysis specification, which was endorsed by most of the OLAP vendors. Since this also used MDX as a query language, MDX became the de-facto standard.

3.6 Market Structure

Given below is a list of top OLAP vendors in 2006 with figures in millions of United States Dollars.

Vendor	**Global Revenue**
Microsoft Corporation	1,801
Hyperion Solutions Corporation	1,077
Cognos	735
Business Objects	416
MicroStrategy	416
SAP AG	330
Cartesis SA	210
Applix	205
Infor	199
Oracle Corporation	159
Others	152
Total	**5,700**

It should be noted that Microsoft was the only vendor that continuously exceeded the industrial average growth during 2000-6. Since the above data was collected, Hyperion has been acquired by Oracle, Business Objects by SAP, Applix by Cognos, and Cognos by IBM.

3.7 Online Transaction Processing,

Online Transaction Processing, or OLTP, refers to a class of systems that facilitate and manage transaction-oriented applications, typically for data entry and retrieval transaction processing. The term is somewhat ambiguous; some understand a "transaction" in the context of computer or database transactions, while others (such as the Transaction Processing Performance Council) define it in terms of business or commercial transaction OLTP has also been used to refer to processing in which the system responds immediately to user requests. An automatic teller machine (ATM) for a bank is an example of a commercial transaction processing application.
The technology is used in a number of industries, including banking, airlines, mailorder, supermarkets, and manufacturing. Applications include electronic banking, order processing, employee time clock systems, e-commerce, and eTrading. The most widely used OLTP system is probably IBM's CICS.

Requirements

Online transaction processing increasingly requires support for transactions that span a network and may include more than one company.For this reason, new OLTP software uses client/server processing and brokering software that allows transactions to run on different computer platforms in a network.

In large applications, efficient OLTP may depend on sophisticated transaction management software (such as CICS) and/or database optimisation tactics to facilitate the processing of large numbers of concurrent updates to an OLTP-oriented database.

For even more demanding decentralised database systems, OLTP brokering programs can distribute transaction processing among multiple computers on a network. OLTP is often integrated into SOA service-oriented architecture and Web services.

Benefits

Online Transaction Processing has two key benefits: simplicity and efficiency.

Reduced paper trails and the faster, more accurate forecasts for revenues and expenses are both examples of how OLTP makes things simpler for businesses. It also provides a concrete foundation for a stable organisation because of the timely updating. Another simplicity factor is that of allowing consumers the choice of how they want to pay, making it that much more enticing to make transactions.

OLTP is proven efficient because it vastly broadens the consumer base for an organisation, the individual processes are faster, and it's available 24/7.

Disadvantages

It is a great tool for any organisation, but in using OLTP, there are a few things to be wary of: the security issues and economic costs.

One of the benefits of OLTP is also an attribute to a potential problem. The worldwide availability that this system provides to companies makes their databases that much more susceptible to intruders and hackers.

For B2B transactions, businesses must go offline to complete certain steps of an individual process, causing buyers and suppliers to miss out on some of the efficiency benefits that the system provides. As simple as OLTP is, the simplest disruption in the system has the potential to cause a great deal of problems, causing a waste of both time and money. Another economic cost is the potential for server failures. This can cause delays or even wipe out an immeasurable amount of data.

3.8 Business Analytics

Business Analyticsis how organisations gather and interpret data in order to make better business decisions and to optimise business processes. Analytical activities are expanding fast in businesses, government agencies and not-for-profit organisations. Analytics are defined as the extensive use of data, statistical and quantitative analysis, explanatory and predictive modeling, and fact-based decision-making. Analytics may be used as input for human decisions; however in business there are also examples of fully automated decisions that require minimal human intervention. In businesses, analytics (alongside data access and reporting) represents a subset of business intelligence (BI).

The Age of Analytics may be thought of as a new subcategory of the Information Age, with the key difference that the early years of the Information Age were a time when information was a scarce resource, whereas today there is an abundance of information. The Age of

Analytics therefore represents another way to think of the activities necessary for success in a knowledge economy or increasingly typical of a modern information society.

Application of Analytics

Many organisations already use analytics in some form. Operating metrics and performance gauges such as the balanced scorecard are familiar to most managers. For instance, a manufacturer may track, interpret and use data to improve how it manages product quality, and a marketing group might base decisions on the long-term analysis of different customer segments. Businesses as diverseas global cement giant CEMEX, California winemaker E & J Gallo Winery, industrial equipment maker John Deere, retailer Tesco, and Bank of America are regularly applying analytics to achieve advantage. For example, Gallo, operating in a business built on using intuition to gauge unpredictable consumer preferences, now quantitatively analyses and predicts the appeal of its wines. Between 2002 and 2005, John Deere saved more than $1 billion by employing a new analytical tool to better optimise inventory.

However, only a handful of companies are using analytics as a foundation for their business strategies. Capital One is among those full-fledged analytical competitors. The financial services provider is very open about its use of data analysis to differentiate among customers based on credit risk, usage and other characteristics, and to match customer characteristics with appropriate product offerings. Harrahs, the world’s largest gaming firm, is another aggressive analytical competitor, particularly in the area of customer loyalty.

Research by global management consultancy Accenture found that high- performance businesses — those that substantially outperform competitors over the long term and across economic, industry and leadership cycles — are twice as likely to use analytics strategically compared with the overall sample, and five times more likely to do so than low performers.

Common to all those aspiring to that level of competitiveness is the need to focus on developing four fundamental assets:

- Committed senior executives: Taking a broad analytical approach to business calls for big changes in culture, process, behaviour and skills for many employees. Such changes must be spearheaded by senior executives who are passionate about analytics and fact-based decision-making.
- A strong base of skills in use of data: It is very important to have a

broad base of employees who are data-savvy — or who can quickly become data-savvy. This calls for hiring, training and rewarding for analytical skills, especially at management levels. It also highlights the need to understand where those skills matter most and where they will matter most in the future.

•Fact-driven business processes: Analytical competitors begin with "a single version of the truth" — not the conflicting views of the same metrics that stymie other companies. What's needed is an integrated, cross-enterprise view of the data — a state that may require business process redesign on a broad scale.

•Technology to capture, sort, and makesense of the data: The processing power to support an analytics thrust is readily available. There is wider use of dedicated "business intelligence appliances" — supercomputer — like machines that can quickly find and sort data in large databases and analyses. Much of the necessary analytical software is also available. "Real-time BI," in which automated decisions are embedded in operational business processes, is gaining ground.

History

Data analysis has been used in business since the dawn of the industrial era — from the time management exercises initiated by Frederick Winslow Taylor in the late 19th century to the measured pacing of the mechanised assembly lines developed by Henry Ford. But it began to command more attention in the late 1960s when computers were used in experiments to aid decision-making. These earliest "decision support systems" addressed repetitive and non-strategic activities such as financial reporting. (One notable exception was at American Airlines, which depended on Sabre, its breakthrough yield management system, to beat its competitors.) Analysis of statistics became more routine in the 1970s with the arrival of packaged computer applications. Bu few executives embraced the strategic use of data; number-crunching was left largely to the statisticians.

Since then, analytics have evolved with the development of enterprise resource planning (ERP) systems, data warehouses, and a wide variety of other hardware and software tools and applications. But until recently, companies have focused on analysing historical data rather than developing predictive analytics for decision-making.

Many companies today are collecting and storing a mind-boggling quantity of data. In just a few years, the common terminology for data volumes has grown from megabytes to gigabytes to terabytes (TB) — a trillion bytes. Some corporate databases are even approaching one petabyte — a quadrillion bytes — in size. The 583 TB in Wal-Mart's data warehouse, for example, is far more than the digita capacity needed if all 17 million of the books in the U.S. Library of Congress were fully formatted. Gargantuan storage facilities are not the only technological frontier: statistical software, high-end 64-bit processors, and specialty “data appliances” can quickly churn through enormous amounts of data—and do so with greater sophistication.

Challenges of analytics

Analytics is dependent on data. If there is no data, there can be no analytics. However, if data is sparse or non-existent, an organisation can conduct surveys or a census to obtain data. In many cases to save expenses, organisations can look for data obtained from situations that are similar but not quite meet the current requirements, and make minor modifications (i.e., a company trying to introduce an energy drink for the first time in a town can use existing data from a survey of athletes that drink carbonated beverages). However, in these cases, businesses should be aware of the risks inherent in using data obtained in such manner.

For many organisations aspiring to be analytical competitors, the primary problem is not that they lack data. It is that they must contend with dirty data. The challenge is that they do not know which data is trustworthy — “clean” — and which contains duplicates, outdated records and erroneous data entries.

According to Gartner, an alarming proportion of all business data is inaccurate. The research firm estimates that at least 25 percent of critical data within Fortune 1,000 companies will continue to be inaccurate through 2007. In a separate study by a leading accounting firm, only a little more than a third of executives were “very confident” in the quality of their corporate data.

A company that finds it has poor-quality data should postpone any plans to compete on analytics and instead should fix its data first. UPS demonstrates the patience that is often necessary. Although the delivery company has been collecting customer information for more than five years, it took more than half that time to validate that data before it was usable.

Unlike mature economies, in some of the fastest growing developing economies, such as India and China, analytics has to contend with "noisy" data wherein the data is incomplete (e.g. credit rating of a customer) or suspect (e.g. demographic information of a mobile telecom customer) or plain missing. A new generation of analytical algorithms that compensates for this "noise" appropriately helps in deployment of analytics solutions without the need to rely on fixing data - something that may never be possible in the near future. Such innovativ algorithms are an example of technology innovation in developing markets that promise to leapfrog existing methods tha have been developed primarily for mature markets and no easily transposable given constraints such as data sanctity.

4.0 CONCLUSION

Processing and analysing data is a very important aspect of decision support systems. In fact, it is the hallmark of DSS. So online analytic processing, online transaction processing and business analytics are important tools in decision support systems. These tools collectively and individually have their various advantages and disadvantages that need to be explored and reevaluated. The electronic nature of the analysis has made them more supportive to decision support systems.

5.0 SUMMARY

- Online Analytical Processing, or OLAP (IPA), is an approach to quickly provide answers to analytical queries that are multi-dimensional in nature. OLAP is part of the broader category business intelligence, which also encompasses relational reporting and data mining.
- At the core of any OLAP system is a concept of an OLAP cube (also called a multidimensional cube or a hypercube). It consists of numeric facts called measures, which are categorised by dimensions.
- It has been claimed that for complex queries OLAP cubes can produce an answer in around 0.1% of the time for the same query on OLTP relational data.
- OLAP systems have been traditionally categorised using different taxonomy.
- Unlike relational databases, which had SQL as the standard query language, and wide-spread APIs such as ODBC, JDBC and OLEDB, there was no such unification in the OLAP world for a long time.
- Online Transaction Processing, or OLTP, refers to a class of systems that facilitate and manage transaction-oriented applications, typically for data entry and retrieval transaction processing.
- Business Analytics is how organisations gather and interpret data in order to make better business decisions and to optimise business processes. Analytical activities are expanding fast in businesses, government agencies and not-for-profit organisations.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention 5 types of online analytic processing (OLAP).
2. Discuss briefly the requirements of online processing transactions (OLTP).

7.0 REFERENCES/FURTHER READING

Codd E.F., Codd S.B., and Salley C.T. (1993). Providing OLAP (On-*Line Analytical Processing) to User-Analysts: An IT Mandate (in English). Codd & Date, Inc.*

Deepak Pareek (2007). Business Intelligence for Telecommunications (in English). CRC Press, 294 pp.

OLAP Council White Paper (in English). OLAP Council (1997). Retrieved on 2008-03-18.

Hari Mailvaganam (2007). Introduction to OLAP - Slice, Dice and *Drill! (in English). Data Warehousing Review.*

Nigel Pendse (2008-03-03). What is OLAP? An Analysis of what the *Often Misused OLAP Term is Supposed to Mean(in English).* OLAP Report.

MicroStrategy, Incorporated (1995). The Case for Relational OLAP (in English).

Surajit Chaudhuri and Umeshwar Dayal (1997). An overview of data *warehousing and OLAP technology (in English). SIGMOD Rec.* **26: 65. ACM. doi:10.1145/248603.248616. ISSN 0163-5808.**

Data Cube: {A} Relational Aggregation Operator Generalising Group-By, Cross-Tab, and Sub-Totals (in English) (1997). J. Data *Mining and Knowledge Discovery 1 (1): pp. 29–53.*

Nigel Pendse (2006-06-27). OLAP Architectures (in English). OLAP *Report.*

Nigel Pendse (2007-08-23). Commentary: OLAP API wars (in English). *OLAP Report.*

Nigel Pendse (2007-08-23). The Origins of Today's OLAP Products (in *English). OLAP Report.*

Nigel Pendse (2006). OLAP Marke (in English). OLAP Report.
Nigel Pendse (2008-03-07). Consolidations in the BI industry (in English).

Transaction Processing Performance Council website.

Davenport, Thomas H.; Jeanne G. Harris (March 2007). Competing on *Analytics: The New Science of Winning. Harvard Business* School Press.

McDonald, Mark; Tina Nunno (February 2007). Creating Enterprise *Leverage: The 2007 CIO Agenda. Stamford, CT: Gartner, Inc.*

Baker, Stephen (January 23, 2006). Math Will Rock Your World. BusinessWeek..

Davenport, Thomas H. (January 1, 2006). Competing on Analytics. Harvard Business Review.

Pfeffer, Jeffrey; Robert Sutton (January 2006). Evidence-Based *Managemen. Harvard Business Review.*

Davenport, Thomas H.; Jeanne G. Harris (Summer 2005). Automated *Decision Making Comes of Age. MIT Sloan Management Review.*

Lewis, Michael (April 2004). Moneyball: The Art of Winning an Unfair *Game. W.W. Norton & Co..*

Bonabeau, Eric (May 2003). Don't Trust Your Gut. Harvard Business Review.

Davenport, Thomas H.; Jeanne G. Harris, David W. De Long, Alvin L. Jacobson. Data to Knowledge to Results: Building an Analytic *Capability. California Management Review 43 (2): 117–138.*

Ranadive, Vivek (2006-01-26). The Power to Predict: How Real Time *Businesses Anticipate Customer Needs, Create Opportunities, and Beat the Competition. McGraw-Hill.*

Zabin, Jeffrey; Gresh Brebach (February 2004). Precision Marketing. John Wiley.

UNIT 2 PREDICTIVE ANALYTICS

CONTENTS

1.0 INTRODUCTION

Predictive analytics encompasses a variety of techniques from statistics and data mining that analyse current and historical data to make predictions about future events. Such predictions rarely take the form of absolute statements, and are more likely to be expressed as values that correspond to the odds of a particular event or behaviour taking place in the future.

Predictive analytics is an area of statistical analysis that deals with extracting information from data and using it to predict future trends and behaviour patterns. The core of predictive analytics relies on capturing relationships between explanatory variables and the predicted variables from past occurrences, and exploiting it to predict future outcomes.

In business, predictive models exploit patterns found in historical and transactional data to identify risks and opportunities. Models capture relationships among many factors to allow assessment of risk or potential associated with a particular set of conditions, guiding decision making for candidate transactions.

One of the most well-known applications is credit scoring, which is used throughout financial services. Scoring models process a customer's credit history, loan application, customer data, etc., in order to rank-order individuals by their likelihood of makingfuture credit payments on time. Predictive analytics are also used in insurance, telecommunications, retail, travel, healthcare, pharmaceuticals and other fields.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define predictive analytics
- identify the different types of predictive analytics
- explain different ways of applying predictive analysis in decision making in life and business
- identify, compare and contrast the various techniques associated with predictive analytics
- name some of the popular tools used in predictive analytics.

3.0 MAIN CONTENT

3.1 Types of Predictive Analytics

Generally, predictive analytics is used to mean predictive modelling, scoring of predictive models, and forecasting. However, people are increasingly using the term to describe related analytic disciplines, such as descriptive modelling and decision modelling or optimisation. These disciplines also involve rigorous data analysis, and are widely used in business for segmentation and decision making but have different purposes and the statistical techniques underlying them vary.

Predictive Models

Predictive models analyse past performance to assess how likely a customer is to exhibit a specific behaviour in the future in order to improve marketing effectiveness. This category also encompasses models that seek out subtle data patterns to answer questions about customer performance, such as fraud detection models. Predictive models often perform calculations during live transactions, for example, to evaluate the risk or opportunity of a given customer or transaction, in order to guide a decision.

Descriptive Models

Descriptive models "describe" relationships in data in a way that is often used to classify customers or prospects into groups. Unlike predictive models that focus on predicting a single customer behaviour (such as credit risk), descriptive models identify many different relationships between customers or products. But the descriptive models do not rank-order customers by their likelihood of taking a particular action the way predictive models do. Descriptive models are often used "offline," for example, to categorise customers by their product preferences and life stage. Descriptive modelling tools can be utilised to develop agent based models that can simulate large number of individualised agents to predict possible futures.

Decision Models

Decision models describe the relationship between all the elements of a decision — the known data (including results of predictive models), the decision and the forecast results of the decision — in order to predict the results of decisions involving many variables. These models can be used in optimisation, a data-driven approach to improving decision logic that involves maximising certain outcomes while minimising others. Decision models are generally used offline, to develop decision logic or a set of business rules that will produce the desired action for every customer or circumstance.

3.2 Current Uses

Although predictive analytics can be put to use in many applications, we outline a few examples where predictive analytics has shown positive impact in recent years.

Analytical Customer Relationship Management (CRM)

Analytical Customer Relationship Management is a frequent commercial application of Predictive Analysis. Methods of predictive analysis are applied to customer data to pursue CRM objectives.

Direct Marketing

Product marketing is constantly faced with the challenge of coping with the increasing number of competing products, different consumer preferences and the variety of methods (channels) available to interact with each consumer. Efficient marketing is a process of understanding the amount of variability and tailoring the marketing strategy for greater profitability.

Predictive analytics can help identify consumers with a higher likelihood of responding to a particular marketing offer. Models can be built using data from consumers' past purchasing history and past response rates for each channel. Additional information about the consumers demographic, geographic and other characteristics can be used to make more accurate predictions. Targeting only these consumers can lead to substantial increase in response rate which can lead to a significant reduction in cost per acquisition. Apart from identifying prospects, predictive analytics can also help to identify the most effective combination of products and marketing channels that should be used to target a given consumer.

Cross-Sell

Often corporate organisations collect and maintain abundant data (e.g. customer records, sale transactions) and exploiting hidden relationships in the data can provide a competitive advantage to the organisation. For an organisation that offers multiple products, an analysis of existing customer behaviour can lead to efficient cross sell of products. This directly leads to higher profitability per customer and strengthening of the customer relationship. Predictive analytics can help analyse customers' spending, usage and other behaviour, and help cross-sell the right product at the right time.

Customer Retention

With the amount of competing services available, businesses need to focus efforts on maintaining continuous consumer satisfaction. In such a competitive scenario, consumer loyalty needs to be rewarded and customer attrition needs to be minimised. Businesses tend to respond to customer attrition on a reactive basis, acting only after the customer has initiated the process to terminate service. At this stage, the chance of changing the customer's decision is almost impossible. Proper application of predictive analytics can lead to a more proactive retention strategy. By a frequent examination of a customer's past service usage, service performance, spending and other behaviour patterns, predictive models can determine the likelihood of a customer wanting to terminate service sometime in the near future. An intervention with lucrative offers can increase the chance of retaining the customer. Silent attrition is the behaviour of a customer to slowly but steadily reduce usage and is another problem faced by many companies. Predictive analytics can also predict this behaviour accurately and before it occurs, so that the company can take proper actions to increase customer activity.

Underwriting

Many businesses have to account for risk exposure due to their different services and determine the cost needed to cover the risk. For example, auto insurance providers need to accurately determine the amount of premium to charge to cover each automobile and driver. A financial company needs to assess a borrower's potential and ability to pay before granting a loan. For a health insurance provider, predictive analytics can analyse a few years of past medical claims data, as well as lab, pharmacy and other records where available, to predict how expensive an enrollee is likely to be in the future. Predictive analytics can help underwriting of these quantities by predicting the chances of illness, default, bankruptcy, etc. Predictive analytics can streamline the process of customer acquisition, by predicting the future risk behaviour of a customer using application level data. Proper predictive analytics can lead to proper pricing decisions, which can help mitigate future risk of default.

Collection Analytics

Every portfolio has a set of delinquent customers who do not make their payments on time. The financial institution has to undertake collection activities on these customers to recover the amounts due. A lot of collection resources are wasted on customers who are difficult or impossible to recover. Predictive analytics can help optimise the allocation of collection resources by identifying the most effective collection agencies, contact strategies, legal actions and other strategies to each customer, thus significantly increasing recovery at the same time reducing collection costs.

Fraud Detection

Fraud is a big problem for many businesses and can be of various types. Inaccurate credit applications, fraudulent transactions, identity thefts and false insurance claims are some examples of this problem. These problems plague firms all across the spectrum and some examples of likely victims are credit card issuers, insurance companies, retail merchants, manufacturers, business to business suppliers and even services providers. This is an area where a predictive model is often used to help weed out the "bads" and reduce a business's exposure to fraud.

Portfolio, Product ~~Or~~or Economy Level Prediction

Often the focus of analysis is not the consumer but the product, portfolio, firm, industry or even the economy. For example a retailer might be interested in predicting store level demand for inventory management purposes. Or the Federal Reserve Board might be interested in predicting the unemployment rate for the next year. These types of problems can be addressed by predictive analytics using Time Series techniques

3.3 Statistical Techniques

The approaches and techniques used to conduct predictive analytics can broadly be grouped into regression techniques and machine learning techniques.

Regression Techniques

Regression models are the mainstay of predictive analytics. The focus lies on establishing a mathematical equation as a model to represent the interactions between the different variables in consideration. Depending on the situation, there is a wide variety of models that can be applied while performing predictive analytics. Some of them are briefly discussed below.

Linear Regression Model

The linear regression model analyses the relationship between the response or dependent variable and a set of independent or predictor variables. This relationship is expressed as an equation that predicts the response variable as a linear function of the parameters. These parameters are adjusted so that a measure of fit is optimised. Much of the effort in model fitting is focused on minimising the size of the residual, as well as ensuring that it is randomly distributed with respect to the model predictions.

The goal of regression is to select the parameters of the model so as to minimise the sum of the squared residuals. This is referred to as **ordinary least squares (OLS) estimation and results in best linear** unbiased estimates (BLUE) of the parameters if and only if the Gauss-Markowitz assumptions are satisfied.

Once the model has been estimated we would be interested to know if the predictor variables belong in the model – i.e. is the estimate of each variable’s contribution reliable? To do this we can check the statistical significance of the model’s coefficients which can be measured using the t-statistic. This amounts to testing whether the coefficient is significantly different from zero. How well the model predicts the dependent variable based on the value of the independent variables can be assessed by using the R^2 statistic. It measures predictive power of the model i.e. the proportion of the total variation in the dependent variable that is “explained” (accounted for) by variation in the independent variables.

Discrete Choice Models

Multivariate regression (above) is generally used when the response variable is continuous and has an unbounded range. Often the response variable may not be continuous but rather discrete. While mathematically it is feasible to apply multivariate regression to discrete ordered dependent variables, some of the assumptions behind the theory of multivariate linear regression no longer hold, and there are other techniques such as discrete choice models which are better suited for this type of analysis. If the dependent variable is discrete, some of those superior methods are logistic regression, multinomial logit and probit models. Logistic regression and probit models are used when the dependent variable is binary.

Logistic Regression

In a classification setting, assigning outcome probabilities to observations can be achieved through the use of a logistic model, which is basically a method which transforms information about the binary dependent variable into an unbounded continuous variable and estimates a regular multivariate model (See Allison's Logistic Regression for more information on the theory of Logistic Regression).
The Wald and likelihood-ratio test are used to test the statistical significance of each coefficient b in the model (analogueous to the t tests used in OLS regression; see above). A test assessing the goodness-of-fit of a classification model is the Hosmer and Lemeshow test.

Multinomial Logistic Regression

An extension of the binary logit model to cases where the dependent variable has more than 2 categories is the multinomial logit model. In such cases collapsing the data into two categories might not make good sense or may lead to loss in the richness of the data. The multinomial logit model is the appropriate technique in these cases, especially when the dependent variable categories are not ordered (for examples colors like red, blue, green). Some authors have extended multinomial regression to include feature selection/importance methods such as Random multinomial logit.

Probit Regression

Probit models offer analternative to logistic regression for modelling categorical dependent variables. Even though the outcomes tend to be similar, the underlying distributions are different. Probit models are popular in social sciences like economics.
A good way to understand the key difference between probit and logit models, is to assume that there is a latent variable z.
We do not observe z but instead observe y which takes the value 0 or 1.
In the logit model we assume that follows a logistic distribution. In the

probit model we assume that follows a standard normal distribution. Note that in social sciences (example economics), probit is often used to model situations where the observed variable y is continuous but takes values between 0 and 1.

Logit vs. Probit

The Probit model has been around longer than the logit model. They look identical, except that the logistic distribution tends to be a little flat tailed. In fact one of the reasons the logit model was formulated was that the probit model was extremely hard to compute because it involved calculating difficult integrals. Modern computing however has made this computation fairly simple. The coefficients obtained from the logit and probit model are also fairly close. However the odds ratio makes the logit model easier to interpret.

For practical purposes the only reasons for choosing the probit model over the logistic model would be:

- There is a strong belief that the underlying distribution is normal
- The actual event is not a binary outcome (e.g. Bankrupt/not bankrupt) but a proportion (e.g. Proportion of population at different debt levels).

Time Series Models

Time series models are used for predicting or forecasting the future behaviour of variables. These models account for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for. As a result standard regression techniques cannot be applied to time series data and methodology has been developed to decompose the trend, seasonal and cyclical component of the series. Modelling the dynamic path of a variable can improve forecasts since the predictable component of the series can be projected into the future.

Time series models estimate difference equations containing stochastic components. Two commonly used forms of these models are autoregressive models (AR) and moving average (MA) models. The Box-Jenkins methodology (1976) developed by George Box and G.M. Jenkins combines the AR and MA models to produce the ARIMA (autoregressive moving average) model which is the cornerstone of stationary time series analysis. ARIMA (autoregressive integrated moving average models) on the other hand are used to describe non- stationary time series. Box and Jenkins suggest differencing a non stationary time series to obtain a stationary series to which an ARMA

model can be applied. Non stationary time series have a pronounced trend and do not have a constant long-run mean or variance.

Box and Jenkins proposed a three stage methodology which includes: model identification, estimation and validation. The identification stage involves identifying if the series is stationary or not and the presence of seasonality by examining plots of the series, autocorrelation and partial autocorrelation functions. In the estimation stage, models are estimated using non-linear time series or maximum likelihood estimation procedures. Finallythe validation stage involves diagnostic checking such as plotting the residuals to detect outliers and evidence of model fit.

In recent years time series models have become more sophisticated and attempt to model conditional heteroskedasticity with models such as ARCH (autoregressive conditional heteroskedasticity) and GARCH (generalised autoregressive conditional heteroskedasticity) models frequently used for financial time series. In addition time series models are also used to understand inter-relationships among economic variables represented by systems of equations using VAR (vector autoregression) and structural VAR models.

Survival or Duration Analysis

Survival analysis is another name for time to event analysis. These techniques were primarily developedin the medical and biological sciences, but they are also widely used in the social sciences like economics, as well as in engineering (reliability and failure time analysis).

Censoring and non-normality which are characteristic of survival data generate difficulty when trying to analyse the data using conventional statistical models such as multiple linear regression. The Normal distribution, being a symmetric distribution, takes positive as well as negative values, but duration by its very nature cannot be negative and therefore normality cannot be assumed when dealing with duration/survival data. Hence the normality assumption of regression models is violated.

A censored observation is defined as an observation with incomplete information. Censoring introduces distortions into traditional statistical methods and is essentially a defect of the sample data. The assumption is that if the data were not censored it would be representative of the population of interest. In survival analysis, censored observations arise whenever the dependent variable of interest represents the time to a terminal event, and the duration of the study is limited in time.

An important concept in survival analysis is the hazard rate. The hazard rate is defined as the probability that the event will occur time t conditional on surviving until time t. Another concept related to the hazard rate is the survival function which can be defined as the probability of surviving to time t.

Most models try to model the hazard rate by choosing the underlying distribution depending on the shape of the hazard function. A distribution whose hazard function slopes upward is said to have positive duration dependence, adecreasing hazard shows negative duration dependence whereas constant hazard is a process with no memory usually characterised by the exponential distribution. Some of the distributional choices in survival models are: F, gamma, Weibull, log normal, inverse normal, exponential etc. All these distributions are for a non-negative random variable.

Duration models can be parametric, non-parametric or semi-parametric. Some of the models commonly used are Kaplan-Meier, Cox proportional hazard model (non parametric).

Classification and Regression Trees

Classification and regression trees (CART) is anon-parametric technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively.

Trees are formed by a collection of rules based on values of certain variables in the modelling data set:

- Rules are selected based on how well splits based on variables' values can differentiate observations based on the dependent variable
- Once a rule is selected and splits a node into two, the same logic is applied to each "child" node (i.e. it is a recursive procedure)
- Splitting stops when CART detects no further gain can be made, or some pre-set stopping rules are met.

Each branch of the tree ends in a terminal node;

- Each observation falls into one and exactly one terminal node
- Each terminal node is uniquely defined by a set of rules.

A very popular method for predictive analytics is Leo Breiman's Random forests or derived versions of this technique like Random multinomial logit.

Multivariate Adaptive Regression Splines

Multivariate adaptive regression splines (MARS) is a non-parametric technique that builds flexible models by fitting piecewise linear regressions.
An important concept associated with regression splines is that of a knot. Knot is where one local regression model gives way to another and thus is the point of intersection between two splines.

In multivariate and adaptive regression splines, basis functions are the tool used for generalising the search for knots. Basis functions are a set of functions used to represent the information contained in one or more variables. Multivariate and Adaptive Regression Splines model almost always creates the basis functions in pairs.
Multivariate and adaptive regression spline approach deliberately overfits the model and then prunes to get to the optimal model. The algorithm is computationally very intensive and in practice we are required to specify an upper limit on the number of basis functions.

Machine Learning Techniques

Machine learning, a branch of artificial intelligence, was originally employed to develop techniques to enable computers to learn. Today, since it includes a number of advanced statistical methods for regression and classification, it finds application in a wide variety of fields including medical diagnostics, credit card fraud detection, face and speech recognition and analysis of the stock market. In certain applications it is sufficient to directly predict the dependent variable without focusing on the underlying relationships between variables. In other cases, the underlying relationships can be very complex and the mathematical form of the dependencies unknown. For such cases, machine learning techniques emulate human cognition and learn from training examples to predict future events.

A brief discussion of some of these methods used commonly for predictive analytics is provided below. A detailed study of machine learning can be found in Mitchell (1997).

3.4 Neural Networks

Neural networks are nonlinear sophisticated modelling techniques that are able to model complex functions. They can be applied to problems of prediction, classification or control in a wide spectrum of fields such as finance, cognitive psychology/neuroscience, medicine, engineering, and physics.

Neural networks are used when the exact nature of the relationship between inputs and output is not known. A key feature of networks is that they learn the relationship between inputs and output through training. There are two types of training in neural networks used by different networks, supervised and unsupervised training, with supervised being the most common one.

Some examples of neural network training techniques are backpropagation, quick propagation, conjugate gradient descent, projection operator, Delta-Bar-Delta etc. Theses are applied to network architectures such as multilayer perceptrons,Kohonen networks, Hopfield networks, etc.

3.5 Radial Basis Functions

A radial basis function (RBF) is a function which has built into it a distance criterion with respect to a centre. Such functions can be used very efficiently for interpolation and for smoothing of data. Radial basis functions have been applied in the area of neural networks where they are used as a replacement for the sigmoidal transfer function. Such networks have 3 layers, the input layer, the hidden layer with the RBF non-linearity and a linear output layer. The most popular choice for the non-linearity is the Gaussian. RBF networks have the advantage of not being locked into local minima as do the feed-forward networks such as the multilayer perceptron.

3.6 Support Vector Machines

Support Vector Machines (SVM) are used to detect and exploit complex patterns in data by clustering, classifying and ranking the data. They are learning machines that are used to perform binary classifications and regression estimations. They commonly use kernel based methods to apply linear classification techniques to non-linear classification problems. There are a number of types of SVM such as linear, polynomial, sigmoid etc.

3.7 Naïve Bayes

Naïve Bayes based on Bayes conditional probability rule is used for performing classification tasks. Naïve Bayes assumes the predictors are statistically independent which makes it an effective classification tool that is easy to interpret. It is best employed when faced with the problem of ‘curse of dimensionality’ i.e. when the number of predictors is very high.

3.8 K-Nearest Neighbours

The nearest neighbour algorithm (KNN) belongs to the class of pattern recognition statistical methods. The method does not impose a priori any assumptions about the distribution from which the modelling sample is drawn. It involves a training set with both positive and negative values. A new sample is classified by calculating the distance to the nearest neighbouring training case. The sign of that point will determine the classification of the sample. In the k-nearest neighbour classifier, the k nearest points are considered and the sign of the majority is used to classify the sample. The performance of the kNN algorithm is influenced by three main factors: (1) the distance measure used to locate the nearest neighbours; (2) the decision rule used to derive a classification from the k-nearest neighbours; and (3) the number of neighbours used to classify the new sample. It can be proved that, unlike other methods, this method is universally asymptotically convergent, i.e.: as the size of the training set increases, if the observations are iid, regardless of the distribution from which the sample is drawn, the predicted class will converge to the class assignment that minimises misclassification error. See Devroy et alt.

3.9 Popular Tools

There are numerous tools available in the marketplace which helps with the execution of predictive analytics. These range from those which need very little user sophistication to those that are designed for the expert practitioner. The difference between these tools is often in the level of customisation and heavy data lifting allowed. For traditional statistical modelling some of the populartools are DAP/SAS, S-Plus, PSPP/SPSS and Stata. For machine learning/data mining type of applications, KnowledgeSEEKER, KnowledgeSTUDIO, Enterprise Miner, GeneXproTools, Viscovery, Clementine, KXEN Analytic Framework, InforSense and Excel Miner are some of the popularly used options. Classification Tree analysis can be performed using CART software. SOMine is a predictive analytics tool based on self-organising maps (SOMs) available from Viscovery Software. R is a very powerful tool that can be used to perform almost any kind of statistical analysis, and is freely downloadable. WEKA is a freely available open-source collection of machine learning methods for pattern classification, regression, clustering, and some types of meta-learning, which canbe used for predictive analytics. RapidMiner is another freely available integrated open-source softwareenvironment for predictive analytics, data mining, and machine learning fully integrating WEKA and providing an even larger number of methods for predictive analytics.

Recently, in an attempt to provide a standard language for expressing predictive models, the Predictive Model Markup Language (PMML) has been proposed. Such an XML-based language provides a way for the different tools to define predictive models and to share these between PMML compliant applications. Several tools already produce or consume PMML documents, these include ADAPA, IBMDB2 Warehouse, CART, SAS Enterprise Miner, and SPSS. Predictive analytics has also found its way into the IT lexicon, most notably in the area of IT Automation. Vendors such as Stratavia and their Data Palette product offer predictive analytics as part of their automation platform, predicting how resources will behave in the future and automate the environment accordingly.

The widespread use of predictive analytics in industry has led to the proliferation of numerous productised solutions firms. Some of them are highly specialised (focusing, for example, on fraud detection, automatic sales lead generation or response modelling) in a specific domain (Fair Isaac for credit card scores) or industry verticals (MarketRx in Pharmaceutical). Others provide predictive analytics services in support of a wide range of business problems across industry verticals (Fifth C). Predictive Analytics competitions are also fairly common and often pit academics and Industry practitioners (see for example, KDD CUP).

4.0 CONCLUSION

Predictive analytics adds great value to businessesto businesses decision making capabilities by allowing them to formulate smart policies on the basis of predictions of future outcomes. A broad range of tools and techniques are available for this type of analysis and their selection is determined by the analytical maturity of the firm as well as the specific requirements of the problem being solved.

5.0 SUMMARY

- Predictive analytics encompasses a variety of techniques from statistics and data mining that analyse current and historical data to make predictions about future events.
- Generally, predictive analytics is used to mean predictive modelling, scoring of predictive models, and forecasting. However, people are increasingly using the term to describe related analytic disciplines, such as descriptive modelling and decision modelling or optimisation.
- Although predictive analytics can be put to use in many applications, we outline a few examples where predictive analytics has shown positive impact in recent years.
- The approaches and techniques used to conduct predictive analytics can broadly be grouped into regression techniques and machine learning techniques.

- Neural networks are nonlinear sophisticated modelling techniques that are able to model complex functions.
- A radial basis function (RBF) is a function which has built into it a distance criterion with respect to a centre
- Support Vector Machines (SVM) are used to detect and exploit complex patterns in data by clustering, classifying and ranking the data.
- Naïve Bayes based on Bayes conditional probability rule is used for performing classification tasks.
- The nearest neighbour algorithm (KNN) belongs to the class of pattern recognition statistical methods.
- There are numerous tools available in the marketplace which helps with the execution of predictive analytics. These range from those which need very little user sophistication to those that are designed for the expert practitioner.

6.0 TUTOR-MARKED ASSIGNMENT

8. List 5 uses of predictive analysis
9. Briefly discuss Classification and Regression Trees (CART)

7.0 REFERENCES/FURTHER READING

L. Devroye, L. Györfi, G. Lugosi (1996). A Probabilistic Theory of *Pattern Recognition. New York: Springer-Verlag.*

John R. Davies, Stephen V. Coggeshall, Roger D. Jones, and Daniel Schutzer, Intelligent Security Systems, in Freedman, Roy S., Flein, .

Robert A., and Lederman, Jess, Editors (1995). Artificial Intelligence in *the Capital Markets. Chicago: Irwin. ISBN 1-55738-811-3.*

Agresti, Alan (2002). Categorical Data Analysis. Hoboken: John Wiley and Sons. ISBN 0-471-36093-7.

Enders, Walter (2004). Applied Time Series Econometrics. Hoboken: John Wiley and Sons. ISBN: 052183919X.

Greene, William (2000). Econometric Analysis. London: Prentice Hall. ISBN: 0-13-013297-7.

Mitchell, Tom (1997). Machine Learning. New York: McGraw-Hill. ISBN 0-07-042807-7

Tukey, John (1977). Exploratory Data Analysis. New York: Addison-Wesley. ISBN: 0.

UNIT 3 SPREADSHEET

CONTENTS

1.0 INTRODUCTION

A spreadsheet is a computer application that simulates a paper worksheet. It displays multiple cells that together makeup a grid consisting of rows and columns, each cell containing either alphanumeric text or numeric values. A spreadsheet cell may alternatively contain a formula that defines how the contents of that cell is to be calculated from the contents of any other cell (or combination of cells) each time any cell is updated. Spreadsheets are frequently used for financial information because of their ability to re-calculate the entire sheet automatically after a change to a single cell is made.

In 1971, Rene K. Pardo and Remy Landau filed a patent on a spreadsheet related algorithm. Visicalc is usually considered the first electronic spreadsheet (although this has been challenged), and it helped turn the Apple II computer into a success and greatly assisted in their widespread application. Lotus 1-2-3 was the leading spreadsheet of DOS era. Excel is now generally considered to have the largest market share.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

• define spreadsheet

- evaluate the history and development of spreadsheet
- identify types of spreadsheet and their uniqueness
- describe the concepts of cell, sheets, and charts which are the major structure of a spreadsheet
- explain the evolution of languages to deal with changes in spreadsheets
- identify the challenges facing spreadsheets
- identify the applications of spreadsheets as decision making tools.

3.0 MAIN CONTENT

3.1 History

Paper Spreadsheets

The word "spreadsheet" came from "spread" in its sense of a newspaper or magazine item (text and/or graphics) that covers two facing pages, extending across the centre fold and treating the two pages as one large one. The compound word "spread-sheet" came to mean the format used to present book-keeping ledgers—with columns for categories of expenditures across the top, invoices listed down the left margin, and the amount of each payment in the cell where its row and column intersect —which were, traditionally, a "spread" across facing pages of a bound ledger (book for keeping accounting records) or on oversized sheets of paper ruled into rows and columns in that format and approximately twice as wide as ordinary paper.

Early Implementations

Batch spreadsheet report generator: The concept of an electronic spreadsheet was outlined in the 1961 paper "Budgeting Models and System Simulation" by Richard Mattessich. The subsequent work by Mattessich (1964a, Chpt. 9, Accounting and Analytical Methods) and its companion volume, Mattessich (1964b, Simulation of the Firm through a Budget Computer Program) applied computerised spreadsheets to accounting and budgeting systems (on main-frame computers in FORTRAN IV). Batch Spreadsheets dealt primarily with the addition or subtraction of entire columns or rows - rather than individual cells.

LANPAR spreadsheet compiler: Key invention in the development of electronic spreadsheets was made by Rene K. Pardo and Remy Landau, who filed in 1971 U.S. Patent ~~4,398,249~~ on4,398,249 on spreadsheet automatic natural order recalculation algorithm in 1970. While the patent was initially rejected by the patent office as being apurely mathematical invention, following 12 years of Appeals, Pardo and

Landau won a landmark court case at the CCPA (Predecessor Court of the Federal Circuit) overturning the Patent Office in 1983 with the world's first software patent - establishing that "something does not cease to become patentable merely because the point of novelty is in an algorithm." This case helped establish the viability of software patents. The actual software was called LANPAR - LANguage for Programming Arrays at Random. This was conceived and entirely developed in the summer of 1969 following Pardo and Landau's recent graduation from Harvard University. Co-inventor Rene Pardo recalls that he felt that one manager at Bell Canada should not have to depend on programmers to program and modify budgeting forms, and he thought of letting users type out forms in any order and having computer calculating results in the right order. The software was developed in 1969.

LANPAR was used by Bell Canada, AT&T and the 18 operating telcos nationwide for their local and national budgeting operations. LANPAR was also used by General Motors. Its uniqueness was the incorporation of natural order recalculation, as opposed to left-to-right, top to bottom sequence for calculating the results in each cell that was used by Visicalc, Supercalc and the first version of Multiplan. Without natural order recalculation the users had to manually recalculate the spreadsheet as many times as necessary until the values in all the cells had stopped changing.

The LANPAR system was implemented on GE400 and Honeywell 6000 on line timesharing systems enabling users to programremotely via computer terminals and modems. Data could be entered dynamically either by paper tape, specific file access, on line, or even external data bases. Sophisticated mathematical expressions including logical comparisons and "if/then" statements could be used in any cell, and cells could be presented in any order.

Autoplan/Autotab spreadsheet programming language: In 1968, three former employees from the General Electric computer company headquartered in Phoenix, Arizona set out to start their own software development house. A. Leroy Ellison, Harry N. Cantrell, and Russell E. Edwards found themselves doing a large number of calculations when making tables for the business plans that they were presenting to venture capitalists. They decided to save themselves a lot of effort and wrote a computer program that produced their tables for them. This program, originally conceived as a simple utility for their personal use, would turn out to be the first software product offered by the company that would become known as Capex Corporation. "AutoPlan" ran on GE’s Time-sharing service; afterward, a version that ran on IBM mainframes was introduced under the name "AutoTab". (National CSS offered a similar product, CSSTAB, which had a moderate timesharing user base by the early 70s. A major application was opinion research tabulation.) AutoPlan/AutoTab was not a WYSIWYG interactive spreadsheet program; it was a simple scripting language for spreadsheets. The user defined the names and labels for the rows and columns, then the formulas that defined each row or column.

Interactive Spreadsheets: Interactive spreadsheets became possible when computer displays became available. Except for LANPAR- which allowed many users in real time and online to timesharing systems to simultaneously program or run their spreadsheets from their computer terminals across the country - earlier implementations were mainly designed around batch programs and certainly none of these batch programs allowed for forward referencing of cells or natural order recalculation. In the early 1970s text based computer displays began to be used as input/output devices for interactive transaction processes. This was sufficient for early interactive spreadsheets to be implemented. The lack of on-line historical material relating to such systems, and their limited coverage in academic and commercial publications, makes it hard to assess their level of innovation and ultimate impact.

APLDOT Modelling Language

An example of an early "industrial weight" spreadsheet was APLDOT, developed in 1976 at the United States Railway Association on an IBM 360/91, running at The Johns Hopkins University Applied Physics Laboratory in Laurel, MD. The application was used successfully for many years in developing such applications as financial and costing models for the US Congress and for Conrail. APLDOT was dubbed a "spreadsheet" because financial analysts and strategic planners used it to solve the same problems they addressed with paper spreadsheet pads.

The spreadsheet concept became widely known in the late 1970s and early 1980s because of Dan Bricklin's implementation of VisiCalc. VisiCalc was the first spreadsheet that combined all essential features of modern spreadsheet applications, such as WYSIWYG interactive user interface, automatic recalculation, status and formula lines, range copying with relative and absolute references, formula building by selecting referenced cells. PC World magazine has called VisiCalc the first electronic spreadsheet.

Bricklin has spoken of watching his university professor create a table of calculation results on a blackboard. When the professor found an error, he had to tediously erase and rewrite a number of sequential entries in the table, triggering Bricklin to think that he could replicate the process on a computer, using the blackboard as the model to view results of underlying formulas. His idea became VisiCalc, the first application that turned the personal computer from a hobby for computer enthusiasts into a business tool.

Screenshot of VisiCalc, the First PC Spreadsheet

VisiCalc went on to become the first "killer app", an application that was so compelling: people would buy a particular computer just to own it. In this case the computer was the Apple II, and VisiCalc was no small part in that machine's success. The program was later ported to a number of other early computers, notably CP/M machines, the Atari 8-bit family and various Commodore platforms. Nevertheless, VisiCalc remains best known as "an Apple II program".

Lotus 1-2-3 and Other MS-DOS Spreadsheets

The acceptance of the IBM PC following its introduction in August, 1981, began slowly, because most of the programs available for it were ports from other 8-bit platforms. Things changed dramatically with the introduction of Lotus 1-2-3 in November, 1982, and release for sale in January, 1983. It became that platform's killer app, and drove sales of the PC due to the improvements in speed and graphics compared to VisiCalc.

Lotus 1-2-3, along with its competitor Borland Quattro, soon displaced VisiCalc. Lotus 1-2-3 was released on January 26, 1983, started outselling then-most-popular VisiCalc the very same year, and for a number of years was the leading spreadsheet for DOS.

Microsoft Excel

Microsoft had been developing Excel on the Macintosh platform for several years at this point, where it had developed into a fairly powerful system. A port of Excel to Windows 2.0 resulted in a fully functional Windows spreadsheet. Themore robust Windows 3.x platforms of the early 1990s made it possible for Excel to take market share from Lotus. By the time Lotus responded with usable Windows products, Microsoft had started compiling their Office suite. Starting in the mid 1990s continuing through 2008, Microsoft Excel has dominated the commercial spreadsheet market.

Apple Numbers

Number is Apple Inc.'s spreadsheet software, part of iWork. It focuses on usability and the elegance of chart presentation. Numbers completed Apple's productivity suite, making it a viable competitor to Microsoft Office. It lacks features such as pivot tables.

OpenOffice.org

OpenOffice.org Calc is a freely available, open-source program modelled after Microsoft Excel. Calc can both open and save in the Excel (XLS) file format. Calc can be acquired as both an installation file and a portable program, capable of being run from a device such as a USB memory drive. It can be downloaded from the OpenOffice.org website.

Web Based Spreadsheets

With the advent of advanced web technologies such as Ajax circa 2005, a new generation of online spreadsheets has emerged. Equipped with a rich Internet application user experience, the best web based online spreadsheets have many of the features seen in desktop spreadsheet applications. Some of them have strong multi-user collaboration features. Some of them offer real time updates from remote sources such as stock prices and currency exchange rates.

3.2 Use of Spreadsheets

The uses of spreadsheets are tremendously diverse. You can use spreadsheet for on-the-moment numeric calculations that you don't need to save, for long-term projects that accumulate monthly or yearly data, and for myriad other applications from billing invoices to preparing financial statement and tax worksheet. In general, you use a spreadsheet to deal with numeric data and especially to perform any kind of analysis to aid in decision making.

Financial Analysis

If there is any one area where spreadsheets are truly pervasive, it is in business. Here are few examples of how spreadsheets aid decision making in businesses.

- The marketing department of a direct-sales company might use a spreadsheet to keep track record of sales totals for its sales staff. They can use the information to calculate commissions for the sales staff and to project materials and product requirements for purchasing and manufacturing departments.
- The purchasing department may use a spreadsheet to keep track of purchases from vendors, including running totals of purchases from individual vendors and prices paid for specific products.
- Employees in the manufacturing departments could use a spreadsheet to keep a record of maintenance performed on

equipment. They can also record receipts of goods from loading docks and transfers of finished goods to a warehouse.

- The payroll department might use a spreadsheet for some types of analysis, even if the payroll system is already automated. For example, they might compile a list of employees who receive a car allowance in addition to commissions, or calculate the balance of an employee's retirement account.
- The accounting department uses spreadsheet to calculate monthly journal entries or to schedule account balances. They may also keep track of the depreciation and book values of company assets, and even prepare their monthly, quarterly, or annual financial statements with a spreadsheet.
- Marketing executives can create impressive slide shows with graphs of past and anticipated product sales. Finance executives can also use a spreadsheet's graphical capabilities to depict the company's financial strengths compared to those of others in the same industry, and to forecast financial trends.

Engineering Analysis

Engineering is another profession that benefits from using spreadsheets for numeric computations.

- A civil engineer might use a spreadsheet for determining the dimensions of drainage channel for a given flow capacity, or to calculate the stress on a concrete beam in a bridge, or to find the weight of a road.
- Electrical engineers may use a spreadsheet to calculate voltages or capacities in the design of electrical circuits, or to create three-dimensional graphs of the output of a circuit.
- A mechanical engineer might use a spreadsheet to calculate the geometric dimensions of objects or to calculate the heat absorption or dissipation of a material being exposed to external pressures.

Graphics and Presentations

Spreadsheets are excellent tools for creating graphs and presentations materials, because these capabilities are built into them. With many spreadsheets, creating a graphic is as simple as using the mouse to select a block of cells, and clicking a graph button.

3.3 Concepts

Cells

A "cell" can be thought of as a box for holding a datum. A single cell is usually referenced by its column and row (A2 would represent the cell below containing the value 10). Its physical size can usually be tailored for its content by dragging its height or width at box intersections (or for entire columns or rows by dragging the column or rows headers).

My Spreadsheet

	A	B	C	D
01	**value1**	**Value2**	**added**	**multiplied**
02	**10**	**20**	**30**	**200**

An array of cells is called a "sheet" or "worksheet". It is analogous to an array of variables in a conventional computer program (although certain unchanging values, once entered, could be considered, by the same analogy, constants). In most implementations, many worksheets may be located within a single spreadsheet. A worksheet is simply a subset of the spreadsheet divided for the sake of clarity. Functionally, the spreadsheet operates as a whole and all cells operate as global variables within the spreadsheet ('read' access only except its own containing cell).

A cell may contain a value or a formula, or it may simply be left empty. By convention, formulas usually begin with = sign.

Values: A value can be entered from the computer keyboard by directly typing into the cell itself. Alternatively, a value can be based on a formula (see below), which might perform a calculation, display the current date or time, or retrieve external data such as a stock quote or a database value.

The Spreadsheet Value Rule: Computer scientist Alan Kay used the term value rule to summarise a spreadsheet's operation: a cell's value relies solely on the formula the user has typed into the cell. The formula may rely on the value of other cells, but those cells are likewise restricted to user-entered data or formulas. There are no 'side effects' to calculating a formula: the only output is to display the calculated result inside its occupying cell. There is no natural mechanism for permanently modifying the contents of a cell unless the user manually modifies the cell's contents. In the context of programming languages, this yields a limited form of first-order functional programming.

Automatic recalculation: A standard of spreadsheets since the mid 80s, this optional feature eliminates the need to manually request the spreadsheet program to recalculate values (nowadays typically the default option unless specifically 'switched off' for large spreadsheets, usually to improve performance). Some earlier spreadsheets required a manual request to recalculate, since recalculation of large or complex spreadsheets often reduced data entry speed. Many modern spreadsheets still retain this option.

Real Time Update: This feature refers to updating a cell's contents periodically when its value is derived from an external source - such as a cell in another "remote" spreadsheet. For shared, web-based spreadsheets, it applies to "immediately" updating cells that have been altered by another user. All dependent cells have to be updated also.

Formula: Animation of a simple spreadsheet that multiplies values in the left column by 2, then sums the calculated values from the right column to the bottom-most cell. In this example, only the values in the **A column are entered (10, 20, 30), and the remainder of cells are** formulas. Formulas in the B column multiply values from the A column using relative references, and the formula in B4 uses the SUM() function to find the sum of values in the B1:B3 range.

A formula identifies the calculation needed to place the result in the cell it is contained within. A cell containing a formula therefore has two display components; the formula itself and the resulting value. The formula is normally only shown when the cell is selected by "clicking" the mouse over a particular cell; otherwise it contains the result of the calculation.

A formula assigns values to a cell or range of cells, and typically has the format:

```
=expression
```

where the expression consists of:

- values, such as 2, 9.14 or 6.67E-11;
- references to other cells, such as, e.g., A1 for a single cell or B1:B3 for a range;
- arithmetic operators, such as +, -, *, /, and others;
- relational operators, such as >=, <, and others; and,
- functions, such as SUM(), TAN(), and many others.

When a cell contains a formula, it often contains references to other cells. Such a cell reference is a type of variable. Its value is the value of the referenced cell or some derivation of it. If that cell in turn references other cells, the value depends on the values of those. References can be relative (e.g., A1, or B1:B3), absolute (e.g., $A$1, or $B$1:$B$3) or mixed row-wise or column-wise absolute/relative (e.g., $A1 is column-wise absolute and A$1 is row-wise absolute).

The available options for valid formulas depend on the particular spreadsheet implementation but, in general, most arithmetic operations and quite complex nested conditional operations can be performed by most of today's commercial spreadsheets. Modern implementations also offer functions to access custom-build functions, remote data, and applications.

A formula may contain a condition (or nested conditions) - with or without an actual calculation - and is sometimes used purely to identify and highlight errors. In the example below, it is assumed the sum of a column of percentages (A1 through A6) is tested for validity and an explicit message put into the adjacent right hand cell.

=IF (SUM(A1:A6) > 100, "More than 100%", SUM(A1:A6))

A spreadsheet does not, in fact, have to contain any formulas at all, in which case it could be considered merely a collection of data arranged in rows and columns (a database) like a calendar, timetable or simple list. Because of its ease of use, formatting and hyper-linking capabilities, many spreadsheets are used solely for this purpose.

Locked Cell: Once entered, selected cells (or the entire spreadsheet) can optionally be "locked" to prevent accidental overwriting. Typically this would apply to cells containing formulas but might be applicable to cells containing "constants" such as a kilogram/pounds conversion factor (2.20462262 to eight decimal places).

Data Format: A cell or range can optionally be defined to specify how the value is displayed. The default display format is usually set by its initial content if not specifically previously set, so that for example "31/12/2007" or "31 Jan 2007" would default to the cell format of "date". Similarly adding a % sign after a numeric value would tag the cell as a percentage cell format. The cell contents are not changed by this format, only the displayed value.

Some cell formats such as "numeric" or "currency" can also specify the number of decimal places.

This can allow invalid operations (such as doing multiplication on a cell containing a date), resulting in illogical results without an appropriate warning.

Text Format: Each cell (like its counterpart the "word" in a word processor) can be separately defined in terms of its displayed format.
Any cell or range of cells can be highlighted in several different ways such as use of bold text, colour, font, text size and so on.

These attributes typically do not alter the data content in any way and some formatting may be lost or altered when copying spreadsheet data between different implementations or software versions. In some implementations, the format may be conditional upon the data within the cell - for example, a value may be displayed red if it is negative.

Named Cells: In most implementations, a cell can be "named" so that even if the cell is "cut and pasted" to a new location within the spreadsheet, its reference always remains intact. Names must be unique within the spreadsheet and, once defined, can then be used instead of a "normal" cell reference.

Cell Reference: A cell reference is the name of some cell in some spreadsheet. Most cell references indicate another cell in the same spreadsheet, but a cell reference can also refer to a cell in a different sheet within the same spreadsheet, or (depending on the implementation) to a cell in another spreadsheet entirely, or to a value from a remote application.

A typical cell reference in "A1" style consists of one or two case-insensitive letters to identify the column (if there are up to 256 columns: A-Z and AA-IV) followed by a row number (e.g. in the range 1-65536). Either part can be relative (it changes when the formula it is in is moved or copied), or absolute (indicated with $ in front of the part concerned of the cell reference). The older "R1C1" reference style consists of the letter R, the row number, the letter C, and the column number; relative row or column numbers are indicated by enclosing the number in square brackets. Most current spreadsheets use the A1 style, some providing the R1C1 style as a compatibility option.

When the computer calculates a formula in one cell to update the displayed value of that cell, cell reference(s) in that cell, naming some other cell(s), cause the computer to fetch the value of the named cell(s).
A cell on the same "sheet" is usually addressed as:-
=A1
A cell on a different sheet of the same spreadsheet is usually addressed

as:-

=SHEET2!A1 (that is; the first cell in sheet 2 of same spreadsheet).

Some spreadsheet implementations allow a cell references to another spreadsheet (not the current open and active file) on the same computer or a local network. It may also refer to a cell in another open and active spreadsheet on the same computer or network that is defined as shareable. These references contain the complete filename, such as:-
='C:\Documents and Settings\Username\My spreadsheets\[main sheet]Sheet1!A1

In a spreadsheet, references to cells are automatically updated when new rows or columns are inserted or deleted. Care must be taken however when adding a row immediately before a set of column totals to ensure that the totals reflect the additional rows values - which often they do not!

A circular reference occurs when the formula in one cell has a reference that directly -- or indirectly, through a chain of references, each one pointing to another cell that has another reference to the next cell on the chain -- points to the one cell. Many common kinds of errors cause such circular references. However, there are some valid techniques that use such circular references. Such techniques, after many recalculations of the spreadsheet, (usually) converge on the correct values for those cells.

Cell Ranges: A reference to a range of cells is typically of the form (A1:A6) which specifies all the cells in the range A1 through to A6. A formula such as "= SUM (A1:A6)" would add all the cells specified and put the result in the cell containing the formula itself.

Sheets

In the earliest spreadsheets, cells were a simple two-dimensional grid. Over time, the model has been expanded to include a third dimension, and in some cases a series of named grids, called sheets. The most advanced examples allow inversion and rotation operations which can slice and project the data set in various ways.

Remote Spreadsheet

Whenever a reference is made to a cell or group of cells that are not located within the current physical spreadsheet file, it is considered as accessing a "remote" spreadsheet. The contents of the referenced cell may be accessed either on first reference with a manual update or more recently in the case of web based spreadsheets, as a near real time value with a specified automatic refresh interval.

Charts

An example is histogram of the heights of 31 Black Cherry trees. Many spreadsheet applications permit charts, graphs or histograms to be generated from specified groups of cells which are dynamically re-built as cell contents change. The generated graphic component can either be embedded within the current sheet or added as a separate object.

3.4 Shortcomings

While spreadsheets are a great step forward in quantitative modelling, they have deficiencies. At the level of overall user benefits, spreadsheets have three main shortcomings.

- Spreadsheets have significant reliability problems. Research studies estimate that roughly 94% of spreadsheets deployed in the field contain errors, and 5.2% of cells in unedited spreadsheets contain errors.
- The practical expressiveness of spreadsheets is limited. Several factors contribute to this limitation. Implementing a complex model requires implementing detailed layouts, cell-at-a-time. Authors have difficulty remembering the meanings of hundreds or thousands of cell addresses that appear in formula.
- Productivity of spreadsheet modellers is reduced by the cell-level focus of spreadsheets. Even conceptually simple changes in spreadsheets (such as changing starting or ending time or time grain, adding new members or a level of hierarchy to a dimension, or changing one conceptual formula that is represented as hundreds of cell formulas) often require large numbers of manual cell-level operations (such as inserting or deleting cells/rows/columns, editing and copying formulas, re-laying out worksheets). Each of these manual corrections increases the risk of introducing further mistakes.

These three deficiencies in high-level benefits have deeper causes that, ironically, flow directly from the signature strength of spreadsheets (that they capture the structure of models in terms of WYSIWYG sheet layout for authors and report users).

- Spreadsheets capture model logic in terms of sheet layout, especially contiguous layout of cells in a table. Spreadsheets have weak or nonexistent methods to capture higher level structures such as named variables, segmentation dimensions, and time series.
- Formulas are subordinated to the cell layout. This forces the sheet layout to carry the structure of the model, not variables and formulas that relate variables. This also causes a large proliferation of cells, formulas and cell-level tasks even when only a few basic concepts

are involved in a model. This forces authors to think and work at the level of cells instead of at the level of the natural concepts and structures of the model.

- Formulas expressed in terms of cell addresses are hard to keep straight and hard to audit. Research shows that spreadsheet auditors who check numerical results and cell formulas find no more errors than auditors who only check numerical results.
- Proliferation of error-prone manual cell-level operations contributes to all three of the high-level problems listed above.

3.5 Spreadsheets as Decision-Making Tools in Financial Modelling

Financial modelling is the process of simulating the financial performance of a companyusing spreadsheet. The purpose of the financial model is to project the future financial performance of the company. Spreadsheet software makes this projection possible because it lets you see the effect of plugging different numbers into the formulas that make up the model.

A simple example of financial model may involve that you run a bicycle messenger service business. The messengers are paid 50 percent of the price of each delivery. Other than the salaries, the only operating expenses are office rent, utilities (electricity, telephone, and garbage), a computer system, and the two-way radio system that is used to communicate with the messengers.

Using a spreadsheet, you can plug into different sales volumes and determine the profit you would make. The information might be vital when planning to move offices or to buy a new two-way radio system.

Naturally, financial models for larger corporations can be far more complex. The most common type of model has areas in the spreadsheet for each of the primary financial statements-balance sheet, income statement, and cash flow statements. Each of these financial statements might be linked to other spreadsheets, so the data can be imported automatically and analysed there.

The whole model can get complicated, but the goal is still the same. It makes you see how different data affects the profitability of the company. Results of financial model help to make the basic business decisions such as:

- How much inventory can a company afford to keep on hand?
- How big a bonus should employees receive?

- Is it better to find less expensive parts or to raise prices?

In the old days, the financial models used to answer these questions were created by accountants or financial manager of a company such as the chief financial officer. But with the ever-widening use of computers, all kinds of business people are now creating and gaining from financial models. Large businesses assume that their new employees are sufficiently adept with spreadsheet to create financial models. Owners and managers of small businesses are equally likely to use such models to help them make decisions. Given the great value that a financial model can have when making business decisions, and the ease with which these models can be made with modern spreadsheet software, the ability to create financial model has become an essential and basic business skill.

4.0 CONCLUSION

A spreadsheet is a tool for calculating, and evaluating numbers. It also offers capabilities for creating reports and presentations to communicate what analysis reveals. Spreadsheet software makes these tasks easy by providing a visual framework to work within, and providing the tools that is needed to make number calculating a breeze. Spreadsheet is an excellent decision support system tool that cuts across disciplines.

5.0 SUMMARY

- A spreadsheet is a computer application that simulates a paper worksheet. It displays multiple cells that together make up a grid consisting of rows and columns, each cell containing either alphanumeric text or numeric values.
- The word "spreadsheet" came from "spread" in its sense of a newspaper or magazine item (text and/or graphics) that covers two facing pages, extending across the centre fold and treating the two pages as one large one.
- The uses of spreadsheets are tremendously diverse. You can use spreadsheet for on-the-moment numeric calculations that you don't need to save, for long-term projects that accumulate monthly or yearly data, and for myriad of other applications from billing invoices to preparing financial statement and tax worksheet.
- A "cell" can be thought of as a box for holding a datum. A single cell is usually referenced by its column and row (A2 would represent the cell below containing the value 10).
- While spreadsheets are a great step forward in quantitative modelling, they have deficiencies. At the level of overall user benefits, spreadsheets have three main shortcomings.
- Financial modelling is the process of simulating the financial performance of a company using spreadsheet. The purpose of the

financial model is to project the future financial performance of the company.

6.0 TUTOR-MARKED ASSIGNMENT

Mention and discuss briefly the uses of spreadsheet in financial analysis.

7.0 REFERENCES/FURTHER READING

Norton, Peter (1995). Introduction to Computers. MacMillan/McGraw-Hill.

UNIT 4 EXECUTIVE INFORMATION SYSTEM I

CONTENTS

1.0 INTRODUCTION

An Executive Information System (EIS) is a type of management information system intended to facilitate and support the information and decision making needs of senior executives by providing easy access to both internal and external information relevant to meeting the strategic goals of the organisation. It is commonly considered as a specialised form of a Decision Support System (DSS).

The emphasis of EIS is on graphical displays and easy-to-use user interfaces. They offer strong reporting and drill-down capabilities. In general, EIS are enterprise-wide DSS that help top-level executives analyse, compare, and highlight trends in important variables so that they can monitor performance and identify opportunities and problems. EIS and data warehousing technologies are converging in the marketplace.

In recent years, the term EIS has lost popularity in favour of Business Intelligence (with the sub areas of reporting, analytics, and digital dashboards).

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define and understand the concept of executive information system (EIS)

- evaluate the development and history of EIS
- identify the major components and contents of EIS
- outline the various applications of EIS in sectors of business
- answer the questions of advantages and disadvantages associated with EIS
- explain what the future holds for EIS.

3.0 MAIN CONTENT

3.1 History

Traditionally, executive information systems were developed as mainframe computer-based programs. The purpose was to package a company's data and to provide sales performance or market research statistics for decision makers, such as financial officers, marketing directors, and chief executive officers, who were not necessarily well acquainted with computers. The objective was to develop computer applications that would highlight information to satisfy senior executives' needs. Typically, an EIS provides data that would only need to support executive level decisions instead of the data for all the company.

Today, the application of EIS is not only in typical corporate hierarchies, but also at personal computers on a local area network. EIS now cross computer hardware platforms and integrate information stored on mainframes, personal computer systems, and minicomputers. As some client service companies adopt the latest enterprise information systems, employees can use their personal computers to get access to the company's data and decide which data are relevant for their decision makings. This arrangement makes all users able to customise their access to the proper company's data and provide relevant information to both upper and lower levels in companies.

3.1 Components

The components of an EIS can typically be classified as:

Hardware

When talking about hardware for an EIS environment, we should focus on the hardware that meet the executive's needs. The executive must be put first and the executive's needs must be defined before the hardware can be selected. The basic computer hardware needed for a typical EIS includes four components:

(1) Input data-entry devices. These devices allow the executive to enter, verify, and update data immediately;

(2) The central processing unit (CPU), which is the kernel because it controls the other computer system components;

(3) Data storage files. The executive can use this part to save useful business information, and this part also helps the executive to search historical business information easily;

(4) Output devices, which provide a visual or permanent record for the executive to save or read. This device refers to the visual output device or printer.

In addition, with the advent of local area networks (LAN), several EIS products for networked workstations became available. These systems require less support and less expensive computer hardware. They also increase access of the EIS information to many more users within a company.

Software

Choosing the appropriate software is vital to design an effective EIS. Therefore, the software components and how they integrate the data into one system are very important. The basic software needed for a typical EIS includes four components:

1. **Text base software. The most common form of text is probably** documents;
2. **Database. Heterogeneous databases residing on a range of vendor-** specific and open computer platforms help executives access both internal and external data;
3. **Graphic base. Graphics can turn volumes of text and statistics into** visual information for executives. Typical graphic types are: time series charts, scatter diagrams, maps, motion graphics, sequence charts, and comparison-oriented graphs (i.e., bar charts);
4. **Model base. The EIS models contain routine and special statistical,** financial, and other quantitative analysis.

Perhaps a more difficult problem for executives is choosing from a range of highly technical software packages. Ease of use, responsiveness to executives' requests, and price are all reasonable considerations. Further, it should be considered whether the package can run on existing hardware.

User Interface

An EIS needs to be efficient to retrieve relevant data for decision makers, so the userinterface is very important. Several types of interfaces can be available to the EIS structure, such as scheduled reports, questions/answers, menu driven, command language, natural language, and input/output. It is crucial that the interface must fit the decision maker's decision-making style. If the executive is not comfortable with the information questions/answers style, the EIS will not be fully utilised. The ideal interface for an EIS would be simple to use and highly flexible, providing consistent performance, reflecting the executive's world, and containing help information.

Telecommunication

As decentralising is becoming the current trend in companies, telecommunications will play a pivotal role in networked information systems. Transmitting data from one place to another has become crucial for establishing a reliable network. In addition, telecommunications within an EIS can accelerate the need for access to distributed data.

3.3 Applications

EIS enables executives to find those data according to user-defined criteria and promote information-based insight and understanding. Unlike a traditional management information system presentation, EIS can distinguish between vital and seldom-used data, and track different key critical activities for executives, which are helpful in evaluating if the company is meeting its corporate objectives. After realising its advantages, people have applied EIS in many areas, especially in manufacturing, marketing, and finance areas.

Manufacturing

Basically, manufacturing is the transformation of raw materials into finished goods for sale, or intermediate processes involving the production or finishing of semi-manufactures. It is a large branch of industry and of secondary production. Manufacturing operational control focuses on day-to-day operations, and the central idea of this process is effectiveness and efficiency. To produce meaningful managerial and operational information for controlling manufacturing operations, the executive has to make changes in the decision processes. EIS provides the evaluation of vendors and buyers, the evaluation of purchased materials and parts, and analysis of critical purchasing areas. Therefore, the executive can oversee and review purchasing operations effectively with EIS. In addition, because production planning and control depends heavily on the plant's data base and its communications with all manufacturing work centres, EIS also provides an approach to improve production planning and control.

Marketing

In an organisation, marketing executives' role is to create the future. Their main duty is managing available marketing resources to create a more effective future. For this, they need make judgments about risk and uncertainty of a project and its impact on company in the short and long term. To assist marketing executives in making effective marketing decisions, an EIS can be applied. EIS provides an approach to sales forecasting, which can allow the market executive to compare sales forecast with past sales. EIS also offers an approach to product price, which is found in venture analysis. The market executive can evaluate pricing as related to competition along with the relationship of product quality with price charged. In summary, EIS software package enables marketing executives to manipulate the data by looking for trends, performing audits of the sales data, and calculating totals, averages, changes, variances, or ratios. All of these sales analysis functions help marketing executives to make final decisions.

Financial

A financial analysis is one of the most important steps to companies today. The executive needs to use financial ratios and cash flow analysis to estimate the trends and make capital investment decisions. An EIS is a responsibility-oriented approach that integrates planning or budgeting with control of performance reporting, and it can be extremely helpful to finance executives. Basically, EIS focuses on accountability of financial performance and it recognises the importance of cost standards and flexible budgeting in developing the quality of information provided for all executive levels. EIS enables executives to focus more on the long-term basis of current year and beyond, which means that the executive not only can manage a sufficient flow to maintain current operations but also can figure out how to expand operations that are contemplated over the coming years. Also, the combination of EIS and EDI environment can help cash managers to review the company's financial structureso that the best method of financing for an accepted capital project can be concluded. In addition, the EIS is a good tool to help the executive to review financial ratios, highlight financial trends and analyse a company's performance and its competitors.

3.4 Advantages and Disadvantages

Advantages

- Easy for upper-level executives to use, extensive computer experience is not required in operations
- Provides timely delivery of company summary information
- Information that is provided is better understood
- Filters data for management
- Improves tracking of information

Disadvantages

- Functions are limited, cannot perform complex calculations
- Hard to quantify benefits and to justify implementation of an EIS
- Executives may encounter information overload
- System may become slow, large, and hard to manage
- Difficult to keep current data
- May lead to less reliable and insecure data
- Small companies may encounter excessive costs for implementation

3.4 Contents of EIS

A general answer to the question of what data is appropriate for inclusion in an Executive Information System is "whatever is interesting to executives." While this advice is rather simplistic, it does reflect the variety of systems currently in use. Executive Information Systems in government have been constructed to track data about Ministerial correspondence, case management, worker productivity, finances, and human resources to name only a few. Other sectors use EIS implementations to monitor information about competitors in the news media and databases of public information in addition to the traditional revenue, cost, volume, sales, market share and quality applications. Frequently, EIS implementations begin with just a few measures that are clearly of interest to senior managers, and then expand in response to questions asked by those managers as they use the system. Over time, the presentation of this information becomes stale, and the information diverges from what is strategically important for the organisation. A "Critical Success Factors" approach is recommended by many management theorists (Daniel, 1961, Crockett, 1992, Watson and Frolick, 1992). Practitioners such as Vandenbosch (1993) found that:

> *While our efforts usually met with initial success, we often found that after six months to a year, executives were almost as bored with the new information as they had been with the old. A strategy we developed to rectify this problem required organisations to create a report of the month. That is, in addition to the regular information provided for management committee meetings, the CEO was charged with selecting a different indicator to focus*

on each month (Vandenbosch, 1993, pp. 8-9).

While the above indicates that selection of data for inclusion in an EIS is difficult, there are several guidelines that help to make that assessment. A practical set of principles to guide the design of measures and indicators to be included in an EIS is presented below (Kelly, 1992b). For a more detailed discussion of methods for selecting measures that reflect organisational objectives, see the section "EIS and Organisational Objectives."

EIS measures must be easy to understand and collect. Wherever possible, data should be collected naturally as part of the process of work. An EIS should not add substantially to the workload of managers or staff.

EIS measures must be based on a balanced view of the organisation's objective. Data in the system should reflect the objectives of the organisation in the areas of productivity, resource management, quality and customer service.

Performance indicators in an EIS must reflect everyone's contribution in a fair and consistent manner. Indicators should be as independent as possible from variables outside the control of managers.

EIS measures must encourage management and staff to share ownership of the organisation's objectives. Performance indicators must promote both team-work and friendly competition. Measures will be meaningful for all staff; people must feel that they, as individuals, can contribute to improving the performance of the organisation.

EIS information must be available to everyone in the organisation. The objective is to provide everyone with useful information about the organisation's performance. Information that must remain confidential should not be part of the EIS or the management system of the organisation.

EIS measures must evolve to meet the changing needs of the organisation.

3.6 Future Trends

The future of executive info systems will not be bound by mainframe computer systems. This trend allows executives escaping from learning different computer operating systems and substantially decreases the implementation costs for companies. Because utilising existing software applications lies in this trend, executives will also eliminate the need to learn a new or special language for the EIS package. Future executive information systems will not only provide a system that supports senior executives, but also contain the information needs for middle managers. The future executive information systems will become diverse because of integrating potential new applications and technology into the systems, such as incorporating artificial intelligence (AI) and integrating multimedia characteristics and ISDN technology into an EIS.

4.0 CONCLUSION

An executive information system is an information system which gives the executive easy access to key internal and external data. EIS has been made possible by the increasing cheapness and sophistication of microcomputer and network technology. An EIS is not only a tool for analysis, but also a tool for interrogating data. The concept of EIS is the same as in the emerging concept of business intelligence.

5.0 SUMMARY

- The future of executive info systems will not be bound by mainframe computer systems. This trend allows executives escaping from learning different computer operating systems and substantially decreases the implementation costs for companies.
- Traditionally, executive information systems were developed as mainframe computer-based programs. The purpose was to package a company's data and to provide sales performance or market research statistics for decision makers, such as financial officers, marketing directors, and chief executive officers, who were not necessarily well acquainted with computers.
- As decentralising is becoming the current trend in companies, telecommunications will play apivotal role in networked information systems.
- Choosing the appropriate software is vital to design an effective EIS.
- When talking about hardware for an EIS environment, we should focus on the hardware that meets the executive's needs.
- In an organisation, marketing executives' role is to create the future. Their main duty is managing available marketing resources to create a more effective future. For this, they need makejudgments about risk and uncertainty of a project and its impact on company in the short and long term.
- EIS enables executives to find those data according to user-defined criteria and promote information-based insight and understanding.
- A general answer to the question of what datais appropriate for inclusion in an Executive Information System is "whatever is

interesting to executives." While this advice is rather simplistic, it does reflect the variety of systems currently in use.

6.0 TUTOR-MARKED ASSIGNMENT

1. Briefly discuss software as a component of EIS
2. Mention 5 disadvantages of EIS

7.0 REFERENCES/FURTHER READING

Thierauf, Robert J. (1994). Executive Information System: A Guide for *Senior Management and MIS Professionals. Quorum Books.*

(January 1994). Executive Information Systems.

Salmeron, Jose L. and Herrero, Ines (July 2005). An AHP-Based *Methodology to Rank Critical Success Factors of Executive Information Systems, Computer Standards & Interfaces, Volume* 28, Issue 1, pp. 1-12.

Salmeron, Jose L. EIS (2003). Success: Keys and Difficulties in Major *Companies. Technovation Volume 23, Issue 1, pp. 35-38.*

Salmeron, Jose L. EIS (2002). Evolution in Large Spanish Businesses. *Information & Management, Volume 40, Issue 1, pp. 41-50.*

Salmeron, Jose L. (2002). EIS Profitability, Costs and Benefits: An *Evolutionary Approach. Industrial Management & Data Systems,Volume 102, Issue 5-6, pp. 284-288.*

Salmeron, Jose L. (2002). EIS Data: Findings from an Evolutionary *Study. Journal of Systems and Software, Volume 64, Issue 2, pp.* 111-114.

UNIT 5 IMPLEMENTING EXECUTIVE INFORMATION SYSTEM

CONTENTS

1.0 INTRODUCTION

An EIS is a tool that provides direct on-line access to relevant information about aspects of a business that are of particular interest to the senior manager.
Many senior managers find that direct on-line access to organisational data is helpful. For example, Paul Frech, President of Lockheed-Georgia, monitored employee contributions to company-sponsored programs (United Way, blood drives) as a surrogate measure of employee morale (Houdeshel and Watson, 1987). C. Robert Kidder, CEO of Duracell, found that productivity problems were due to salespeople in Germany wasting time calling on small stores and took corrective action (Main, 1989).

Information systems have long been used to gather and store information, to produce specific reports for workers, and to produce aggregate reports for managers. However, senior managers rarely use these systems directly, and often find the aggregate information to be of little use without the ability to explore underlying details (Watson & Rainer, 1991, Crockett, 1992).

An Executive Information System (EIS) is a tool that provides direct on-line access to relevant information in a useful and navigable format. Relevant information is timely, accurate, and actionableinformation about aspects of a business that are of particular interest to the senior manager. The useful and navigable format of the system means that it is specifically designed to be used by individuals with limited time, limited keyboarding skills, and little direct experience with computers. An EIS is easy to navigate so that managers can identify broad strategic issues, and then explore the information to find the root causes of those issues.

Executive Information Systems differ from traditional information systems in the following ways:

- •are specifically tailored to executive's information needs
- •are able to access data about specific issues and problems as well as aggregate reports
- •provide extensive on-line analysis tools including trend analysis, exception reporting & "drill-down" capability
- •access a broad range of internal and external data
- •are particularly easy to use (typically mouse or touch-screen driven)
- •are used directly by executives without assistance
- •present information in a graphical form

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- • differentiate EIS from traditional information systems
- • identify the purposes of an EIS
- • identify the numerous challenges to the effectiveness of EIS in an organisation
- • identify the major indices for an effective EIS implementation in organisation.
- • describe how EIS fits into organisational objectives.

3.0 MAIN CONTENT

3.1 Purpose of EIS

The primary purpose of an Executive Information System is to support managerial learning about an organisation, its work processes, and its interaction with the external environment. Informed managers can ask better questions and make better decisions. Vandenbosch and Huff (1992) from the University of Western Ontario found that Canadian firms using an EIS achieved better business results if their EIS promoted managerial learning. Firms with an EIS designed to maintain managers' "mental models" were less effective than firms with an EIS designed to build or enhance managers' knowledge.

This distinction is supported by Peter Senge in The Fifth Dimension. He illustrates the benefits of learning about the behaviour of systems versus simply learning more about their states. Learning more about the state of a system leads to reactive management fixes. Typically these reactions feed into the underlying system behaviour and contribute to a downward spiral. Learning more about system behaviour and how various system inputs and actions interrelate will allow managers to make more proactive changes to create long-term improvement.

A secondary purpose for an EIS is to allow timely access to information. All of the information contained in an EIS can typically be obtained by a manager through traditional methods. However, the resources and time required to manually compile information in a wide variety of formats, and in response to ever changing and ever more specific questions usually inhibit managers from obtaining this information. Often, by the time a useful report can be compiled, the strategic issues facing the manager have changed, and the report is never fully utilised.

Timely access also influences learning. When a manager obtains the answer to a question, that answer typically sparks other related questions in the manager's mind. If those questions can be posed immediately, and the next answer retrieved, the learning cycle continues unbroken. Using traditional methods, by the time the answer is produced, the context of the question may be lost, and the learning cycle will not continue. An executive in Rockart & Treacy's 1982 study noted that:

> *Your staff really can't help you think. The problem with giving a question to the staff is that they provide you with the answer. You learn the nature of the real question you should have asked when you muck around in the data (p. 9).*

A third purpose of an EIS is commonly misperceived. An EIS has a powerful ability to direct management attention to specific areas of the organisation or specific business problems. Some managers see this as an opportunity to discipline subordinates. Some subordinates fear the directive nature of the system and spend a great deal of time trying to outwit or discredit it. Neither of these behaviours is appropriate or productive. Rather, managers and subordinates can work together to determine the root causes of issues highlighted by the EIS.

The powerful focus of an EIS is due to the maxim "what gets measured gets done." Managers are particularly attentive to concrete information about their performance when it is available to their superiors This focus is very valuable to an organisation if the information reported is actually important and represents a balanced view of the organisation's objectives.

Misaligned reporting systems can result in inordinate management attention to things that are not important or to things which are important but to the exclusion of other equally important things. For example, a production reporting system might lead managers to emphasise volume of work done rather than quality of work. Worse yet, productivity might have little to do with the organisation's overriding customer service objectives.

3.2 Barriers to Effectiveness

There are many ways in which an EIS can fail. Dozens of high profile, high cost EIS projects have been cancelled, implemented and rarely used, or implemented and used with negative results. An EIS is a high risk project precisely because it is intended for use by the most powerful people in an organisation. Senior managers can easily misuse the information in the system with strongly detrimental effects on the organisation. Senior managers can refuse to use a system if it does not respond to their immediate personal needs or is too difficult to learn and use.

Unproductive Organisational Behaviour Norms

Issues of organisational behaviour and culture are perhaps the most deadly barriers to effective Executive Information Systems. Because an EIS is typically positioned at the top of an organisation, it cancreate powerful learning experiences and lead to drastic changes in organisational direction. However, there is also great potential for misuse of the information. Green, Higgins and Irving (1988) found that performance monitoring can promote bureaucratic and unproductive behaviour, can unduly focus organisational attention to the point where other important aspects are ignored, and can have a strongly negative impact on morale.

The key barrier to EIS effectiveness, therefore, is the way in which the organisation uses the information in the system. Managers must be aware of the dangers of statistical data, and be skilled at interpreting and using data in an effective way. Even more important is the manager's ability to communicate with others about statistical data in a non-defensive, trustworthy, and constructive manner. Argyris (1991) suggests a universal human tendency towards strategies that avoid embarrassment or threat, and towards feelings of vulnerability or incompetence. These strategies include:

- stating criticism of others in a way that you feel is valid but also in a way that prevents others from deciding for themselves
- failing to include any data that others could use to objectively evaluate your criticism

•stating your conclusions in ways that disguise their logical implications and denying those implications if they are suggested.

To make effective use of an EIS, mangers must have the self-confidence to accept negative results and focus on the resolution of problems rather than on denial and blame. Since organisations with limited exposure to planning and targeting, data-based decision-making, statistical process control, and team-based work models may not have dealt with these behavioural issues in the past, they are more likely to react defensively and reject an EIS.

Technical Excellence

An interesting result from the Vandenbosch & Huff (1988) study was that the technical excellence of an EIS has an inverse relationship with effectiveness. Systems that are technical masterpieces tend to be inflexible, and thus discourage innovation, experimentation and mental model development.

Flexibility is important because an EIS has such a powerful ability to direct attention to specific issues in an organisation. A technical masterpiece may accurately direct management attention when the system is first implemented, but continue to direct attention to issues that were important a year ago on its first anniversary. There is substantial danger that the exploration of issues necessary for managerial learning will be limited to those subjects that were important when the EIS was first developed. Managers must understand that as the organisation and its work changes, an EIS must continually be updated to address the strategic issues of the day.

A number of explanations as to why technical masterpieces tend to be less flexible are possible. Developers who create a masterpiece EIS may become attached to the system and consciously or unconsciously dissuade managers from asking for changes. Managers who are uncertain that the benefits outweigh the initial cost of a masterpiece EIS may not want to spend more on system maintenance and improvements. The time required to create a masterpiece EIS may mean that it is outdated before it is implemented.

While usability and response time are important factors in determining whether executives will use a system, cost and flexibility are paramount. A senior manager will be more accepting of an inexpensive system that provides 20% of the needed information within a month or two than with an expensive system that provides 80% of the needed information after a year of development. The manager may also find that the inexpensive system is easier to change and adapt to the evolving needs of the business. Changing a large system would involve throwing away parts of a substantial investment. Changing the inexpensive system means losing a few weeks of work. As a result, fast, cheap, incremental approaches to developing an EIS increase the chance of success.

Technical Problems

Paradoxically, technical problems are also frequently reported as a significant barrier to EIS success. The most difficult technical problem -- that of integrating data from a wide range of data sources both inside and outside the organisation -- is also one of the most critical issues for EIS users. A marketing vice-president, who had spent several hundred thousand dollars on an EIS, attended a final briefing on the system. The technical experts demonstrated the many graphs and charts of sales results, market share and profitability. However, when the vice-president asked for a graph of market share and advertising expense over the past ten years, the system was unable to access historical data. The project was cancelled in that meeting.

The ability to integrate data from many different systems is important because it allows managerial learning that is unavailable in other ways. The president of a manufacturing company can easily get information about sales and manufacturing from the relevant VPs. Unfortunately, the information the president receives will likely be incompatible, and learning about the ways in which sales and manufacturing processes influence each other will not be easy. An EIS will be particularly effective if it can overcome this challenge, allowing executives to learn about business processes that cross organisational boundaries and to compare business results in disparate functions.

Another technical problem that can kill EIS projects is usability. Senior managers simply have the choice to stop using a system if they find it too difficult to learn or use. They have very little time to invest in learning the system, a low tolerance for errors, and initially may have very little incentive to use it. Even if the information in the system is useful, a difficult interface will quickly result in the manager assigning an analyst to manipulate the system and print out the required reports. This is counter-productive because managerial learning is enhanced by the immediacy of the question - answer learning cycle provided by an EIS. If an analyst is interacting with the system, the analyst will acquire more learning than the manager, but will not be in a position to put that learning to its most effective use.

Usability of Executive Information Systems can be enhanced through the use of prototyping and usability evaluation methods. These methods ensure that clear communication occurs between the developers of the system and its users. Managers have an opportunity to interact with systems that closely resemble the functionality of the final system and thus can offer more constructive criticism than they might be able to after reading an abstract specification document. Systems developers also are in a position to listen more openly to criticisms of a system in a prototype is expected to be disposable. Several evaluation products are available including observation and monitoring, software logging, experiments and benchmarking, etc. (Preece et al, 1994). The most appropriate methods for EIS design are those with an ethnographic flavour because the experience base of system developers is typically so different from that of their user population (senior executives).

Misalignment between Objectives and EIS

A final barrier to EIS effectiveness was mentioned earlier in the section on purpose. As noted there, the powerful ability of an EIS to direct organisational attention can be destructive if the system directs attention to the wrong variables. There are many examples of this sort of destructive reporting. Grant, Higgins and Irving (1988) report the account of an employee working under a misaligned reporting system thus:

> *I like the challenge of solving customer problems, but they get in the way of hitting my quota. I'd like to get rid of the telephone work. If (the company) thought dealing with customers was important, I'd keep it; but if it's just going to be production that matters, I'd gladly give all the calls to somebody else.*

Traditional cost accounting systems are also often misaligned with organisational objectives, and placing these measures in an EIS will continue to draw attention to the wrong things. Cost accounting allocates overhead costs to direct labour hours. In some cases the overhead burden on each direct labour hour is as much as 1000%. A manager operating under this system might decide to sub-contract 100 hours of direct labor at $20 per hour. On the books, this $2,000 saving is accompanied by $20,000 of savings in overhead. If the sub-contractor charges $5,000 for the work, the book savings are $2,000 + $20,000 - $5,000 = $17,000. In reality, however, the overhead costs for an idle machine in a factory do not go down much at all. The sub-contract actually ends up costing $5,000 - $2,000 = $3,000. (Peters, 1987)

3.3 Characteristics of Successful EIS Implementations

Find an Appropriate Executive Champion

EIS projects that succeed do so because at least one member of the senior management team agrees to champion the project. The executive champion need not fully understand the technical issues, but must be a person who works closely with all of the senior management team and understands their needs, work styles and their current methods of obtaining organisational information. The champion's commitment must include a willingness to set aside time for reviewing prototypes and implementation plans, influencing and coaching other members of the senior management team, and suggesting modifications and enhancements to the system.

Deliver a Simple Prototype Quickly

Executives judge a new EIS on the basis of how easy it is to use and how relevant the information in the system is to the current strategic issues in the organisation. As a result, the best EIS projects begin as a simple prototype, delivered quickly, that provides data about at least one critical issue. If the information delivered is worth the hassle of learning the system, a flurry of requirements will shortly be generated by executives who like what they see, but want more. These requests are the best way to plan an EIS that truly supports the organisation, and are more valuable than months of planning by a consultant or analyst.

One caveat concerning the simple prototype approach is that executive requests will quickly scatter to questions of curiosity rather than strategy in an organisation where strategic direction and objectives are not clearly defined. A number of methods are available to support executives in defining business objectives and linking them to performance monitors in an EIS. These are discussed further in the section on EIS and Organisational Objectives below.

Involve Your Information Systems Department

In some organisations, the motivation for an EIS project arises in the business units quite apart from the traditional information systems (IS) organisation. Consultants may be called in, or managers and analysts in the business units may take the project on without consulting or involving IS. This is a serious mistake. Executive Information Systems rely entirely on the information contained in the systems created and maintained by this department. IS professionals know best what information is available in an organisation's systems and how to get it. They must be involved in the team. Involvement in such a project can also be beneficial to IS by giving them a more strategic perspective on how their work influences the organisation.

Communicate & Train to Overcome Resistance

A final characteristic of successful EIS implementations is that of communication. Executive Information Systems have the potential to drastically alter the prevailing patterns of organisational communication and thus will typically be met with resistance. Some of this resistance is simply a matter of a lack of knowledge. Training on how to use statistics and performance measures can help. However, resistance can also be rooted in the feelings of fear, insecurity and cynicism experienced by individuals throughout the organisation. These attitudes can only be influenced by a strong and vocal executive champion who consistently reinforces the purpose of the system and directs the attention of the executive group away from unproductive and punitive behaviours.

EIS and Organisational Culture

Henry Mintzberg (1972) has argued that impersonal statistical data is irrelevant to managers. John Dearden (1966) argued that the promise of real-time management information systems was a myth and would never be of use to top managers. Grant, Higgins, and Irving (1988) argue that computerised performance monitors undermine trust, reduce autonomy and fail to illuminate the most important issues.

Many of these arguments against EISs have objective merit. Managers really do value the tangible tidbits of detail they encounter in their daily interactions more highly than abstract numerical reports. Rumours suggest a future, while numbers describe a past. Conversations are rich in detail and continuously probe the reasons for the situation while statistics are vague approximations of reality. When these vague approximations are used to intimidate or control behaviour rather than to guide learning, they really do have a negative impact on the organisation.

Yet both of these objections point to a deeper set of problems -- the assumptions, beliefs, values and behaviours that people in the organisation hold and use to respond to their environment. Perhaps managers find statistical data to be irrelevant because they have found too many errors in previous reports? Perhaps people in the organisation prefer to assign blame rather than discover the true root cause of problems. The culture of an organisation can have a dramatic

influence on the adoption and use of an Executive Information System.
The following cultural characteristics will contribute directly to the success or failure of an EIS project.

Learning vs Blaming: A learning organisation is one that seeks first to understand why a problem occurred, and not who is to blame. It is a common and natural response for managers to try to deflect responsibility for a problem on to someone else. An EIS can help to do this by indicating very specifically who failed to meet a statistical target, and by how much. A senior manager, armed with EIS data, can intimidate and blame the appropriate person. The blamed person can respond by questioning the integrity of the system, blaming someone else, or even reacting in frustration by slowing work down further.

In a learning organisation, any unusual result is seen as an opportunity to learn more about the business and its processes. Managers who find an unusual statistic explore it further, breaking it down to understand its components and comparing it with other numbers to establish cause and effect relationships. Together as a team, management uses numerical results to focus learning and improve business processes across the organisation. An EIS facilitates this approach by allowing instant exploration of a number, its components and its relationship to other numbers.

Continuous Improvement vs Crisis Management: Some organisations find themselves constantly reacting to crises, with little time for any proactive measures. Others have managed to respond to each individual crisis with an approach that prevents other similar problems in the future. They are engaged in a continual cycle of improving business practices and finding ways to avoid crisis.
Crises in government are frequently caused by questions about organisational performance raised by an auditor, the Minister, or members of the Opposition. An EIS can be helpful in responding to this sort of crisis by providing instant data about the actual facts of the situation. However, this use of the EIS does little to prevent future crises.

An organisational culture in which continual improvement is the norm can use the EIS as an early warning system pointing to issues that have not yet reached the crisis point, but are perhaps the most important areas on which to focus management attention and learning. Organisations with a culture of continuous improvement already have an appetite for the sort of data an EIS can provide, and thus will exhibit less resistance.

Team Work vs Hierarchy: An EIS has the potential to substantially disrupt an organisation that relies upon adherence to a strict chain of command. The EIS provides senior managers with the ability to micro-manage details at the lowest levels in the organisation. A senior manger with an EIS report who is surprised at the individual results of a front-line worker might call that person directly to understand why the result is unusual. This could be very threatening for the managers between the senior manager and the front-line worker. An EIS can also provide lower-level managers with access to information about peer performance and even the performance of their superiors.

Organisations that are familiar with work teams, matrix managed projects and other forms of interaction outside the chain of command will find an EIS less disruptive. Senior managers in these organisations have learned when micro-management is appropriate and when it is not. Middle managers have learned that most interactions between their superiors and their staff are not threatening to their position. Workers are more comfortable interacting with senior managers when the need arises, and know what their supervisor expects from them in such an interaction.

Data-based Decisions vs Decisions in a Vacuum: The total quality movement, popular in many organisations today, emphasises a set of tools referred to as Statistical Process Control (SPC). These analytical tools provide managers and workers with methods of understanding a problem and finding solutions rather than allocating blame and passing the buck. Organisations with training and exposure to SPC and analytical tools will be more open to an EIS than those who are suspicious of numerical measures and the motives of those who use them.

It should be noted that data-based decision making does not deny the role of intuition, experience, or negotiation amongst a group. Rather, it encourages decision-makers to probe the facts of a situation further before coming to a decision. Even if the final decision contradicts the data, chances are that an exploration of the data will help the decision-maker to understand the situation better before a decision is reached. An EIS can help with this decision-making process.

Information Sharing vs Information Hoarding: Information is power in many organisations, and managers are motivated to hoard information rather than to share it widely. For example, managers may hide information about their own organisational performance, but jump at any chance to see information about performance of their peers.

A properly designed EIS promotes information sharing throughout the organisation. Peers have access to information about each other's domain; junior managers have information about how their performance contributes to overall organisational performance. An organisation that is comfortable with information sharing will have developed a set of "good manners" for dealing with this broad access to information. These behavioural norms are key to the success of an EIS.

Specific Objectives vs Vague Directions: An organisation that has experience developing and working toward Specific, Measurable, Achievable and Consistent (SMAC) objectives will also find an EIS to be less threatening. Many organisations are uncomfortable with specific performance measures and targets because they believe their work to be too specialised or unpredictable. Managers in these organisations tend to adopt vague generalisations and statements of the exceedingly obvious in place of SMAC objectives that actually focus and direct organisational performance. In a few cases, it may actually be true that numerical measures are completely inappropriate for a few aspects of the business. In most cases, managers with this attitude have a poor understanding of the purpose of objective and target-setting exercises. Some business processes are more difficult to measure and set targets for than others. Yet almost all business processes have at least a few characteristics that can be measured and improved through conscientious objective setting. (See the following section on EIS and Organisational Objectives.)

3.4 EIS and Organisational Objectives

A number of writers have discovered that one of the major difficulties with EIS implementations is that the information contained in the EIS either does not meet executive requirements, or meets executive requirements, but fails to guide the organisation towards its objectives. As discussed earlier, organisations that are comfortable in establishing and working towards Specific, Measurable, Achievable, and Consistent (SMAC) objectives will find it easier to create an EIS that actually drives organisational performance. Yet even these organisations may have difficulty because their stated objectives do not represent all of the things that are important.

Crockett (1992) suggests a four step process for developing EIS information requirements based on a broader understanding of organisational objectives. The steps are:

(1) Identify critical success factors and stakeholder expectations,
(2) Document performance measures that monitor the critical success factors and stakeholder expectations,
(3) Determine reporting formats and frequency, and
(4) Outline information flows and how information can be used.

Crockett begins with stakeholders to ensure that all relevant objectives and critical success factors are reflected in the EIS.

Kaplan and Norton (1992) suggest that goals and measures need to be developed from each of four perspectives: financial, customer, internal business, and innovation and learning. These perspectives help managers to achieve a balance in setting objectives, and presenting them in a unified report exposes the tough tradeoffs in any management system. An EIS built on this basis will not promote productivity while ignoring quality, or customer satisfaction while ignoring cost.

Meyer (1994) raises several questions that should be asked about measurement systems for teams. Four are appropriate for evaluating objectives and measures represented in an EIS. They are:

- •Are all critical organisational outcomes tracked?
- •Are all "out-of-bounds" conditions tracked? (Conditions that are serious enough to trigger a management review.)
- •Are all the critical variables required to reach each outcome tracked?
- •Is there any measure that would not cause the organisation to change its behaviour?

In summary, proper definition of organisational objectives and measures is a helpful precondition for reducing organisational resistance to an EIS and is the root of effective EIS use. The benefits of an EIS will be fully realised only when it helps to focus management attention on issues of true importance to the organisation.

4.0 CONCLUSION

Though executive information system is an excellent tool to aid decision making, the success depends so much on how to effectively implement it. Culture, technical issues and organisational goals go a long way to determine the effectiveness of adopting EIS in an organisation.

5.0 SUMMARY

- An EIS is a tool that provides direct on-line access to relevant information about aspects of a business that are of particular interest to the senior manager.
- The primary purpose of an Executive Information System is to support managerial learning about an organisation, its work processes, and its interaction with the external environment.
- There are many ways in which an EIS can fail. Dozens of high profile, high cost EIS projects have been cancelled, implemented and rarely used, or implemented and used with negative results
- Issues of organisational behaviour and culture are perhaps the most deadly barriers to effective Executive Information Systems.
- Flexibility is important because an EIS has such a powerful ability to direct attention to specific issues in an organisation.
- EIS projects that succeed do so because at least one member of the

senior management team agrees to champion the project.

- A number of writers have discovered that one of the major difficulties with EIS implementations is that the information contained in the EIS either does not meet executive requirements, or meets executive requirements, but fails to guide the organisation towards its objectives.

6.0 TUTOR-MARKED ASSIGNMENT

1. In what ways does an EIS differ from traditional information systems?
2. Discuss briefly the involvement of information systems department as a characteristic of a successful EIS implementation

7.0 REFERENCES/FURTHER READING

Agor, Weston H. (1986). How Top Executives Use Their Intuition To *Make Important Decisions. Business Horizons. January-February* 1986, pp. 49-53.

Anderson, Joseph V. (1993). Mind Mapping: A Tool for Creative *Thinking. Business Horizons. January-February 1993, pp. 41-46.*

Argyris, Chris (1991). Teaching Smart People How to Learn. Harvard *Business Review. May-June 1991, pp. 99-109.*

Barrow, Craig (1990). Implementing an Executive Information System: *Seven Steps for Success. Journal of Information Systems Management. Spring 1990, pp. 41-46.*

Bird, Jane (1993). A System for the Perplexed. Management Today.April 1993, pp. 71-74.

Carroll, Chris and Larkin, Chris (1992). Executive Information *Technology: A Strategic Necessity at Motorola Codex. Information Systems Management. Summer 1992, pp. 21-29.*

Cook, Gary J. and Swain, Monte R. (1993). A Computerised Approach *to Decision Process Tracing for Decision Support System Design. Decision Sciences. Volume 24, pp. 931-952.*

Cooper, Robin and Kaplan, Robert S. (1991). Profit Priorities from *Activity-Based Costing. Harvard Business Review. May-June* 1991, pp. 130-135.

Crockett, Fess (1992). Revitalising Executive Information Systems.

Sloan Management Review. Summer 1992, pg 39-47.

Daft, Richard L. and Lengel, Robert H. (1986). Organisational *Information Requirements, Media Richness and Structural Design. Management Science. Volume 32, pp. 554-571.*

Davenport, Thomas H., Eccles, Robert G., and Prusak, Laurence. (1992). Information Politics. Sloan Management Review. Fall 1992, pp. 53-65.

Dearden, John. (1966). Myth of Real-Time Management Information. *Harvard Business Review. May-June 1966, pp. 123-132.*

Denton, Keith D. (1991). What's Wrong With these Employees? *Business Horizons. September-October 1991, pp. 45-49.*

Fisher, Marshall L., Hammond, Janice H., Obermeyer, Walter R., and Raman, Anath. (1994). Making Supply Meet Demand in an *Uncertain World. Harvard Business Review. May-June 1994, pp.* 83-93.

Forgionne, Guisseppi A. (1991). Decision Technology Systems: A Step *toward Complete Decision Support. Information Systems Managemen. Fall 1991, pp. 34-43.*

Grant, Rebecca A., Higgins, Christopher A., and Irving, Richard H. (1988). Computerised Performance Monitors: are They Costing *You Customers? Sloan Management Review. Spring 1988, pp.* 39-45.

Huseman, Richard C.; Miles, Edward W. (1988). Organisational *Communication in the Information Age: Implications of Computer-Based Systems. Journal of Management. 14(2), pp.* 181-204.

Jones, Jack William, and McLeod, Raymond Jr. (1986). The Structure *of Executive Information Systems: An Exploratory Analysis. Decision Sciences. Volume 17, pp. 220-249.*

Kallman, Ernest A. and McKinnon, William P. (1989). Case Study: *Evaluating Is By Using An Executive Information System. Journal of Information Systems Management. Fall 1989,* pp58-68.

Kaplan, Robert S., and Norton, David P. (1992). The Balanced *Scorecard - Measures That Drive Performance. Harvard Business Review. January-February 1992, pp. 71-79.*

Keen, Peter G.W. (1980). Decision Support Systems: Translating *Analytic Techniques into Useful Tools. Sloan Management Review. Spring 1980, pp. 33-44.*

Kelly, Floyd (1992a). Program Implementation: Measures of Success. *Working Paper, Canadian Ministry of Housing. December 9,* 1992.

King, William R., Premkumar, G., and Ramamurthy, K. (1990). An *Evaluation of the Role and Performance of a Decision Support System in Business Education. Decision Sciences. Volume 21, pp.* 642-659.

Knight, Charles F. (1992). Emerson Electric: Consistent Profits, *Consistently. Harvard Business Review. January-February 1992,* pp. 57-70.

Krachenberg, A. Richard, Henke, John W. Jr., and Lyons, Thomas F. (1993). The Isolation of Upper Management. Business Horizons. July-August 1993, pp. 41-47.

Leventhal, Laura Marie, Teasley, Barbee Mynatt, Instone, Keith, Rohlman, Diane Schertler, and Farhat, John. (1993). Sleuthing in Hyperhomes: an Evaluation of using Hypertext vs. A Book to Answer Questions. Behaviour and Information Technology. Volume 12, pp. 149-164.

McKinnon, Sharon M., and Bruns, William J. Jr. (1992). The *Information Mosaic, Harvard Business School Press. Boston MA.*

McLeod, Raymond Jr., and Jones, Jack W. (1986). Making Executive *Information Systems More Effective. Business Horizons.* September-October 1986, pp. 29-37.

Meyer, Christopher. (1994). How The Right Measures Help Teams *Excel. Harvard Business Review. May-June 1994, pp. 95-103.*

O'Shea, Timothy J. (1989). Low-Cost Approaches To Executive *Information Systems. Journal of Information Systems Management. Spring 1989, pp34-41.*

Osborne, David, and Gaebler, Ted. (1992). Reinventing Government. Addison-Wesley, Reading MA.

Posner, Bruce G., and Rothstein, Lawrence R. (1994). Reinventing the business of government: An Interview with Change Catalyst

David Osborne. Harvard Business Review. May-June 1994, pp. 132-143.

Preece, Jenny, Rogers, Yvonne, Sharp, Helen, Benyon, David, Holland, Simon, and Carey, Tom (1994). Human Computer Interaction. Addison-Wesley. Wokingham, England.

Rockart, John F. and DeLong, David W. (1986). Identifying the *Attributes of Successful Executive Support System Implementation. Transactions of the Sixth International Conference on Decision Support Systems.*

Rockart, John F. and Treacy, Michael E. (1982). The CEO Goes On-*Line. Harvard Business Review. January-February, 1982.*

Rodgers, T. J. (1990). No Excuses Management. Harvard Business *Review.July-August 1990, pp. 84-98.*

Sabourin, Helen (1992). Resource Management Information. Defence *2000 Achievements Report, pp. 44-47.*

Schaffer, Robert H. (1974). Demand Better Results - And Get Them. *Harvard Business Review. November-December 1974.*

Scott, Derek. (1993). Visual Search in Modern Human-Computer *Interfaces. Behaviour and Information Technology. Volume 12,* pp. 174-189.

Stevenson, Mark (1994). He Sees All, He Knows All. Canadian *Business. Spring 1994, pp. 30-35.*

Tan, Joseph K. H. and Benbasat, Izak. (1993). The Effectiveness Of *Graphical Presentation for Information Extraction: A Cumulative Experimental Approach. Decision Sciences. Volume* 24, pp. 167-191.

Umanath, Narayan S., Scamell, Richard W., and Das, Sidhartha R. (1990). An Examination of Two Screen/Report Design Variables *in an Information Recall Context. Decision Sciences. Volume 21,* pp. 216-240.

Vandenbosch, Betty. (1993). Executive Support Systems Impact Viewed *from a Learning Perspective. Doctoral Dissertation, University of* Western Ontario.

Determining Information Requirements. Information Systems

Management. Spring 1992, pp. 37-43.

Watson, Hugh J.; Rainer, R. Kelly Jr. (1991). A Manager's Guide to *Executive Support Systems. Business Horizons. March-April* 1991, pp. 44-50.

MODULE 3

UNIT 1 GEOGRAPHIC INFORMATION SYSTEM

CONTENTS

1.0 INTRODUCTION

A geographic information system (GIS), also known as a **geographical information system, is an information system for** capturing, storing, analysing, managing and presenting data which are spatially referenced (linked to location).

In the strictest sense, it is any information system capable of integrating, storing, editing, analysing, sharing, and displaying geographically referenced information. In a more generic sense, GIS applications are tools that allow users to create interactive queries (user created searches), analyse spatial information, edit data, maps, and present the results of all these operations. Geographic information science is the science underlying the geographic concepts, applications and systems, taught in degree and GIS Certificate programs at many universities.

Geographic information system technology can be used for scientific investigations, resource management, asset management, environmental impact assessment, urban planning, cartography, criminology, geographic history, marketing, and logistics to name a few. For example, GIS might allow emergency planners to easily calculate emergency response times in the event of a natural disaster, GIS might be used to find wetlands that need protection from pollution, or GIS can be used by a company to site a new business location to take advantage of a previously under-served market.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define geographic information systems (GIS)
- identify the major components of GIS
- explain the steps to take in developing an effective GIS
- evaluate the development and history of the concept of GIS
- identify and understand the various techniques adopted in GIS execution.

3.0 MAIN CONTENT

3.1 History of Development

About 15,500 years ago, on the walls of caves near Lascaux, France, Cro-Magnon hunters drew pictures of the animals they hunted. Associated with the animal drawings are track lines and tallies thought to depict migration routes. While simplistic in comparison to modern technologies, these early records mimic the two-element structure of modern geographic information systems, an image associated with attribute information.

The year 1962 saw the development of the world's first true operational GIS in Ottawa, Ontario, Canada by the Federal Department of Forestry and Rural Development. Developed by Dr. Roger Tomlinson, it was called the "Canada Geographic Information System" (CGIS) and was used to store, analyse, and manipulate data collected for the Canada Land Inventory (CLI).

By the early 1980s, M&S Computing (later Intergraph), Environmental Systems Research Institute (ESRI) and CARIS (Computer Aided Resource Information System) emerged as commercial vendors of GIS software, successfully incorporating many of the CGIS features, combining the first generation approach to separation of spatial and attribute information with a second generation approach to organising attribute data into database structures.

3.2 Components of GIS

The basic components of a GIS are software, data and hardware.

GIS Software

As with other information technology, GIS must be implemented in a manner that easily allows applications to support your organisation's workflows and business requirements.

How to Choose

By following the 10-stage implementation process outlined here, you should have a good idea of what GIS software best meets your needs.
The software you choose depends on the information products you want, your data requirements, and the functionality requirements you discovered during the Determine Technology Requirements stage.
Basic things to consider include:

- Will I be successful if I use this software from this vendor?
- Is the software based on common IT standards?
- Will support be adequate from the vendor?

What Every GIS Platform Should Have

GIS software can be increasingly thought of as IT infrastructure for assembling large, sophisticated multi-user systems. Look for a GIS platform that provides all the capabilities necessary to suppor this vision:

- Geographic database to store and manage all geographic objects
- A Web-based network for distributed geographic information management and sharing
- Desktop and server applications for;
- Data compilation
- Information query
- Spatial analysis and geo-processing
- Cartographic production
- Image visualisation and exploitation
- GIS data management
- Modular software components (engines) to embed GIS logic in other applications and build custom applications
- Geographic information services for multitier and centralised GIS systems
- A comprehensive GIS platform that meets all geographic requirements

GIS Data

The backbone of GIS is good data. Inaccurate data can result in inaccurate models and maps, skewing the results of your analysis and ultimately resulting in poor decisions. "Garbage in, garbage out," as the adage says.

The past 10 years has seen an explosion in the amount of data available, much of it free, with the advent of the Internet and proliferation of commercial sources of data. Internet mapping and Web services technology has made it possible for anyone anywhere to share or access data from around the globe.

This wide availability makes it critical to understand what GIS data is, how it is used, and how to select the right data for your needs.

Data Types: Geography data comes in three basic forms:

1. **Map data. Map data contains the** location and shape of geographic features. Maps use threebasic shapes to present real-world features: points, lines, and areas (called polygons).

2. **Attribute data. Attribute (tabular) data is the descriptive data that** GIS links to map features. Attribute data is collected and compiled for specific areas like states, census tracts, cities, and so on and often comes packaged with map data. When implementing a GIS, the most common sources of attribute data are your own organisation's databases combined with data sets you buy or acquire from other sources to fill in gaps.

3. **Image data. Image data ranges from satellite images and aerial** photographs to scanned maps (maps that have been converted from printed to digital format).

Using Data in GIS

GIS stores information about the world as a collection of themed layers that can be used together. A layer can be anything that contains similar features such as customers, buildings, streets, lakes, or postal codes. This data contains either an explicit geographic reference, such as a latitude and longitude coordinate, or an implicit reference such as an address, postal code, census tract name, forest stand identifier, or road name.

To work, a GIS requires explicit references. A GIS can create these explicit references from implicit references by an automated process called geo-coding—tying an implicit reference like an address to a specific point on the earth.

Find Data

Data is free or fee-based and comes from your own organisation; commercial, nonprofit, educational, and governmental sources; and other GIS software users.

The following links provide a good start for finding data for your needs.

- ESRIData

Data created by experts and designed to fit your business needs.

- Geodata.gov

Geodata.gov is part of the Geospatial One-Stop E-Gov initiative providing access to geospatial data and information.

- Geography Network

The Geography Network is a collaborative, multi-participant system for publishing, sharing, and using digital geographic information on the Internet.

- GISDataDepot

This site offers free data from all over the world, or you can pay a fee for a custom-cut CD.

Hardware—GIS Infrastructure

Hardware is really a simplistic term used to describe the technology infrastructure needed to support your GIS implementation. The infrastructure you develop depends on the system requirements you determined you needed during that phase of implementation planning.
Using Web services for your GIS needs requires minimal investment in infrastructure, while an enterprise GIS implementation requires careful planning and a fairly significant investment in computer equipment, network technology, database connectivity, and whatever else is required.

GIS is very much part of the information technology (IT) infrastructure for any organisation, not a separate or niche technology. Your IT group will help you determine the infrastructure you need and ensure that it is compatible with the overall IT infrastructure for the organisation. Depending on the scale of your implementation, your IT group will help you determine how to address issues such as security, data integrity, compatibility with other critical IT infrastructure and applications, and so on. Partnering early on with your IT group in planning the implementation of your GIS will significantly increase the success of your implementation and use of GIS.

3.3 Implementing GIS

Careful planning is the key to successfully implementing GIS in the enterprise.
Implementing GIS presents a unique set of challenges. Even the most well-funded projects can fail because of poor planning.

Here we outline a 10-stage process for successfully deploying GIS.

3. Consider the strategic purpose.
4. Plan for the planning.
5. Determine technology requirements.
6. Determine the end products.
7. Define the system scope.
8. Create a data design.
9. Choose a data model.
10. Determine system requirements.
11. Analyse benefits and costs.
12. Make an implementation plan.

All deployment situations are unique. To ensure the greatest chance of success, it is important that you understand all the stages outlined before adapting the methodology to your needs.

Consider the Strategic Purpose

When planning for a successful GIS, you must think first about what you want from it. Everything else you do come from that vision to ensure that the GIS will ultimately work towards serving some useful and important purpose.

Start by considering the strategic purpose of the organisation within which the system will be developed. What are its goals, objectives, and mandates?

This stage of the planning ensures that the process and the final system fit within organisational context and truly support the strategic objectives of the organisation. This stage also allows you to assess how information created by the GIS will impact the business strategy of the organisation.

Plan for the Planning

GIS planning should not be taken lightly. Forget about actually implementing a GIS for the moment. Just planning a GIS takes a commitment of resources and people.

Before you begin, you need to know that your organisation understands the distinction between planning and implementing and that it is prepared to provide the resources needed to make the planning happen. Making the case means understanding what needs to be done and what it will take to get it done. The end result of this stage is a project proposal that makes that case and explicitly seeks approval to launch the formal planning process.

Commitment to the planning process is essential to a successful GIS implementation, especially in municipal government agencies and other public-sector organisations. The project proposal helps to secure the political commitment to the planning process. This is the moment to introduce the GIS planning process to the most senior executives of your organisation. Arrange to keep them fully informed of th progress. If you receive approval for your planning projec and a commitment of resources at this point, your chances o having a successful GIS are high.

Determine Technology Requirements

Conduct a Technology Seminar

Once you have the approval for your project plan, the in-house GIS planning team can be activated. Defining the specific GIS requirements is the primary task in the planning process. You must meet with your customers or clients (those who will use the system or use the output from the system) to begin gathering specifics about the actual requirements from the user's perspective.

A highly effective method of soliciting input on the needs of your organisation is to hold one or more in-house technology seminars.

In addition to its information-gathering purpose, the technology seminar is also an ideal opportunity for you to explain to key personnel the nature of GIS, its potential benefits, and the planning process itself. By involving stakeholders at this early stage, you help to ensure participation in the subsequent planning work ahead so that all participants will appreciate the scope of the planning process. The technology seminar is also the place where initial identification of information products begins.

Knowing what you want to get out of your GIS is the key to a successful implementation. The "stuff" you get from your GIS can be described in the form of "information products."

This stage must be carefully undertaken. You'll need to talk to the users about what their job involves and what information they need to perform their tasks. Ultimately, you need to determine things like how the information products should be made and how frequently, what data is needed to make them, how much error can be tolerated, and the benefits of the new information produced.
This stage should result in a document that includes a description of all the information products that can be reasonably foreseen, together with details of the data and functions required to produce these products.

Define the System Scope

Once the information products have been described, you can begin to define the scope of the entire system. This involves determining what data to acquire, when it will be needed, and the data volumes that need to be handled. This includes assessing the probable timing of the production of the information products. It may become clear that it will be possible to use one input data source to generate more than one information product, and this can now be built into your development programme. Each refinement helps clarifyyour needs and increases your chance of success.

Create a Data Design

In GIS, data is a major factorbecause spatial data is a relatively complicated thing. In the conceptual system design phase of the planning process, you review the requirements identified in the earlier stages and use them to begin developing database design.

A good GIS project plan includes collaboration and an iterative approach to database design. Designing the conceptual, logical, and physical structure of the database will involve input from key people including project managers, subjectmatter experts, and technical resources.

Choose a Logical Data Model

A logical data model describes those parts of the real world that concern your organisation. The database may be simple or complex but must fit together in a logical manner so that you can easily retrieve the data you need and efficiently carry out the analysis tasks required.

There are several options available for your system's database design. You should also consider data accuracy, update requirements error tolerance, and data standards at this stage, as these issues will affect system design.

Determine System Requirements

The system requirements stage is where you examine the system functions and user interface needed, along withthe interface, communications, hardware, and software requirements. This should be the first time in the planning process that you examine software and hardware products.

By reviewing the information product descriptions, you will summarise the functions needed to produce them. Issues of interface design effective communications (particularly in distributed systems), and appropriate hardware and software configurations should also be considered during this planning phase. Plan your system carefully.

GIS Cost Waterfall

Cost of a change:

$1 – in the requirements stage
$10 – in the design stage
$100 – in the construction stage
$1000 – in the implementation stage

Source: Database Programming & Design, October, 1996, p. 57-64

Analyse Benefits and Costs

Following conceptual system design, you need to work out the best way to implement the system you have designed. This is where you plan how the system will be taken from the planning stage to actual implementation. You may also need to conduct a benefit-cost analysis to make your business case for the system.

Until now, the focus of the planning methodology has been on what you need to put in place to meet your requirements. The focus at this stage switches to how to put the system in place—an acquisition plan. Issues such as institutional interactions, legal matters, existing legacy hardware and software, security, staffing, and training are addressed at this stage.

The acquisition plan that results from this stage of the planning process will contain your implementation strategy and benefit-cost analysis. This plan can be used both to secure funding for your system and as a guide for the actual implementation of the system.

Make an Implementation Plan

The final report equips you with all the information you need to implement a successful GIS. It will become your GIS planning book to help you through the implementation process. Developing the final report should be the result of a process of communication between the GIS team and management, so that no parts of the report come as a surprise to anyone.

The report should contain a review of the organisation's strategic business objectives, the information requirements study, details of the conceptual system design, recommendations for implementation, time planning issues, and funding alternatives.

3.4 Techniques used in GIS

Data Creation

Modern GIS technologies use digital information, for which various digitised data creation methods are used. The most common method of data creation is digitisation, where a hard copy map or survey plan is transferred into a digital medium through the use of a computer-aided design (CAD) program, and geo-referencing capabilities. With the wide availability of ortho-rectified imagery (both from satellite and aerial sources), heads-up digitising is becoming the main avenue through which geographic data is extracted. Heads-up digitising involves the tracing of geographic data directly on top of the aerial imagery instead of through the traditional method of tracing the geographic form on a separate digitising tablet.

Relating Information from Different Sources

If you could relate information about the rainfall of your state to aerial photographs of your country, you might be able to tell which wetlands dry up at certain times of the year. A GIS, which can use information from many different sources in many different forms, can help with such analyses. The primary requirement for the source data consists of knowing the locations for the variables. Location may be annotated by x, y, and z coordinates of longitude, latitude, and elevation, or by other geo-code systems like ZIP Codes or by highway mile markers. Any variable that can be located spatially can be fed into a GIS. Several computer databases that can be directly entered into a GIS are being produced by government agencies and non-government organisations. Different kinds of data in map form can be entered into a GIS.

Data Representation

GIS data represents real world objects (roads, land use, elevation) with digital data. Real world objects can be divided into two abstractions: discrete objects (a house) and continuous fields (rain fall amount or elevation). There are two broad methods used to store data in a GIS for both abstractions: Raster and Vector.

Raster: A raster data type is, in essence, any type of digital image. Anyone who is familiar with digital photography will recognise the pixel as the smallest individual unit of an image. A combination of these pixels will create an image, distinct from the commonly used scalable vector graphics which are the basis of the vector model. While a digital image is concerned with the output as representation of reality, in a photograph or art transferred to computer, the raster data type will reflect an abstraction of reality. Aerial photos are one commonly used form of raster data, with only one purpose, to display a detailed image on a map or for the purposes of digitisation. Other raster data sets will contain information regarding elevation, a DEM, or reflectance of a particular wavelength of light, LANDSAT.

Digital elevation model, map (image), and vector data

Raster data type consists of rows and columns of cells, with each cell storing a single value. Raster data can be images (raster images) with each pixel (or cell) containing a color value. Additional values recorded for each cell may be a discrete value, such as land use, a continuous value, such as temperature, or a null value if no data is available.

Vector: A simple vector map, using each of the vector elements: points for wells, lines for rivers, and a polygon for the lake.

In a GIS, geographical features are often expressed as vectors, by considering those features as geometrical shapes. Different geographical features are expressed by different types of geometry.

Advantages and Disadvantages: There are advantages and disadvantages to using a raster or vector data model to represent reality. Raster data sets record a value for all points in the area covered which may require more storage space than representing data in a vector format that can store data only where needed. Raster data also allows easy implementation of overlay operations, which are more difficult with vector data. Vector data can be displayed as vector graphics used on traditional maps, whereas raster data will appear as an image that, depending on the resolution of the raster file, may have a blocky appearance for object boundaries. Vector data can be easier to register, scale, and re-project. This can simplify combining vector layers from different sources. Vector data are more compatible with relational database environment. They can be part of a relational table as a normal column and processes using a multitude of operators.

The file size for vector data is usually much smaller for storage and sharing than raster data. Image or raster data can be 10 to 100 times larger than vector data depending on the resolution. Another advantage of vector data is that it can be easily updated and maintained.

Non-Spatial Data: Additional non-spatial data can also be stored besides the spatial data represented by the coordinates of vector geometry or the position of a raster cell. In vector data, the additional data are attributes of the object. For example, a forest inventory polygon may also have an identifier value and information about tree species. In raster data the cell value can store attribute information, but it can also be used as an identifier that can relate to records in another table.

There is also software being developed to support spatial and non-spatial decision-making. In this software, the solutions to spatial problems are integrated with solutions to non-spatial problems. The end result it is hoped, with these Flexible Spatial Decision-Making Support Systems (FSDSS) will be that non experts can use GIS and spatial criteria with their other non spatial criteria to view solutions to multi-criteria problems that will support decision making.

Data Capture

Data capture—entering information into the system—consumes much of the time of GIS practitioners. There are a variety of methods used to enter data into a GIS where it is stored in a digital format.

Existing data printed on paper or PET film maps can be digitised or scanned to produce digital data. A digitiser produces vector data as an operator traces points, lines, and polygon boundaries from a map. Scanning a map results in raster data that could be further processed to produce vector data.

Survey data can be directly entered into a GIS from digital data collection systems on survey instruments using a technique called Coordinate Geometry (COGO). Positions from a Global Positioning System (GPS), another survey tool, can also be directly entered into a GIS.

Remotely sensed data also plays an important role in data collection and consist of sensors attached to a platform. Sensors include cameras, digital scanners and LIDAR, while platforms usually consist of aircraft and satellites.

Satellite remote sensing provides another important source of spatial data. Here satellites use different sensor packages to passively measure the reflectance from parts of the electromagnetic spectrum or radio waves that were sent out from an active sensor such as radar. Remote

sensing collects raster data that can be further processed using different bands to identify objects and classes of interest, such as land cover.

Raster-to-Vector Translation

Data restructuring can be performed by a GIS to convert data into different formats. For example, a GIS may be used to convert a satellite image map to a vector structure by generating lines around all cells with the same classification, while determining the cell spatial relationships, such as adjacency or inclusion.

4.0 CONCLUSION

Geographic information systems have come of age as a very important tool in decision making in environmental sciences and weather forecasting. GIS combines computerised maps and databases to organise and analyse location-specific information. It assists in assessing and forecasting property damage and co-ordinating relief efforts in cases of natural disaster as in Hurricanes and Tornados.

5.0 SUMMARY

- A geographic information system (GIS), also known as a geographical information system, is an information system for capturing, storing, analysing, managing and presenting data which are spatially referenced (linked to location).
- About 15,500 years ago, on the walls of caves near Lascaux, France, Cro-Magnon hunters drew pictures of the animals they hunted. Associated with the animal drawings are track lines and tallies thought to depict migration routes.
- The basic components of a GIS are software, data and hardware.
- Careful planning is the key to successfully implementing GIS in the enterprise. Implementing GIS presents a unique set of challenges. Even well-funded projects can fail because of poor planning.
- A GIS can also convert existing digital information, which may not yet be in map form, into forms it can recognise and use.
- Modern GIS technologies use digital information, for which various digitised data creation methods are used.
- GIS data represents real world objects (roads, land use, elevation) with digital data. Real world objects can be divided into two abstractions: discrete objects (a house) and continuous fields (rain fall amount or elevation).
- Satellite remote sensing provides another important source of spatial data. Here satellites use different sensor packages to passively measure the reflectance from parts of the electromagnetic spectrum or radio waves that were sent out from an active sensor such as radar.

- There are advantages and disadvantages to using a raster or vector data model to represent reality.

6.0 TUTOR-MARKED ASSIGNMENT

2. List the 10 stages to successfully deploy GIS.
3. Discuss briefly data capture as a technique used in GIS.

7.0 REFERENCES/FURTHER READING

GIS.com (2008). The Guide to Geographical Information Systems, *Implementing GIS*.

UNIT 2 EXPERT SYSTEMS

CONTENTS

1.0 INTRODUCTION

An expert system is a software system that attempts to reproduce the performance of one or more human experts, most commonly in a specific problem domain, and is a traditional application and/or subfield of artificial intelligence. A wide variety of methods can be used to simulate the performance of the expert.

However common to most or all are:

1) The creation of a so-called "knowledgebase" which uses some knowledge representation formalism to capture the subject matter experts (SME) knowledge,
2) A process of gathering that knowledge from the SME and codifying it according to the formalism, which is called knowledge engineering and
3) Expert systems may or may not have learning components but a third common element is that once the system is developed it is proven by being placed in the same real world problem solving situation as the human SME, typically as an aid to human workers or a supplement to some information system.

As a premiere application of computing and artificial intelligence, the topic of expert systems has many points of contact with general systems theory, operations research, business process reengineering and various topics in applied mathematics and management science.

Two illustrations of actual expert systems can give an idea of how they work. In one real world case at a chemical refinery a senior employee was about to retire and the company was concerned that the loss of his expertise in managing a fractionating tower would severely impact operations of the plant. A knowledge engineer was assigned to produce an expert system reproducing his expertise saving the company the loss of the valued knowledge asset. Similarly a system called Mycin was developed from the expertise of best diagnosticians of bacterial infections whose performance was found to be good as or better than the average clinician. An early commercial success and illustration of another typical application (a task generally considered overly complex for a human)was an expert system fielded by DEC in the 1980s to quality check the configurations of their computers prior to delivery. The eighties were the time of greatest popularity of expert systems and interest lagged after the onset of the AI Winter.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define the concept of expert systems
- explain some major expert systems and languages of the shelf
- identify the topical issues related to expert systems
- explain the various applications of expert systems
- evaluate the advantages and disadvantages associated with expert systems.

3.0 MAIN CONTENT

3.1 Overview

The most common form of expert system is a computer program, with a set of rules that analyses information (usually supplied by the user of the system) about a specific class of problems, and recommends one o more courses of user action. The expert system ma also provide mathematical analysis of the problem(s). The expert system utilises what appears to be reasoning capabilities to reach conclusions.

A related term is wizard. A wizard is an interactive computer program that helps a user to solve a problem. Originally the term wizard was used for programs that construct a database search query based on criteria supplied by the user. However, some rule-based expert systems are also called wizards. Other "Wizards" are a sequence of online forms that guide users through a series of choices, such as the ones which manage the installation of new software on computers, and these are not expert systems.

3.2 Prominent Expert Systems and Languages

- ART,- An early general-purpose programming language used in the development of expert systems.
- CADUCEUS (expert system) - Blood-borne infectious bacteria.
- CLIPS - Programming language used in the development of expert systems
- Drools - An open source offering from JBOSS labs
- Dendral - Analysis of mass spectra
- Dipmeter Advisor - Analysis of data gathered during oil exploration
- Jess - Java Expert System Shell. A CLIPS engine implemented in Java used in the development of expert systems
- KnowledgeBench – expert system for building new product development applications
- MQL 4 - MetaQuotes Language 4, a customised language for financial strategy programming
- Mycin - Diagnose infectious blood diseases and recommend antibiotics (by Stanford University)
- NETeXPERT - A mission-critical Operational SupportSystems framework with rules, policies, object modeling, and adapters for Network Operations Center automation
- NEXPERT Object - An early general-purpose commercial backwards-chaining inference engine used in the development of expert systems
- Prolog - Programming language used in the development of expert systems
- Forth - Programming language used in the development of expert systems
- R1/Xcon - Order processing
- SHINE Real-time Expert System - Spacecraft Health Inference Engine
- STD Wizard - Expert system for recommending medical screening tests
- PyKe - Pyke is a knowledge-based inference engine (expert system)

3.3 The Study of Expert Systems

Knowledge Representation

Knowledge Representation is an issue that arises in both cognitive science and artificial intelligence. In cognitive science, it is concerned with how people store and process information. In artificial intelligence (AI) the primary aim is to store knowledge so that programs can process it and achieve the verisimilitude of human intelligence. AI researchers have borrowed representation theories from cognitive science. Thus there are representation techniques such as frames, rules and semantic networks which have originated from theories of human information processing. Since knowledge is used to achieve intelligent behaviour, the fundamental goal of knowledge representation is to represent knowledge in a manner as to facilitate inference i.e. drawing conclusions from knowledge.

Knowledge Engineer

Knowledge engineers are concerned with the representation chosen for the expert's knowledge declarations and with the inference engine used to process that knowledge. He / she can use the knowledge acquisition component of the expert system to input the severa characteristics known to be appropriate to a good inference technique, including:

- A good inference technique is independent of the problem domain.
- In order to realise the benefits of explanation, knowledge transparency, and reusability of the programs in a new problem domain, the inference engine must not contain domain specific expertise.
- Inference techniques may be specific to a particular task, such as diagnosis of hardware configuration. Other techniques may be committed only to a particular processing technique.
- Inference techniques are always specific to the knowledge structures.
- Successful examples of rule processing techniques are forward chaining and backward chaining.

3.4 Expert Systems Topics

Chaining

There are two main methods of reasoning when using inference rules: backward chaining and forward chaining.

Forward chaining starts with the data available and uses the inference rules to conclude more data until a desired goal is reached. An inference engine using forward chaining searches the inference rules until it finds one in which the if-clause is known to be true. It then concludes the then-clause and adds this information to its data. It would continue to do this until a goal is reached. Because the data available determines which inference rules are used, this method is also called data driven.

Backward chaining starts with a list of goals andworks backwards to see if there is data which will allow it to conclude any of these goals. An inference engine using backward chaining would search the inference rules until it finds one which has a then-clause that matches a desired goal. If the if-clause of that inference rule is not known to be true, then it is added to the list of goals. For example, suppose a rule base contains two rules:

(1) If Fritz is green then Fritz is a frog.
(2) If Fritz is a frog then Fritz hops.

Suppose a goal is to conclude that Fritz hops. The rule base would be searched and rule (2) would be selected because its conclusion (the then clause) matches the goal. It is not known that Fritz is a frog, so this "if" statement is added to the goal list. The rule base is again searched and this time rule (1) is selected because its then clause matches the new goal just added to the list. This time, the if-clause (Fritz is green) is known to be true and the goal that Fritz hops is concluded. Because the list of goals determines which rules are selected and used, this method is called goal driven.

Certainty Factors

One advantage of expert systems over traditional methods of programming is that they allow the use of "confidences" (or "certainty factors"). When ahuman reasons he does not always conclude things with 100% confidence. He might say, "If Fritz is green, then he is probably a frog" (after all, he might be a chameleon). This type of reasoning can be imitated by using numeric values called confidences.
For example, if it is known that Fritz is green, it might be concluded with 0.85 confidence that he is a frog; or, if it is known that he is a frog, it might be concluded with 0.95 confidence that he hops. These numbers are similar in nature to probabilities, but they are not the same. They are meant to imitate the confidences humans use in reasoning rather than to follow the mathematical definitions used in calculating probabilities.

Expert System Architecture

The following general points about expert systems and their architecture have been illustrated.

1. The sequence of steps taken to reach a conclusion is dynamically synthesised with each new case. It is not explicitly programmed when the system is built.
2. Expert systems can process multiple values for any problem parameter. This permits more than one line of reasoning to be pursued and the results of incomplete (not fully determined) reasoning to be presented.
3. Problem solving is accomplished by applying specific knowledge rather than specific technique. This is a key idea in expert systems technology. It reflects the belief that human experts do

not process their knowledge differently from others, but they do possess different knowledge. With this philosophy, when one finds that their expert system does not produce the desired results, work begins to expand the knowledge base, not to re-program the procedures.

There are various expert systems in which a rule base and an inference engine cooperate to simulate the reasoning process that a human expert pursues in analysing a problem and arriving at a conclusion. In these systems, in order to simulate the human reasoning process a vast amount of knowledge needed to be stored in the knowledge base. Generally, the knowledge base of such an expert system consisted of a relatively large number of "if then" type of statements that were interrelated in a manner that, in theory at least, resembled the sequence of mental steps that were involved in the human reasoning process.

Because of the need for large storage capacities and related programs to store the rule base, most expert systems have, in the past, been run only on large information handling systems. Recently, the storage capacity of personal computers has increased to a point where it is becoming possible to consider running some types of simple expert systems on personal computers.

In some applications of expert systems, the nature of the application and the amount of stored information necessary to simulate the human reasoning process for that application is just too vast to store in the active memory of a computer. In other applications of expert systems, the nature of the application is such that not all of the information is always needed in the reasoning process. An example of this latter type application would be the use of an expert system to diagnose a data processing system comprising many separate components, some of which are optional. When that type of expert system employs a single integrated rule base to diagnose the minimum system configuration of the data processing system, much of the rule base is not required since many of the components which are optional units of the system will not be present in the system. Nevertheless, earlier expert systems require the entire rule base to be stored since all the rules were, in effect, chained or linked together by the structure of the rule base.

When the rule base is segmented, preferably into contextual segments or units, it is then possible to eliminate portions of the rule base containing data or knowledge that is not needed in a particular application. The segmenting of the rule base also allows the expert system to be run with systems or on systems having much smaller memory capacities than was possible with earlier arrangements since each segment of the rule base can be paged into and out of the system as needed. The segmenting of

the rule base into contextual segments requires that the expert system manage various intersegment relationships as segments are paged into and out of memory during execution of the program. Since the system permits a rule base segment to be called and executed at any time during the processing of the first rule base, provision must be made to store the data that has been accumulated up to that point so that at some time later in the process, when the system returns to the first segment, it can proceed from the last point or rule node that was processed. Also, provision must be made so that data that has been collected by the system up to that point can be passed to the second segment of the rule base after it has been paged into the system and data collected during the processing of the second segment can be passed to the first segment when the system returns to complete processing that segment.
The user interface and the procedure interface are two important functions in the information collection process.

End User

The end-user usually sees an expert system through an interactive dialogue, an example of which follows:

Q. Do you know which restaurant you want to go to?
A. No
Q. Is there any kind of food you would particularly like?
A. No
Q. Do you like spicy food?
A. No
Q. Do you usually drink wine with meals?
A. Yes
Q. When you drink wine, is it French wine?
A. Why

As can be seen from this dialogue, the system is leading the user through a set of questions, the purpose of which is to determine a suitable set of restaurants to recommend. This dialogue begins with the system asking if the user already knows the restaurant choice (a common feature of expert systems) and immediately illustrates a characteristic of expert systems; users may choose not to respond to any question. In expert systems, dialogues are not pre-planned. There is no fixed control structure. Dialogues are synthesised from the current information and the contents of the knowledge base. Because of this, not being able to supply the answer to a particular question does not stop the consultation.

Explanation System(move up to next page)

Another major distinction between expert systems and traditional systems is illustrated by the following answer given by the system when the user answers a question with another question, "Why", as occurred in the above example. The answer is:

A. I am trying to determine the type of restaurant to suggest. So far Chinese is not a likely choice. It is possible that French is a likely choice. I know that if the diner is a wine drinker, and the preferred wine is French, then there is strong evidence that the restaurant choice should include French.

It is very difficult to implement a general explanation system (answering questions like "Why" and "How") in a traditional computer program. An expert system can generate an explanation by retracing the steps of its reasoning. The response of the expert system to the question WHY is an exposure of the underlying knowledge structure. It is a rule; a set of antecedent conditions which, if true, allow the assertion of a consequent. The rule references values, and tests them against various constraints or asserts constraints onto them. This, in fact, is a significant part of the knowledge structure. There are values, which may be associated with some organising entity. For example, the individual diner is an entity with various attributes (values) including whether they drink wine and the kind of wine. There are also rules, which associate the currently known values of some attributes with assertions that can be made about other attributes. It is the orderly processing of these rules that dictates the dialogue itself.

Expert Systems versus Problem-Solving Systems

The principal distinction between expert systems and traditional problem solving programs is the way in which the problem related expertise is coded. In traditional applications, problem expertise is encoded in both program and data structures.
In the expert system approach all of the problem related expertise is encoded in data structures only; none is in programs. This organisation has several benefits.

An example may help contrast the traditional problem solving program with the expert system approach. The example is the problem of tax advice. In the traditional approach data structures describe the taxpayer and tax tables, and a program in which there are statements representing an expert tax consultant's knowledge, such as statements which relate information about the taxpayer to tax table choices. It is this representation of the tax expert's knowledge that is difficult for the tax expert to understand or modify.

In the expert system approach, the information about taxpayers and tax computations is again found in data structures, but now the knowledge describing the relationships between them is encoded in data structures as well. The programs of an expert system are independent of the problem domain (taxes) and serve to process the data structures without regard to the nature of the problem area they describe. For example, there are programs to acquire the described data values through user interaction, programs to represent and process special organisations of description, and programs to process the declarations that represent semantic relationships within the problem domain and an algorithm to control the processing sequence and focus.

The general architecture of an expert system involves two principal components: a problem dependent set of data declarations called the knowledge base or rule base, and a problem independent (although highly data structure dependent) program which is called the inference engine.

Individuals involved with Expert Systems

There are generally three individuals having an interaction with expert systems. Primary among these is the end-user; the individual who uses the system for its problem solving assistance. In the building and maintenance of the system there are two other roles: the problem domain expert who builds and supplies the knowledge base providing the domain expertise, and a knowledge engineer who assists the experts in determining the representation of their knowledge, enters this knowledge into an explanation module and who defines the inference technique required to obtain useful problem solving activity. Usually, the knowledge engineer will represent the problem solving activity in the form of rules which is referred to as a rule-based expert system. When these rules are created from the domain expertise, the knowledge base stores the rules of the expert system.

User Interface

The function of the user interface is to present questions and information to the user and supply the user's responses to the inference engine.

Any values entered by the user must be received and interpreted by the user interface. Some responses are restricted to a set of possible legal answers, others are not. The user interface checks all responses to ensure that they are of the correct data type. Any responses that are restricted to a legal set of answers are compared against these legal answers. Whenever the user enters an illegal answer, the user interface informs the user that his answer was invalid and prompts him to correct it.

3.5 Application of Expert Systems

Expert systems are designed and created to facilitate tasks in the fields of accounting, medicine, process control, financial service, production, human resources etc. Indeed, the foundation of a successfu expert system depends on a series of technical procedures and development that may be designed by certain technicians and related experts.

A good example of application of expert systems in banking area is expert systems for mortgages. Loan departments are interested in expert systems for mortgages because of the growing cost of labour which makes the handling and acceptance of relatively small loans less profitable. They also see in the application of expert systems a possibility for standardised, efficient handling of mortgage loan, and appreciate that for the acceptance of mortgages there are hard and fast rules which do not always exist with other types of loans.

While expert systems have distinguished themselves in AI research in finding practical application, their application has been limited. Expert systems are notoriously narrow in their domain of knowledge—as an amusing example, a researcher used the "skin disease" expert system to diagnose his rust bucket car as likely to have developed measles—and the systems were thus prone to making errors that humans would easily spot. Additionally, once some of the mystique had worn off most programmers realised that simple expert systems were essentially just slightly more elaborate versions of the decision logic they had already been using. Therefore, some of the techniques of expert systems can now be found in most complex programs without any fuss about them.

An example and a good demonstration of the limitations of, an expert system used by many people is the Microsoft Windows operating system troubleshooting software located in the "help" section in the taskbar menu. Obtaining expert / technical operating system support is often difficult for individuals not closely involved with the development of the operating system. Microsoft has designed their expert system to provide solutions, advice, and suggestions to common errors encountered throughout using the operating systems.

Another 1970s and 1980s application of expert systems — which we today would simply call AI — was in computer games. For example, the computer baseball games Earl Weaver Baseball and Tony La Russa Baseball each had highly detailed simulations of the game strategies of those two baseball managers. When a human played the game against the computer, the computer queried the Earl Weaver or Tony La Russa

Expert System for a decision on what strategy to follow. Even those choices where some randomness was part of the natural system (such as when to throw a surprise pitch-out to try to trick a runner trying to steal a base) were decided based on probabilities supplied by Weaver or La Russa. Today we would simply say that "the game's AI provided the opposing manager's strategy."

3.6 Advantages and Disadvantages

Advantages

- Provides consistent answers for repetitive decisions, processes and tasks
- Holds and maintains significant levels of information
- Encourages organisations to clarify the logic of their decision-making
- Never "forgets" to ask a question, as a human might.

Disadvantages

- Lacks common sense needed in some decision making
- Cannot make creative responses as human expert would in unusual circumstances
- Domain experts not always able to explain their logic and reasoning
- Errors may occur in the knowledge base, and lead to wrong decisions
- Cannot adapt to changing environments, unless knowledge base is changed.

3.7 Types of Problems Solved by Expert Systems

Expert systems are most valuable to organisations that have a high-level of know-how experience and expertise that cannot be easily transferred to other members. They are designed to carry the intelligence and information found in the intellect of experts and provide this knowledge to other members of the organisation for problem-solving purposes.

Typically, the problems to be solved are of the sort that would normally be tackled by a medical or other professional. Real experts in the problem domain (which will typically be very narrow, for instance "diagnosing skin conditions in human teenagers") are asked to provide "rules of thumb" on how they evaluate the problems, either explicitly with the aid of experienced systems developers, or sometimes implicitly, by getting such experts to evaluate test cases and using computer programs to examine the test data and (in a strictly limited manner) derive rules from that. Generally, expert systems are used for problems for which there is no single "correct" solution which can be encoded in a conventional algorithm — one would not write an expert system to find shortest paths through graphs, or sort data, as there are simply easier ways to do these tasks.

Simple systems use simple true/false logicto evaluate data. More sophisticated systems are capable of performing at least some evaluation, taking into account real-world uncertainties,usingsuch methods as fuzzy logic. Such sophistication is difficult to develop and still highly imperfect.

4.0 CONCLUSION

Expert systems describe computer programs which allow users to benefit from expert knowledge and information, and also advice. It is an offshoot of a much broader field of artificial intelligence. The specialist applications of expert systems are usually intended to widen the knowledge of a computer user who is already an expert in his field, rather than to provide information to a non-expert and novice but Presently non-experts are beginning to hook unto expert systems to proffer some level of solution to their problems.

5.0 SUMMARY

- An expert system is a software system that attempts to reproduce the performance of one or more human experts, most commonly in a specific problem domain, and is a traditional application and/or subfield of artificial intelligence.
- The most common form of expert system is a computer program, with a set of rules, that analyses information (usually supplied by the user of the system) about a specific class of problems, and recommends one or more courses of user action.
- Knowledge representation is an issue that arises in both cognitive science and artificial intelligence. In cognitive science, it is concerned with how people store and process information.
- There are two main methods of reasoning when using inference rules: backward chaining and forward chaining.
- Expert systems are designed and created to facilitate tasks in the fields of accounting, medicine, process control, financial service, production, human resources etc. Indeed, the foundation of a successful expert system depends on a series of technical procedures and development that may be designed by certain technicians and related experts.

6.0 TUTOR-MARKED ASSIGNMENT

1. Mention 5 characteristics of good inference techniques as used by knowledge engineers.

2. Provide 5 disadvantages of expert systems.

7.0 REFERENCES/FURTHER READING

James Ignizio, (1991). Introduction to Expert Systems. ISBN: 0-07-909785-5.

Joseph C. Giarratano, Gary Riley (2005). Expert Systems, Principles *and Programming. ISBN: 0-534-38447-1*.

Peter Jackson (1998). Introduction to Expert Systems. ISBN: 0-201-87686-8.

WALKER, Adrian et al (1990). Knowledge Systems and Prolog. Addison-Wesley. ISBN: 0-201-52424-4.

UNIT 3 ARTIFICIAL INTELLIGENCE

CONTENTS

1.0 INTRODUCTION

What is Artificial Intelligence

Organisations are using artificial intelligence (AI) to capture individual and collective knowledge and to codify and internalise their knowledge base.

AI can be defined as the effort to develop computer-based system (both hardware and software) that behaves as humans. Such systems will be to learn natural languages, accomplish coordinated physical tasks (robotics), use perceptual apparatus that inform their physical behaviour and language (oral and visual perception systems), and emulate human expertise and decision making (expert systems). Such systems also exhibit logic, reasoning, intuition, and the just-plain-common-sense qualities that we associate with human beings. Another important element is intelligence machines, the physical hardware that performs these tasks.

Successful AI systems are based on human expertise, knowledge and selected reasoning patterns, but they do not exhibit the intelligence of human beings. Existing intelligence systems do not come up with new and novel solutions to problems. Existing systems extend power ex experts but in no way substitute for them or capture much of their intelligence. Briefly, existing systems lack the common sense and generality of naturally intelligent human beings.

2.0 OBJECTIVES

At the end of this unit, the students should be able to:

- define and understand the concept of artificial intelligence
- explain the major sub-field of artificial intelligence (AI)
- outline the development and history of AI
- identify AI characteristics that appeal to business functions
- identify the various tools used in the field of AI

- evaluate the major applications of AI in lifegenerally.
- .

3.0 MAIN CONTENT

3.1 Components of AI

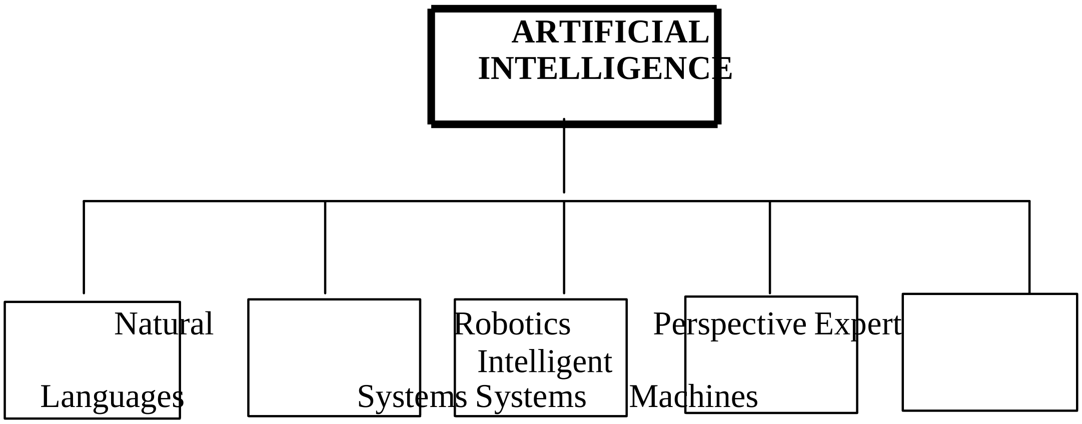


Figure 1: Components of Artificial Intelligence

3.2 Why Business is interested in AI

Although AI applications are much more limited than human intelligence, they are of great interest to business for the following reasons:

- To preserve expertise that might be lost through the retirement, resignation, or death of an acknowledged expert.
- To store information in an active form- to create an organisational knowledge base- that many employees can examine, much like electronic textbook or manual so that other may learn rules of thumb not found in the textbooks.
- To create a mechanism that is not subject to human feelings such as fatigue and worry. This may be especially useful when jobs may be environmentally, physically, or mentally dangerous to humans. These systems may also be useful advisers in times of crisis.
- To eliminate routine and unsatisfying jobs held by people.
- To enhance the organisation's knowledge base by suggesting solutions to specific problems that are too massive and complex to be analysed by human beings in a short period of time.

3.3 Tools of AI Research

In the course of 50 years of research, AI has developed a large number of tools to solve the most difficult problems in computer science. A few of the most general of these methods are discussed below.

Search and Optimisation

Many problems in AI can be solved in theory by intelligently searching through many possible solutions: Reasoning can be reduced to performing a search. For example, logical proof can be viewed as searching for a path that leads from premises to conclusions, where each step is the application of an inference rule. Planning algorithms search through trees of goals and sub goals, attempting to find a path to a target goal, a process called means-ends analysis. Robotics algorithms for moving limbs and grasping objects use local searches in configuration space. Many learning algorithms use search algorithms based on optimisation.

Simple exhaustive searches are rarely sufficient for most real world problems: the search space (the number of places to search) grows quickly to astronomical numbers. The result is a search that is too slow or never completes. The solution, for many problems, is to use "heuristics" or "rules of thumb" that eliminate choices that are unlikely to lead to the goal (called "pruning the search tree"). Heuristics supply the program with a "best guess" for what path the solution lies on.

A very different kind of search came to prominence in the 1990s, based on the mathematical theory of optimisation. For many problems, it is possible to begin the search with a form of a guess and then refine the guess incrementally until no more refinements can be made. These algorithms can be visualised as blind hill climbing: we begin the search at a random point on the landscape, and then, by jumps or steps, we keep moving our guess uphill, until we reach the top. Other optimisation algorithms are simulated annealing, beam search and random optimisation.

Evolutionary computation uses a form of optimisation search. For example, they may begin with a population of organisms (the guesses) and then allow them to mutate and recombine, selecting only the fittest to survive each generation (refining the guesses). Forms of evolutionary computation include swarm intelligence algorithms (such as ant colony or particle swarm optimisation) and evolutionary algorithms (such as genetic algorithms and genetic programming).

Logic

Logic was introduced into AI research by John McCarthy in his 1958

Advice Taker proposal. The most important technical development was

J. Alan Robinson's discovery of the resolution and unification algorithm for logical deduction in 1963. This procedure is simple, complete and entirely algorithmic, and can easily be performed by digital computers. However, a naive implementation of the algorithm quickly leads to a combinatorial explosionor an infinite loop. In 1974, Robert Kowalski suggested representing logical expressions as Horn clauses (statements in the form of rules: "if p then q"), which reduced logical deduction to backward chaining or forward chaining. This greatly alleviated (but did not eliminate) the problem.

Logic is used for knowledge representation and problem solving, but it can be applied to other problems as well. For example, the satplan algorithm uses logic for planning, and inductive logic programming is a method for learning.

There are several different forms of logic used in AI research.

- Propositional or sentential logic is the logic of statements which can be true or false.
- First-order logic also allows the use of quantifiers and predicate and can express facts about objects, their properties, and their relations with each other.
- Fuzzy logic, a version of first-order logic which allows the truth of a statement to be represented as a value between 0 and 1, rather than simply True (1) or False (0). Fuzzy systems can be used for uncertain reasoning and have been widely used in modern industrial and consumer product control systems.
- Default logics, non-monotonic logics and circumscription are forms of logic designed to help with default reasoning and the qualification problem.
- Several extensions of logic have been designed to handle specific domains of knowledge, such as: description logics; situation calculus, event calculus and fluent calculus (for representing events and time) causal calculus; belief calculus; and modal logics.

Probabilistic Methods for Uncertain Reasoning

Many problems in AI (in reasoning, planning, learning, perception and robotics) require the agent to operate with incomplete or uncertain information. Starting in the late 80s and early 90s, Judea Pearl and others championed the use of methods drawn from probability theory and economics to devise a number of powerful tools to solve these problems[120].

Bayesian networks are very general tool that can be used for a large number of problems: reasoning (using the Bayesian inference algorithm), learning (using the expectation-maximisation algorithm), planning (using decision networks) and perception (using dynamic Bayesian networks).

Probabilistic algorithms can also be used for filtering, prediction, smoothing and finding explanations for streams of data, helping perception systems to analyse processes that occur over time[126] (e.g., hidden Markov models and Kalman filters).

Planning problems have also taken advantages of other tools from economics, such as decision theory and decision analysis, information value theory, Markov decision processes, dynamic decision networks game theory and mechanism design.

Classifiers and Statistical Learning Methods

The simplest AI applications can be divided into two types: classifiers ("if shiny then diamond") and controllers ("if shiny then pick up"). Controllers do however also classify conditions before inferring actions, and therefore classification forms a central part of many AI systems.

Classifiers are functions that use pattern matching to determine a closest match. They can be tuned according to examples, making them very attractive for use in AI. These examples are known as observations or patterns. In supervised learning, each pattern belongs to a certain predefined class. A class can be seen as a decision that has to be made. All the observations combined with their class labels are known as a data set.

When a new observation is received, that observation is classified based on previous experience. A classifier can be trained in various ways; there are many statistical and machine learning approaches.

A wide range of classifiers are available, each with its strengths and weaknesses. Classifier performance depends greatly on the characteristics of the data to be classified. There is no single classifier that works best on all given problems; this is also referred to as the "no free lunch" theorem. Various empirical tests have been performed to compare classifier performance and to find the characteristics of data that determine classifier performance. Determining a suitable classifier for a given problem is however still more an art than science.

The most widely used classifiers are the neural network, kernel methods such as the support vector machine, k-nearest neighbor algorithm, Gaussian mixture model, naive Bayes classifier, and decision tree. The performances of these classifiers have been compared over a wide range of classification tasks [139] in order to find data characteristics that determine classifier performance.

Neural Networks

A neural network is an interconnected group of nodes, akin to the vast network of neurons in the human brain.

The study of artificial neural networks began in the decade before the field AI research was founded. In the 1960s Frank Rosenblatt developed an important early version, the perceptron. Paul Werbos developed the back propagation algorithm for multilayer perceptrons in 1974, which led to a renaissance in neural network research and connectionism in general in the middle 1980s. The Hopfield net, a form of attractor network, was first described by John Hopfield in 1982.

Common network architectures which have been developed include the feedforward neural network, the radial basis network, the Kohonen self-organising map and various recurrent neural networks. Neural networks are applied to the problem of learning, using such techniques as Hebbian learning, competitive learning and the relatively new field of Hierarchical Temporal Memory which simulates the architecture of the neocortex.

3.4 Specialised Languages

AI researchers have developed several specialised languages for AI research:

•IPL, one of the first programming languages, developed by Alan Newell, Herber Simon and J. C. Shaw.
•Lisp was developed by John McCarthy at MIT in 1958. There are many dialects of Lisp in use today.
•Prolog, a language based on logic programming, was invented by French researchers Alain Colmerauer and Phillipe Roussel, in collaboration with Robert Kowalski of the University of Edinburgh.
•STRIPS, a planning language developed at Stanford in the 1960s.
•Planner developed at MIT around the same time.

AI applications are also often written in standard languages like C++ and languages designed for mathematics, such as Matlab and Lush.

3.5 Evaluating Artificial Intelligence

How can one determine if an agent is intelligent? In 1950, Alan Turing proposed a general procedure to test the intelligence of an agent now known as the Turing test. This procedure allows almost all the major problems of artificial intelligence to be tested. However, it is a very difficult challenge and at present all agents fail.

Artificial intelligence can also be evaluated on specific problems such as small problems in chemistry, hand-writing recognition and game-playing. Such tests have been termed subject matter expert Turing tests. Smaller problems provide more achievable goals and there are an ever-increasing number of positive results.

The broad classes of outcome for an AI test are:

- optimal: it is not possible to perform better
- strong super-human: performs better than all humans
- super-human: performs better than most humans
- sub-human: performs worse than most humans

For example, performance at checkers (draughts) is optimal, performance at chess is super-human and nearing strong super-human and performance at many everyday tasks performed by humans is sub-human.

3.6 Competitions and Prizes

There are a number of competitions and prizes to promote research in artificial intelligence. The main areas promoted are: general machine intelligence, conversational behaviour, data-mining, driverless cars, robot soccer and games.

3.7 Applications of Artificial Intelligence

Artificial intelligence has been used in a wide range of fields including medical diagnosis, stock trading, robot control, law, scientific discovery and toys.

Many practical applications are dependent on artificial neural networks, networks that pattern their organisation in mimicry of a brain's neurons, which have been found to excel in pattern recognition.

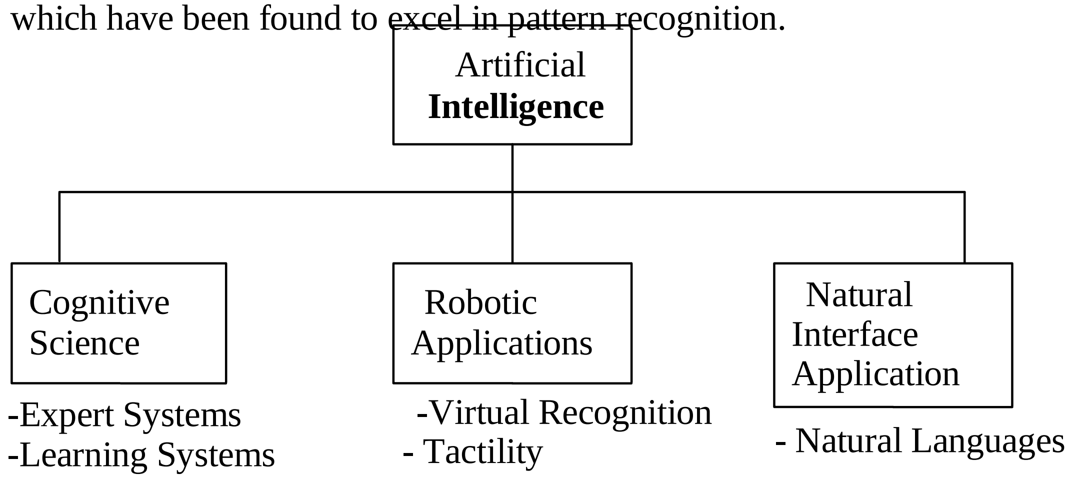


Figure 2: The major Application Areas of Artificial Intelligence

Computer Science

AI researchers have created many tools to solve the most difficult problems in computer science.Many of their inventions have been adopted by mainstream computer science and are no longer considered a part of AI. According to Russell & Norvig (2003, p. 15), all of the following were originally developed in AI laboratories:

- time sharing
- interactive interpreters
- graphical user interfaces and the computer mouse
- rapid development environments
- the linked list data type
- automatic storage management
- symbolic programming
- functional programming
- dynamic programming
- object-oriented programming

Business

Banks use artificial intelligence systems to organise operations, invest in stocks, and manage properties. In August 2001, robots beat humans in a simulated financial trading competition.
Financial institutions have long used artificial neural network systems to detect charges or claims outside of the norm, flagging these for human investigation.

Medical Clinics

A medical clinic can use artificial intelligence systems to organise bed schedules, make a staff rotation, and provide medical information.

They may also be used for medical diagnosis.

Artificial neural networks are used for medical diagnosis (such as in Concept Processing technology in EMR software), functioning as machine differential diagnosis.

Heavy Industry

Robots have become common in many industries. They are often given jobs that are considered dangerous to humans. Robots have proven effective in jobs that are very repetitive which may lead to mistakes or accidents due to a lapse in concentration and other jobs which humans may find degrading. General Motors uses around 16,000 robots for tasks such as painting, welding, and assembly. Japan is the leader in using and producing robots in the world. In 1995, 700,000 robots were in use worldwide; over 500,000 of which were from Japan.

Toys and Games

The 1990s saw some of the first attempts to mass-produce domestically aimed types of basic Artificial Intelligence for education, or leisure. This prospered greatly with the Digital Revolution, and helped introduce people, especially children, to a life of dealing with various types of AI, specifically in the form of Tamagotchis and Giga Pets, the Internet (example: basic search engine interfaces are one simple form), and the first widely released robot, Furby. A mere year later an improved type of domestic robot was released in the form of Aibo, a robotic dog with intelligent features and autonomy.

Aviation

The Air Operations Division AOD, uses for the rul based expert systems. The AOD has use for artificial intelligenc for surrogate operators for combat and training simulators, mission management aids, support systems for tactical decision making, and post processing of the simulator data into symbolic summaries.
The use of artificial intelligence in simulators is proving to be very useful for the AOD. Airplane simulators are using artificial intelligence in order to process the data taken from simulated flights. Other than simulated flying, there is also simulated aircraft warfare. The computers are able to come up with the best success scenarios in these situations. The computers can also create strategies based on the placement, size, speed, and strength of the forces and counter forces. Pilots may be given assistance in the air curing combat by computers. The artificial intelligent programs can sort the information and provide the pilot with the best possible manoeuvres, not to mention getting rid of certain manoeuvres that would be impossible for a sentient being to perform.
Multiple aircraft are needed to get good approximations for some calculations so computer simulated pilots are used to gather data. These computer simulated pilots are also used to train future air traffic controllers.

The system used by the AOD in order to measure performance was the Interactive Fault Diagnosis and Isolation System, or IFDIS. It is a rule based expert system put together by collecting information from TF-30 documents and the expert advice from mechanics that work on the TF-30. This system was designed to be used for the development of the TF-30 for the RAAF F-111C. The performance system was also used to replace specialised workers. The system allowed the regular workers to communicate with the system and avoid mistakes, miscalculations, or having to speak to one of the specialised workers.

The AOD also uses artificial intelligence in speech recognition software. The air traffic controllers are giving directions to the artificial pilots and the AOD wants the pilots to respond to the ATC's with simple responses. The programs that incorporate the speech software must be trained, which means they use neural networks. The program used, the Verbex 7000, is still a very early program that has plenty of room for improvement. The improvements are imperative because ATCs use very specific dialogue and the software needs to be able to communicate correctly and promptly every time.

The Artificial Intelligence supported Design of Aircraft, or AIDA, is used to help designers in the process of creating conceptual designs of aircraft. This program allows the designers to focus more on the design itself and less on the design process. The software also allows the user to focus less on the software tools. The AIDA uses rule based systems to compute its data. This is a diagram of the arrangement of the AIDA modules. Although simple, the program is proving effective.

In 2003, NASA's Dryden Flight Research Center, and many other companies, created software that could enable a damaged aircraft to continue flight until a safe landing zone can be reached. The Intelligent Flight Control System was tested on an F-15, which was heavily modified by NASA. The software compensates for all the damaged components by relying on the undamaged components. The neural network used in the software proved to be effective and marked a triumph for artificial intelligence.

The Integrated Vehicle Health Management system, also used by NASA, on board an aircraft must process and interpret data taken from the various sensors on the aircraft. The system needs to be able to determine the structural integrity of the aircraft. The system also needs to implement protocols in case of any damage tothe vehicle.

4.0 CONCLUSION

Artificial Intelligence (AI) is the concept that computers can be programmed to carry out certain logical processes which are normally associated with human intelligence rather than with machines, such as learning, adaptation and self-correction. The application of AI is numerous and cuts across all works and life. It is one other good thing that has happened to human development. Many businesses have adopted the concept to enhance their productivity and efficiency.

5.0 SUMMARY

- Organisations are using artificial intelligence (AI) to capture individual and collective knowledge and to codify and internalise their knowledge base. AI can be defined as the effort to develop computer-based system (both hardware and software) that behaves as humans.
- Successful AI systems are based on human expertise, knowledge and selected reasoning patterns, but they do not exhibit the intelligence of human beings.
- Although AI applications are much more limited than human intelligence, they are of great interest to business.
- In the course of 50 years of research, AI has developed a large number of tools to solve the most difficult problems in computer science.
- AI researchers have developed several specialised languages for AI research.
- In 1950, Alan Turing proposed a general procedure to test the intelligence of an agent now known as the Turing test. This procedure allows almost all the major problems of artificial intelligence to be tested.
- There are a number of competitions and prizes to promote research in artificial intelligence.
- Artificial intelligence has been used in a wide range of fields including medical diagnosis, stock trading, robot control, law, scientific discovery and toys.

6.0 TUTOR-MARKED ASSIGNMENT

1. Give 5 reasons why business community favour Artificial Intelligence.
2. Mention 5 areas in computer science that Artificial Intelligence research has helped to create tools for problem solving.

7.0 REFERENCES/FURTHER READING

Russell, Stuart J. & Norvig, Peter (2003). Artificial Intelligence: A *Modern Approach (2nd ed.). Upper Saddle River, NJ: Prentice*

Hall, ISBN 0-13-790395-2.

Nordlander, Tomas Eric (2001). AI Surveying: Artificial Intelligence in *Business. (MS Thesis), De Montfort University*.

Norton, Peter (1995). Introduction to Computers. Macmillan/McGraw-Hill.