City University
London

Conceptual Modeling and Systems Theory with an application using Real Options Analysis

BY

KONSTANTINA MILIOTI

THESIS SUBMITTED FOR THE
AWARD OF THE DEGREE OF DOCTOR
OF PHILOSOPHY
IN

SYSTEMS & MODELING

CENTRE FOR SYSTEMS RESEARCH
DEPARTMENT OF ELECTRICAL, ELECTRONIC &
INFORMATION ENGINEERING
CITY UNIVERSITY
LONDON EC1V 0HB

April 2007

“Conceptual Modeling and Systems Theory with an Application using Real Options Analysis”

Στον πατέρα μου, την μητέρα μου και την Φωτεινή…

Στους Λαιστρυγόνες, τους Κύκλωπες, και τον άγριο Ποσειδώνα…

Σὰ βγεῖς στὸν πηγαιμὸ γιὰ τὴν Ἰθάκη,
νὰ εὔχεσαι νἄναι μακρὺς ὁ δρόμος,
γεμάτος περιπέτειες, γεμάτος γνώσεις.
Τοὺς Λαιστρυγόνας καὶ τοὺς Κύκλωπας,
τὸν θυμωμένο Ποσειδῶνα μὴ φοβᾶσαι....

...Πάντα στὸ νοῦ σου νἄχεις τὴν Ἰθάκη....

...Ἡ Ἰθάκη σ' ἔδωσε τ' ὡραῖο ταξεῖδι.
Χωρὶς αὐτὴν δὲν θἄβγαινες στὸν δρόμο.
Ἀλλὰ δὲν ἔχει νὰ σὲ δώσει πιά.

Κι ἂν πτωχικὴ τὴν βρεῖς, ἡ Ἰθάκη δὲν σὲ γέλασε.
Ἔτσι σοφὸς ποὺ ἔγινες, μὲ τόση πεῖρα,
ἤδη θὰ τὸ κατάλαβες ἡ Ἰθάκες τί σημαίνουν.

Κωνσταντίνος Π. Καβάφης (1863-1933)

Table of Contents

Volume 1 of 2

List of Figures

List of tables

Acknowledgements

This single page seems painfully inadequate for me to thank all the people that helped me and supported me through the years of my research.

I owe an immeasurable debt to my supervisor Prof. Nicos Karcanias, who has guided me through "dim lit paths". His valuable comments and constructive criticism, his continuous support and understanding, became my cornerstone the last four years. Thank you for stressing out that "Greece will always be there, waiting for you…just enjoy the journey..." A true friend and family.

Words cannot express my gratitude to my family who gave me the opportunity to pursue my dreams in the UK. Your unconditional love and support is my safety blanket. Fotini thank you for the long, crazy conversations that shifted the weight from my worries. Mum you are my rock…still…forever. Dad you are my shining star.

A special thank you goes to my second supervisor, Prof. David Stupples, who introduced me to a new area for me, that of finance.

Things would be extremely difficult for me if I did not have people around me to keep me happily "dazed and confused"…and these are my friends and fellow colleagues. Stefane thank you for keeping me sane when things seemed out of this world. I would also like to thank my colleagues from the Control, Information and Biomedical and Mechanical Engineering research centers…for going through the same phase and *actually* understanding. A special thank you goes to Maria, Vangelis, Khalid, Leyla, Stamatina, Nick and Lefteris for their hospitality the last few months. Mr. G… thank you for "waking me up".

Angry Poseidona…I do not fear you!

Declaration

The University Librarian of the City University may allow this thesis to be copied in whole or in part without any reference to the author. This permission covers only single copies, made for study purposes, subject to normal conditions of acknowledgements.

Abstract

This thesis introduces an abstract conceptualization of the notion of the system and examines its relations to the various issues relating to the problem of conceptual modeling and its links to the development of formal models and control/information architectures. The overall effort is placed on creating a basis that will help to move from a purely conceptual level of describing a system to a more formal one, enabling decision making and driving the analysis away from experience and intuition.

The major drive is the development of a basis of general system theory in a way that expresses all aspects which are relevant to conceptualisation, modeling and control of processes, but are so general to provide a meaningful framework for general processes. For that purpose, a general systemic framework, suitable for discussing a number of important issues arising in the analysis and design of processes, is introduced.

A particular case study-the future projection of prices of electricity, generated by gas and nuclear in the UK- has been used in this thesis to enable the conceptualisation process, as well as initiate the decision making process with the use of Real Options analysis. Through this application, a spherical understanding of the requirements needed for the purpose of modeling analysis and decision making was acquired.

Section 1

Modelling and Systems:

Background Theory

Chapter 1

Introduction

The study of engineering and technological systems has benefited immensely by the use of a system's framework, which, when supported by formal modelling tools, enables the emergence of analysis and design methodologies. System's concepts have been employed at the 'design of the system' stage, as well as when formulating control decision- making strategies for the final system. So far, system theory has evolved based on paradigms linked to physical or engineering systems, and on simple classes of mathematical models that permit the development of formal methodologies for analysis and synthesis-design. Major challenges emerge nowadays in problem areas associated with new paradigms, where the rich modelling tools of the engineering domain do not always apply. Hence, developing a successful conceptual framework and efficient methods for these alternative areas is an important challenge. An effort to develop a formal approach requires the definition of a generic system framework, which is of the conceptual type, and selection / development of relevant modelling tools that will support analysis and design. It is a central objective of this thesis to clarify and unify the alternative notions of a "conceptual model" and thus create a basis for developing formal methods, control/information architectures and decision making strategies in a systematic way.

The central feature of interdisciplinary work is the effort to bridge diverse areas and disciplines and eventually lead to some form of integration. Integration is an area with its own dynamics and crucial for its development is the presence of a common language, concepts, and methodologies which may act as facilitators for integration. This task is of holistic nature and this makes the need for development of a general systems framework of paramount importance. Such a framework is essential for modelling and the shaping of relevant analysis and design tools.

Understanding systems and the basic problems associated with them, gives to the scientist a platform to stand on and reflect the generalized framework to create a model for the system. Understanding the system is equivalent to the development of a model which is a problem that up to now has attracted a lot of effort, but still lacks an integrated methodology and approach that goes beyond the specifics of the considered application domain. Model construction, model reduction, data mining and many other model related problems are still in the focus of many areas that introduce considerable challenges in the study of solutions of complex problems.

The model is the interpretation of the system as seen by the modeller-observer. This action of trying to observe systems, comprehend their behaviour and then describe and communicate this knowledge expresses the art of modelling. Once a model is developed, one may use the model to explore the properties of the system. This kind of analysis drives the modelling process, and thus, improves it. Developing models is a wide and complicated area of research. To handle systems in a way different than that of the "soft" systems approach, we need a proper development of the process of conceptualization and its links to the successive formal modeling stages and this requires following:

- An understanding of the abstract system
- The conceptualisation itself, and
- The development of decision making based on the structure given from the two previous stages.

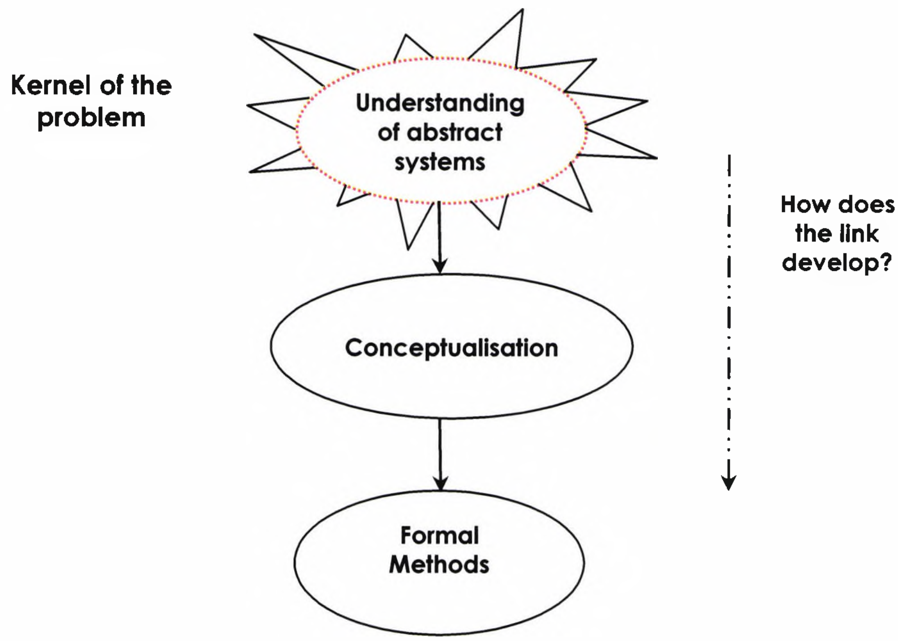


Figure 1.1: Development of the links of the stages in the modelling process.

This thesis aims to demonstrate that these stages are strongly interrelated and for the development of formal modelling and decision making methods, the proper way is by developing a strong understanding of the first two stages. This thesis focuses on the first two stages of this approach and examines key aspects by:

- Introducing a general systemic framework, motivated by our understanding of "hard" systems, that enables the extension of standard system notion to the more general set up of non-engineering applications.
- Formalizing the process of conceptualisation and integrating it into the overall modelling process.
- Using the abstract system definition and the notions of conceptual models to develop control and measurement architectures for the problem of "integrated operations".
- Using a case study to demonstrate the flow of the modelling process, and by emphasising the need of systems thinking and conceptualisation in decision making problems.

- Focussing on the open issue of the transition from the abstract and conceptual description to a more formal one and thus creating a natural path for moving from a purely conceptual level of describing a model to a more formal one and thus laying the foundations for their eventual integration.

The major drive in this thesis is the development of a basis of general system theory in a way that expresses all aspects which are relevant to conceptualisation, modeling and control of technological processes, but are so general to provide a meaningful framework for general processes. A significant contribution is the understanding of the role of the general system in the specification of conceptual models and subsequently the transition towards the development of formal models. For that reason the paradigm of integration of operations in an industrial enterprise will be used. With the help of that paradigm, we will demonstrate how the area of overall process of operations involves processes of different nature expressing functionalities of the problem, and therefore the need to develop generic features of the control and information architecture at such early stages is very important.

In more detail, a general systemic framework, suitable for discussing a number of important issues arising in the analysis and design of processes, is introduced. It will be described as a means of integrating diverse areas where system concepts are used. A general approach towards the definition of a system will be given. The work here uses the existing methodologies, but aims at redefining notions, concepts and introduces new ones reflecting the needs of the new paradigms. The development of a detailed and relevant mathematical framework that bridges the abstract with the concrete descriptions of specific domains is a future task that can be undertaken when the relevant concepts have been appropriately defined. This future challenge is closely related with the development of formal methods towards the end of the modelling process and is something that is not in the core of this thesis.

In addition, the basic notions of a model and the modelling process will be described, along with preliminary descriptions of the types of models, as they have been defined for the purpose of this thesis. Several issues regarding the modelling process, such as the role of the modeller and data will be identified. Furthermore, a class of modelling questions will be shaped and we will provide a range of problems that are critical for the overall modelling exercise and require special research effort for their study.

Special attention will be given in the preliminary analysis, regarding the specifications, the requirements, the relevant domains, and the assumptions related to the applications. This stage of analytical argumentation that can be thought as an early study of the system (problem) will be referred here as conceptualisation. Conceptualisation of the system is the first stage of the modelling process that is based on the "general knowledge" about the system and relies on information from the past. It produces a structure, a framework that has to be substantiated and populated by "current" and "specific" knowledge. This is where the interaction of conceptualisation and the "measurement process" begins. Conceptualisation provides the fundamentals of the structure of the model, whereas measurement comes to provide the means to identify details, fix parameters and validate assumptions. Problem articulation has been seen and developed before by Soft systems approach but the Conceptualisation was never seen as an integral part of the modelling process. The abstract system is central for the development of a proper costruction of the conceptual model and provides the key for bridging conceptualization and the formal methods. This thesis is emphasizing the indivisible nature, and critical role of conceptualisation in the modelling process.

Another key issue in this thesis is the effort to bridge conceptualisation with something more concrete, which is the development of control/information architectures on the basis of conceptualization. In our effort to understand the system, we begin to collect data that may lead to useful observations, will enhance our understanding of the model and develop decision making processes on the basis of conceptualization and partial models describing aspects of the system behaviour. This thesis will use a case study to demonstrate how this bridging can be effective, by first defining a problem and then using its conceptual model try to find the best suited solution, according to the identified specifications and constraints. We will look into the electricity production problem in the UK, and more specifically look into the possibility of substituting the percentage given by natural gas, with another energy source. This necessity for decision making will prompt the discussion of Real Options analysis, and more specifically the Binomial Model, a tool usually used for economic forecasting with the advantage of providing options for the user.

In the following we give a brief description of the layout of this thesis. Chapter 2 of this thesis provides a new approach to define abstract systems using the experiences from the well understood engineering problems, but abstracted to a level that is suitable for general processes. A general approach towards the definition of a system will be given which embodies experience from the different areas of engineering type of applications and lays down the fundamentals for the model independent definition of fundamental system concepts. The characterization of these notions requires the specific context of the formal model and thus goes beyond the scope of the thesis. Then a classification of the systems on account to their nature, and what they try to incorporate in their description will be given. This journey will emphasise the need for a foundation of that theory at a level suitable for all disciplines that use it.

In chapter 3 our task is to reconsider existing concepts and notions from the general Systems area, detach them from the influences of specific paradigms and generalise them appropriately to make them relevant for the new challenges. This is a first attempt to identify the areas and issues where abstract system theory requires extensions and modifications in order to cope with the challenges of the new paradigms, such as that of Business Processes (BP), Life-cycle issues in Engineering Problems (LCEP), Data Structures (DS) etc.

Chapter 4 will try to give a structure to the process of modelling and the "pre-modelling" stages, with an open mind, but always treasuring the central role of the idea of the system. A discussion regarding scientific methods is given and then an introduction to general concepts and issues concerning modelling and how modelling relates to systems. The importance of the observer in the modelling process and his or her irreplaceability is emphasised. The issue of mapping the system to the model will emerge, along with how data mining and knowledge management help in that process and the further stages of modelling. A review of some soft approaches to problem solving will take place and finally the attention will be focused on the conceptual modelling; the definition and its role in the subsequent stages of modelling.

The main purpose of Chapter 5 is to identify the different stages in modelling, from early design to formal representation, giving special attention to the most critical phase of the modelling process, the conceptualisation. At present it appears that the opportunity to gain clarity within the design process is being wasted, with the process of design being generally poorly understood, which in turn leads to designers having no real structure or common focus [Mac et al].

In Chapter 6 a discussion regarding the use of general conceptualization and modelling in the development of control structures and architectures is considered. We used as a driving paradigm that of an industrial enterprise, to demonstrate the integration of overall operations. The issue of development of the generic features of the control and information architecture at such early stages is very important and expresses a new step in our effort to design the overall system that is part of the overall process of *conceptual design* of the system. This discussion will provide an introduction to the main issues related to the subsequent stages of conceptual modeling which is the conceptual design process.

Chapter 7 will examine the setup for decision making, using primarily economic and financial tools. Special attention will be given to economic forecasting tools as a means to provide insight and elucidate certain grey areas where decisions are prominent. An overiew of some of the main, more efficient and widely used economic tools, will be examined. Real options analysis is the chosen method that will be used for the case study, and that choice is justified, giving a point by point description of the approach, how and when it is used, what are its strenghts and how it is compared to other leading tools.

Chapter 8 will introduce the problem area and define the problem, which will be the case study for this thesis. In this section the problem of electricity generation in the UK is considered in order to demonstrate the development of a conceptual model and the use of

Real Option Analysis for decision making. The conceptualisation of that specific problem will be discussed and the focus will be put on what kind of decisions need to be taken and how a decision making mechanism can help to show a way forward.

In chapter 9 the use of ROA will be demonstrated on the specific problem. The analysis will take place using the binomial lattice approach, to compare the two options that have been identified in the previous chapter. The comparison of the two evaluation lattices is a more graphical way to see where one of the two options becomes more competitive than the other, while the option valuation lattice pinpoints the exact point in time where the successor energy source can take effect. What is needed from this analysis is the planning horizon for the energy source substitution.

In Chapter 10 a general overview on the subject of uncertainty in modelling will take place. A discussion will be made on how the *model* uncertainty, incorporates *real world* uncertainty, as well as uncertainty as viewed by the *decision maker-modeller*. Furthermore the outcome of the formal model, realized in the previous chapter, will be discussed and measured against uncertainties and risks. An overview of possible risks and the uncertainties associated with the specific problem will take place. Furthermore, various issues regarding the use of Real Options Analysis for the specific problem will be identified in an effort to conclude the discussion of the case study. The various issues regarding the outcome of the model, the timing of the proposed solution, reaction of the public, etc, will be discussed as well. Finally, Chapter 11 provides an overview of the research and specifies issues for further research.

Chapter 2

Background to Systems

The notion of the "system" is something that has been an engaging issue for a long time and for many and diverse areas. This notion has been used on various occasions and time periods from different areas, to describe ideas or things that bond together to create a whole, successfully. It is something that can be seen used in almost every part of life, and although this use can only emphasize its importance, it has become the axis of separation between the various areas that use that notion. This chapter provides a background information regarding systems. A general approach towards the definition of a system will be given. Furthermore, their use from the different areas and the various issues emerging from the difference in divergence will be identified; issues emerging from a system's different viewpoints, from different priorities when identifying the system, from the different terminology used to describe it and the difference its observer makes. Then a classification of the systems on account to their nature, and what they try to incorporate in their description will be given. A small historical retrospection of the notion of system will be introduced, showing how the various ideas and views were shaped. This journey will emphasise the need for a foundation of that theory at a level suitable for all disciplines that use it.

2.1 The notion of the system

At this point the word "system" needs to be defined and an emphasis needs to be given on the general properties-assumptions of its nature. A system is not something presented to the observer; it is something to be recognized by him. Often the word does not refer to physical entities, but rather to conceptual ones a way of organizing our thoughts about the real world. A system is usually defined, in a physical way, as the grouping of several independent but interrelated objects (parts), so as the product of the grouping will be a functioning "entity" (whole). The system is surrounded by the system environment, with which it is always in communication, through a system boundary. The definition of the system involves in an active way the specification of the boundary that encloses the activities under consideration. The external world to the system may be known or unknown and always communicates with the system. A system is a distinct entity, a unified whole with its own identity. Another similar, often used common sense definition is the following: 'A system is a set of interacting units or elements that form an integrated whole, performing or if designed, intended to perform some action'. Reduced to everyday language it can be expressed as any structure that exhibits *order*, *pattern* and *purpose* [Skytt].

Any system corresponds to a reality; it is part of the world independently of the fact that the world could be either the physical or the non-physical. The parts that make up the system can be called objects. An *object* is a general unit (abstract, or physical) defined in terms of its attributes that is its qualities, and the possible relations between them. Objects at this stage can be assumed to be the most primitive form of elements, allowing the observer this way to be free to express them at different levels of detail. To be more precise, take the example of a human organism; the system is the human body, so one observer sees the objects of the system as the organs of the human and another observer sees the cells as the objects of the system. Thus the elements can be as big or as small depending on the observer's aim, the situation, the purpose of the observation, etc. The parts that make up the system are linked with each other and thus share some kind of relationship. Those relationships are not necessarily fixed or always predictable since the interactions with the environment can change the objects and the relationships between them. Most systems are dynamic and can never be thought as fixed.

Previously characteristics such as the order, the pattern and the purpose of the system were emphasized. Further to those characteristics of any system, is the *behaviour* of the system. The behaviour of the system can be defined as the aggregate manifestation of time event driven evaluation of attributes associated with the system. The most fundamental concept related to a system is the realization that the system can be thought as a box that tries to transform its input map to an output map. This means that there are sets of activities contained in the system, necessary to transform some inputs into some outputs. Thus as an example take a designed physical system such as a chemical process; the input could be raw material of some kind and the output a range of products.

The reason why we need to understand the behaviour of systems comes as a twofold; understanding the way they work and thus extracting relevant and useful information about their existence and how this information can be used in a similar situation and secondly use of the system for a given purpose. Thus it can be said that systems can have either of an educational character for the observer or of a performing character. The latest implies that if the system needs to be used to achieve a particular objective, some measure of the degree of that achievement must be derived. This is defined as measure of performance. That measure then is compared to the desired outcome and decisions are taken to include further activities in the system so as to improve the degree of achievement. The information collected according to that measure will be used by some decision making procedure to take control action through control mechanisms. Thus, for example, if the system objective is defined as the satisfaction of a perceived market need, the measure of the performance must be related to how well the particular sector of the market is satisfied. Based upon the information collected, action can be taken to improve the product or improve the market definition.

The idea of the system puts emphasis on the connections between components and seeing a degree of organisation. It stresses the role of each component in contributing to the behaviour of the system. If a component was not present in the system, the behaviour of the system would be different. Every system does something and more particularly every subsystem does something too. These subsystems might be independent entities and have a purpose, goal and behaviour of their own. While, though, under the power of the system, their behaviour is closely dependent by the interacting parts, trying to conform to the rules of the system. What is understood, therefore, is that the system is not the aggregation of its parts, the whole is more than the sum of the parts; the behaviour of the system is different than the behaviour of the individual parts, but given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole. The organisation of the various different parts is producing the behaviour of the system.

The concept of 'system' is universal; anything can be regarded as a system and represented by a coherent collection of its *attributes*, or as a collection of its related *parts*. Systems may be as enormous as galaxies, as miniscule as atoms, as simple as a brick wall, or as complex as a human cell [Kap].

System definition is the most critical phase in any project. The functionality and performance of a system are defined in this stage, which becomes the basis for determining the system's specifications and capabilities. The axiomatic definition of the system is also required later on in the process of modeling as an integral part of the product model.

The system of interest may exist, irrespectively of the observer, just as the solar system exists. A system may be planned and built purposefully and constructed deliberately and evolve gradually, like any man-made system (e.g. the road system of a continent), or natural systems. It may evolve naturally over a long time period, like the nervous system of a species, the food chain in a forest or the human society.

Other well known features of systems include the following [Skytt]:

- The universe is an organisation of systems; that is simple systems are synthesized into more complex systems from subatomic particles to civilizations.

- All systems, or forms of organisation, have some characteristics in common and it is assumed that statements concerning these characteristics are universally applicable generalizations.
- Every system has a set of boundaries that indicates some degree of differentiation between what is included and excluded in the system.
- Everything that exists, whether formal, existential or psychological, is an organized system of energy, matter and information.
- Most human constructed systems, if they are to attain their goal, must transform inputs to outputs.
- Systems are generally complex wholes made up of smaller subsystems. This nesting of systems within other systems is what is implied by organisation.

Systems are linked to specific "activities" which are performed. The system as previously described is a notion widely accepetd and used in various areas to describe different ideas. Following are some examples of the use of the notion of the system, as they appear in different domains.

- An engineering example of a system is often a circuit.
- In the natural world, we say that there are systems. For example, the solar system of nine planets orbiting the sun. In the human body, we refer to such systems as the nervous system, the circulatory system, the digestive system, the reproductive system, and the respiratory system. The entire body is also referred to as a *system* in terms of physiology.
- Laws are a *system* which governs human social behavior. Grammar is a *system* which governs language usage.
- In computer science and information science, a system could also be a method or an algorithm. A computer system consists of a set of hardware and software which performs algorithmic processes. This still fits with the definition of components which are connected together (in this case in order to facilitate the flow of information). System can also be used referring to a framework, be it software or hardware, designed to allow software to run.

- In operations research and organizational development (OD), organizations are viewed as human systems of interacting components such as sub-systems, processes and organizational structures.

It is clear that the notion of systems emerges in many and diverse areas and it is a fundamental notion appearing in every possible aspect of life. The terminology in each area might differ but what is meant by the outcome of the collection of parts to create a whole remains the same throughout.

2.2 Basic problems with system notions

Each science pursues its studies from a certain point of view. Systems science has as well its specific point of view and it is to understand systems, their structures and properties, their environment, the interactions between them and the environment. Political science concentrates on the society's political and administrative organization. Business economics is concerned with the commercial organization, geography with the physical structure and philosophy with the pattern of thought, views of life and ideologies. Engineering is concerned with the practical application of scientific knowledge in the design and synthesis of structures. In thesis the distinctive effort will be the use of tried and tested concepts and approaches developed within the engineering field, to more general problems not identified with engineering. The difference in the perspective used by each science to describe or define a system creates an array of problems.

2.2.1 General systems issues

The concept of a system is the idea of a whole entity, which under a range of conditions maintains its identity, provides a way of viewing and interpreting the universe as an organisation of interconnected and interrelated wholes (based on the admission that a system is a set of objects-wholes). A system provides the generic structure of the entity each time considered. The behaviour of systems is not always exemplary. Economic systems are subject to inflation and depreciation; biological systems are subject to decay and disease; ecological systems are subject to pollution. What can be realised is that all kinds of systems are subject to external disturbances and do require care and the cost of this supervision is often an important factor in decision-making.

The concept of a system is very general. There are reasons that make difficult the study o certain systems. Some of those reasons are unfolded below:

- The number of attributes necessary to describe or characterise a system may not be too many.
- The large scale nature of the system, expressed in terms of dimensionality or multi-component nature.
- Not all attributes are necessarily observable. In fact, the structure or configuration of the system is rarely self-evident.
- In large, complex, systems there is a scope for many possible "configurations". Selection of one, linked to specific objectives, may have far-reaching repercussions.
- Systems do not necessarily have a fixed or static nature but they evolve in time.
- The system environment may inflict changes in the organisation and nature of the system. Such changes are not generally under the control of the observer and their influences as the system evolves in time are not apparent at the outset.
- Any system design must take into consideration the fact that future disturbances may arise which are not present in the existing system and the control system must itself evolve in order to respond effectively to the future disturbances.

Whether a given system is considered large or small essentially depends upon value judgement. What is considered complex from one point of view could be of simple structure from another point of view. As a rule, complexity in behaviour arises due to complexity in structure. Thus a useful indicator revealing the complexity of a system lies in the complexity of the structure of the system. The general problem of operating a complex system with limited resources and limited amount of time for observation, data processing and implementation of control, generates new kinds of questions that have not yet been precisely formulated and certainly not resolved. [Vem]

2.2.1.1 System complexity

The original Latin word *complexus*, which signifies "twisted together" may be interpreted in the following way: in order to have a complex you need two or more components, which are joined in such a way that it is difficult to separate them. Similarly, the Oxford Dictionary defines something as "complex" if it is "made of (usually several) closely connected parts". Intuitively then, a system would be more complex if more parts could be distinguished, and if more connections between them existed. More parts to be represented means more extensive models, which require more time to be searched or computed. Since the components of a complex cannot be separated without destroying it, the method of analysis or decomposition into independent modules cannot be used to develop or simplify such models. This implies that complex entities will be difficult to model, that eventual models will be difficult to use for prediction or control, and that problems will be difficult to solve. This accounts for the connotation of *difficult*, which the word "complex" has received in later periods [Heyl].

Complex Systems is a term that emerges in many disciplines and domains and has many interpretations, implications and problems associated with it. How complex or simple a structure is depends critically upon the way in which it is described. Most of the complex structures in the world are enormously redundant, and this redundancy can be used to simplify their description. But to use it there is an apparent need to find the right representation [Simon].

It is argued that variation and selection automatically produce differentiation (variety) and integration (dependency), for living as well as non-living systems. Structural complexity is produced by spatial differentiation and the selection of fit linkages between components. Functional complexity follows from the need to increase the variety of actions in order to cope with more diverse environmental perturbations, and the need to integrate actions into higher-order complexes in order to minimize the difficulty of decision-making. Both processes produce a hierarchy of nested super systems or metasystems, and tend to be self-reinforcing. Though simplicity is a selective factor, it does not tend to arrest or reverse overall complexity.

The various parts of any system have to be in balance. Some design procedure is essential in order to ensure that an engine is in balance with the vehicle it drives, or that the heart is evolved with sufficient capacity to pump the blood around the organism. It ensures that the capabilities of the systems which regulate are sufficient to deal with the complexity of the problems which they have to deal with. Variety is the tool that helps dealing with these issues.

The definition of variety is straightforward. It is the number of states in which a system can exist. Variety is used to measure complexity. If a system is complicated is has a large variety. If it's simple the variety is low. The variety of a regulator must be at least as large as that of the system it regulates; this statement is usually referred to as Ashby's *Law of Requisite Variety*, as it says that the regulator must have enough (requisite) variety to adequately do its job. This can be expressed in a number of ways. It's often expressed as “only variety can absorb variety”. In organisational terms it means that the capabilities of the regulators have to balance the complexity of the situation they are charged with regulating. This regulation could be the traditional view of management, or it could be the regulation of a jazz band where the rules by which the music unfolds are under the control of the people making it, or the regulation of body temperature. In all these cases Ashby's Law is applicable. If the systems which regulate don't have enough (or *requisite*) variety to match the complexity of the regulated, then regulation will fail. The system will be out of control [VSM].

Complexity is a common characteristic of many technological, production, environmental, societal, and financial, business problems, etc. Below, some classes of the categories of complex nature, which is manifested in many different ways, are given:

- Lack of knowledge or difficulties in characterising the behaviour of the basic process, or sub processes (*Unit Behavioural Complexity*).
- Complexity of computational engine associated with a sub process (*Computational Complexity*).

- Difficulties in characterising the interconnection topology of sub processes and/or variability, uncertainty of this topology during the system lifecycle (*Interconnection Topology Complexity*).
- Large scale dimensionality and possibly multicomponent character that impacts on methodologies and computations (*Large Scale – Multicomponent Complexity*)
- Heterogeneous nature of sub processes, which in a given interconnection topology, results in hybrid forms of overall behaviour (*Hybrid Behavioural Complexity*).
- Organisational alternatives for the functioning, information and decision making (control) structures in respond to goals and operational requirements (*Organisational Complexity*).
- Variability and/or uncertainty on the system's environment during the lifecycle (changing goals, requirements, disturbances, structural changes) which require flexibility in organisation and operability (*Lifecycle Complexity*).

The above demonstrate that the area of complex systems is multidimensional and has a multidisciplinary character.

2.2.2 A system and its viewpoints

The diversification of the system notion is evident in every aspect of all possible areas. It is evident that each discipline will identify differently a system, describe it with different terminology, and use the system to satisfy different purposes. What is not obvious though is, that every system has different viewpoints; they are all true, valid and correct, but their relevance depends on the area, the purpose, the observer and basically the angle from which the system is observed.

It is important to understand how the existence of different system viewpoints can be realized, with the help of an example. Imagine the situation when given a specific "entity" there is the need to identify a system, within that entity; given that the entity is a chair, how do the different systems emerge? In what context should the different systems emerge, how does every different area's purpose affect that purpose of identification, what attributes does the observer bring into the design and identification of the system? Going back to the example of the chair, whether the chair will be identified as a system comprised of different parts of wood, or a bigger system within which there are nested several other different systems (such as the molecules of the wood), it depends on the purpose the identification of the system, the different ideas and perspectives of the various disciplines, the observer, etc.

An example of how different areas are identifying systems having different viewpoints is the following; Cybernetics is a "theory of machines", but it treats, not things but *ways of behaving*. It does not ask "what *is* this thing?" but "*what does it do?*" Thus it is very interested in such a statement as "this variable is undergoing a simple harmonic oscillation", and is much less concerned with whether the variable is the position of a point on a wheel, or a potential in an electric circuit. It is thus essentially functional and behaviouristic. Cybernetics started by being closely associated in many ways with physics, but it depends in no essential way on the laws of physics or on the properties of matter. It deals with all forms of behaviour in so far as they are regular, or determinate, or reproducible. The materiality is irrelevant, and so is the holding or not of the ordinary laws of physics [Ashby].

In the following sections an overview of those emerging issues and problem areas will be discussed.

2.2.3 Differentiations of system notions due to different priorities and issues.

Each science perceives the notion of the system differently, and using the language appropriate to each different discipline, they try to develop a holistic way of describing it. The problems surface with the realisation that the term "holistic" is sufficiently predetermined by the boundaries set by each discipline. That is, each area will only try to explain the term system inside a relevant bounded area. This bounded area is as is, because each discipline will see and subsequently try to define specific angles of the same problem.

As previously discussed the system is not presented to the observer, it is something to be recognized by him/her. Because systems can be considered as mental constructions and not actually existing physical things it is relevant to ask, "What are the needs of the observer of the system that lead the observer to hypothesize this particular system and not some other system?" Depending on the field of interest, the observer will "see" the system to fit his specifications. Those specifications are not the same for each area. They are sculpted to take shape depending on the priorities and the issues of each area. Depending on the purpose of existence of the area, the system is identified and fitted to support that purpose. Organizational culture and behavior are among the most powerful, and ambiguous, forces within any discipline.

A simple example will clarify any blurred points. A computer from an engineering point of view is an integration of mechanical and electronic parts that interact and work together for purposeful outcome. A computer from the information science point of view is the software and the processes related to that. The same "world system" is interpreted differently and thus it can construct two different systems in different areas. The reason for that differentiation in this case is the difference of interests in those areas. An engineer will only want to look whether the mechanical parts are properly fitted, working and according to the specifications of the customer, with the final test that the computer can be switched on. An IT person would want, for example, to check the software inside the computer, the data related to that and the continuous flow of data in between parts. It is evident, therefore, how the conception of a system in each area affects the process of identifying the system, specifying the boundaries and generally pinpointing the specifics in the process of system design.

The issues arising that need description in each discipline, are different, and thus the hierarchisation of what needs to be described and what is of no importance when introducing the system notion in that area, is subsequently different in every area. It can be, therefore, seen that the purpose of the system and the issues each discipline wants to focus on is of prime importance and gives the perspective any area needs.

2.2.4 Differentiations of system notions due to different terminology.

The problems with system notions continue with the differentiation of the language used to describe them. As previously observed, each science perceives the notion of the system differently, and using the language appropriate to each different discipline, they try to explore a holistic way of thinking. Technical terminology that is specific to the linguistic community enables people to communicate complex concepts as efficiently as possible. Precise terms present information more effectively than everyday language, but they are only useful when their definitions are universally understood and the mapping between a term and the concept it represents is unique.

The issue with using different terminology is evident within the limits of every area, but becomes a problem when there is a need to compare, contrast or even try to exchange and join ideas about a subject in between areas. The language used by every area is most of the times unique and even when trying to describe the same system, the outcome of the description will evidently be unique in each area. This way though, the knowledge is contained within the bounded limits of each area, and every new idea, conceptualization or even experimentation for something new, remains "unfound" for the rest of the areas. The term "unfound" is in a fluid state, since the system could already be recognized by different areas, but the different terminology could provide a barrier of miscommunication between the areas for any exchange of ideas, thus every new aspect of the system will remain "unfound" by at least one of the communicating areas.

It is evident that there are times where the problem of system notions comes down to a difference of semantics. The problem of the *semantic matching* of the terms used to define the system and the specifics of the system, in between areas, is of immense importance not only for the knowledge that can be transferred from area to area, but for any kind of integration towards systems notions. Only if the meaning of the terms used is clear, is it possible to compare and contrast between disciplines. In other words, the problem is to make explicit the *intended models* of the vocabulary used to convey and request information, clarifying the assumptions implicit in the terms adopted for concepts, relations, attributes.

An example to illustrate the effect of the difference of semantics between areas is the following: "Consider a classical control system...To be in good control the system must form a feedback loop, so that disturbances and other external forces from "reality" (for example noise or other disturbances from other control systems) are counteracted by compensating actions so as to make the measured state as close as possible to the desired state, or at least stable within a region of its state space....The measurement function relates a state of the world to a particular representation (of the system); a decision function (by some *agent*) that representation to the choice of a particular action of another. That action has consequences for the state of the world...." [Jos]. This extract has been taken out of paper written in the area of Computer and Information Technology. Although, the main concept of what is a control system is the same as the one perceived by Control Engineering, the term "*agent*", is a new term for this area. The same action, in the Control engineering area, is taken by the *controller*. This example is just a simple demonstration of the differences in terminology and not in concepts, which only makes stronger the argument for a unifying framework of systems concept, for a more productive exchange of ideas between areas.

There is a need to try to study the interaction from multiple perspectives, holistically, to provide a general language with which to tie together various areas in interdisciplinary communication, to join together the many splintered disciplines with a "law of laws", applicable to them all and integrating all scientific knowledge [Skytt].

2.2.5 The observer of the system and the qualities brought by him

Another issue arising regarding system notions is the way the observer will see the system and try to model it afterwards. The notion of the observer as well as modelling will be thoroughly discussed in a following chapter. At this point an introductory discussion will take place.

Identifying a system closely depends as seen earlier on the context of the discipline the system will exist, the observer and the distance of the observer from the system. The observer is the person who will identify the system and try to model it. A system needs an observer with an interest and this interest will influence the viewpoint of the observer and even what he or she will choose to include in his or her 'system'. The observer is somehow concerned with how things *ought to be* in order to attain goals and to function. Understanding the system will be personal to whoever undertakes it; the observer is influenced by his/her own past experiences, carrying knowledge and obviously his or her own views and qualities. As it can be understood a mixture of the above, cannot be parted from the observer and it passed on his or her methodology of identifying a system and later on modelling it. The observer is always introducing his/her viewpoint in identifying and describing the system, its goals, its objectives, etc, but this is an integral part of the process.

Furthermore and independently of the qualities of the observer, identifying a system depends on the distance the observer has from the system, that is what aspect of the real world is needed to be captured and handled in such way as to compose a system relevant. Ultimately, this distance-and therefore the relevance-is set by the discipline and the purpose of identification and existence of the system.

For example, a typical building seen from the outside has a distinguishable structure on two or three levels: the building as a whole, the windows and doors, and perhaps the individual bricks. By noting that the building has distinguishable structure down to the level of bricks, implicitly the molecular, atomic and particle structure of those bricks is ignored, since it seems irrelevant to how the building is constructed or used. This is possible because the structure of the bricks is independent of the particular molecules out of which they are built: it does not really matter whether they are made out of concrete, clay, plaster or even plastic. On the other hand, in the example of the human body, the functioning of the cells critically depends on which molecular structures are present, and that is why it is much more difficult to ignore the molecular level when building a useful model of the body [Heyl].

The complexity of how the system is perceived, that is, the model, is determined much more by the limitations of its observer, than by its subject level. *Conceptual models* are aids to understanding and they can be no more complex than the limitation to understanding set by the human brain. However, once the validity of a conceptual model has been accepted and it is developed into a specific, practically useful form, it can become very complex.

Variety, as a measure of complexity, will depend upon what is distinguished by the observer, and in realistically complex systems determining what to distinguish is a far from trivial matter. What the observer does is picking up those distinctions which are somehow the most important, creating high-level classes of similar phenomena, and neglecting the differences which exist between the members of those classes. Depending on which distinctions the observer makes, he or she may see their variety and dependency (and thus the complexity of the model) to be larger or smaller, and this will also determine whether the complexity is seen to increase or decrease [Heyl].

The importance of the observer in the whole process is not something that can be bypassed. The observer is an inseparable part of the procedure of identifying and then modelling a system. This importance will be discussed further on.

2.3 Classification of systems

As previously observed, the word 'system' has many interpretations depending upon the context in which it is used. It can mean a procedure, a process, a network or a computer based data processing package. The definition can be refined by first of all deriving a classification in terms of types of systems and then developing a set of concepts appropriate to each type [Wil].

A very basic and simple but otherwise concrete classification of systems can be made by separating the physical entities from the ontological entities.

- By physical entities, it is meant all that is included in the physical world, that is either *naturally evolved* or *man made* (engineered). All physical entities are easily understood and can be easily depicted as something tangible. An example of a physical entity is the human body; naturally evolved, this system is made up of interconnected parts, the organs that act together as a whole to perform the actions meant to be performed by the human body, in ideal and healthy conditions. Another example of a physical system, this time of the engineered/man made type, is a car; this entity is composed by different mechanical parts and engineered in such a way as to conform to the purpose and role of the car.

- An ontological entity is a more abstract concept. It includes all non tangible systems. They can be *naturally evolved* or *man made*. The term ontological here is used in the context of pure semantics, as depicted in Webster dictionary; it means "a particular theory about the nature of being and the kinds of existents". In the philosophical tradition, ontology refers to a particular system of categories that seeks to study and structure a certain portion of reality. With that definition in mind, it is easier to understand that systems of that category are not as transparent as are the ones that fit into the category of the physical entities. An example of such an ontological system is any piece of software. The software can be thought of as a system consisting of various parts (algorithms) that work together towards the system's purpose. The software is not a physical entity but a conceptual one. This however does not negate its existence or its role as a system.

The important distinction between the physical and ontological systems is that the physical, either naturally evolved or engineered, are entities that are tangible in a way. The ontological, on the other hand, are mental constructions, as perceived and engineered by man on most occasions. There are occasions though that the ontological systems are naturally emerging and are not man made. An example of such an ontological system is any organisation. The notion organisation here is used in its abstract version denoting the managerial high level activities; that is how the different levels of an enterprise are communicating with each other, how the decisions are passed throughout the enterprise. The enterprise can be a firm, a software package or the human body. The concept of organization is naturally emerging and is an ontological system, with different parts trying to communicate.

Previously the terms *naturally evolved* and *man made,* were used. This observation about the source of creation for the system creates another distinction and thus another classification of systems; this classification can be made between the constructed or engineered systems with the ones that are naturally emerging. This classification comes as additional information about the distinction between physical and ontological systems.

A general observation can be made, at this point, regarding the physical and ontological/abstract systems. Physical systems, either naturally emerging or engineered are ruled by rules of physics. That is, for example if a system of a ball trying to balance itself on a free rotating beam, is considered, the rules of physics, gravity, rotation, friction, will drive the system through to a “balance”. That balance is something that in physical systems is evident that will happen and that systems are looking out for; that balance is equilibrium and all physical systems are driven towards that equilibrium. Ontological/abstract systems on the other hand, are ruled by the rules of logic and the notion of equilibrium is something not relevant to them. What is relevant is that any kind of continuity and consistency is driven by logic. To illustrate that argument, consider the following example. A software package is an example of abstract/ontological, engineered system. A software package has a purpose and a pattern. Since the pattern and order of the parts and thus the system is of an abstract nature rules of physics are impossible to be applied. What does make sense and is the main ingredient holding the system together in these cases is rules of logic; do the several different parts and subroutines of the package seem coherent and blend logically together so as to lead to the purpose of the software? Obviously this is what all systems need to be satisfying, but for physical systems the equilibrium is something inevitable unlike for ontological systems it is a way that needs to be lit so as to be followed.

The following diagram will summarise the classification of systems so far.

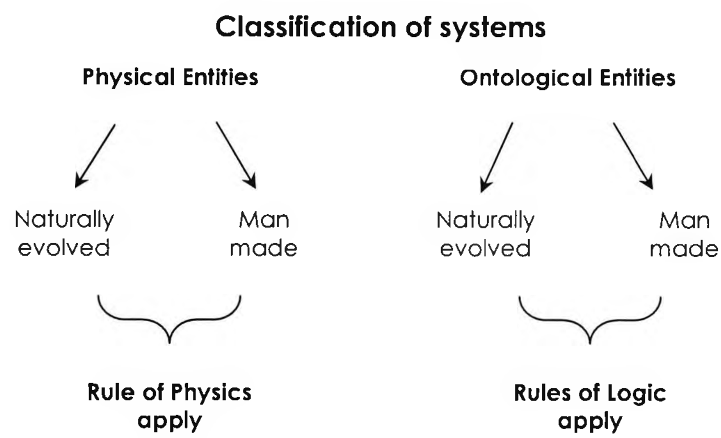


Figure 2.1: Classification of systems

Further to the discussion about classification of the systems comes a more *basic* categorization of systems in the following way:

- *Natural systems.* Natural systems are systems found in the real-physical world. Physical systems which make up the Universe in a hierarchy from subatomic systems through to biological systems, systems of ecology to galactic systems. Natural systems are a product of evolution.
- *Designed systems-Man made systems.* These can be both physical (bridges, machines) and abstract (mathematics, language, philosophy). They include engineering systems, computing/software systems, as previously discussed. Man made systems can easily interact with natural systems to create a bridge of relation. An example of such an interaction are, medical systems, where the physical world, and more precisely the biological systems, create a platform for designed systems to blossom.
- *Human activity systems-Man related systems.* Generally describing human beings undertaking purposeful activity such as man-machine systems, industrial activity, political systems, social and cultural systems, economical systems, psychological, management related systems. Most human activity will exist within a social system

where the elements will be human beings and the relationships will be interpersonal. Example of such systems would be a family, a community, an organisation etc. A general observation can be made here regarding a further classification of the human activity systems; it will be shown that the man related systems, have a hierarchisation, from a unit, to a grouping of units, to larger more complex systems, that is easier to pinpoint and understand, than in other classes.

- The unit. It can be said that the Psychological systems are ultimately revolved and related to the person which is the unit, and for this discipline cannot be broken down into smaller systems.
- The group. People usually work as part of a formal or informal group that influences the way in which they react to their work, help mould their opinions and goals and is in turn influenced by their relationship to it. Here the interpersonal behaviour is dominant.
    - Organisational systems are interested in the "group"; the unit here is lost as individuality and its effects can only be seen through the group. Superimposed upon the group is the control and influence exerted by the wider system or organisation to which the groups belong. The objectives of the organisation and of its component parts must be in satisfactory harmony with the relevant needs and desires of the system.
- Social/Economic/Political systems are far more complex and large than the ones described earlier. These systems are a grouping of the previously mentioned systems-a grouping of units and groups (organisational systems). In those systems, depending on the circumstances, the levels of nested units and groups are high.

In bibliography there exist different kinds of system classifications, but most of them follow the above idea of classification: separate natural systems from man made or man related. Differentiations of some kind obviously exist; for example, some researchers argue that social and cultural systems should be separated and treated as a class of their own, or that there exist several different sub-classes in each class, but this is only a matter of definition.

2.4 General Observations

At this point some general observations as well as a summary regarding the notions mentioned so far and the issues emerging from the various areas will be made:

- The notion of the system is multidimensional. It can be used to describe different parts of the "organisation". The term organisation here is used, in a very abstract form, denoting every kind and form of organisation from a discipline to a compilation of interdisciplinary areas. The problem of system integration in systems is examined here and it is considered nowadays as a major technological challenge; this, however, is perceived by different people in different areas from entirely different viewpoints. The dominant trend is to treat the problem as a software problem and neglect the multidisciplinary nature of the task and the very many different aspects of the problem, apart from software and data. The practical significance of integration has created some urgency in working out solutions to difficult problems and this has led to the development of interdisciplinary teams empowered with the task to create such solutions.

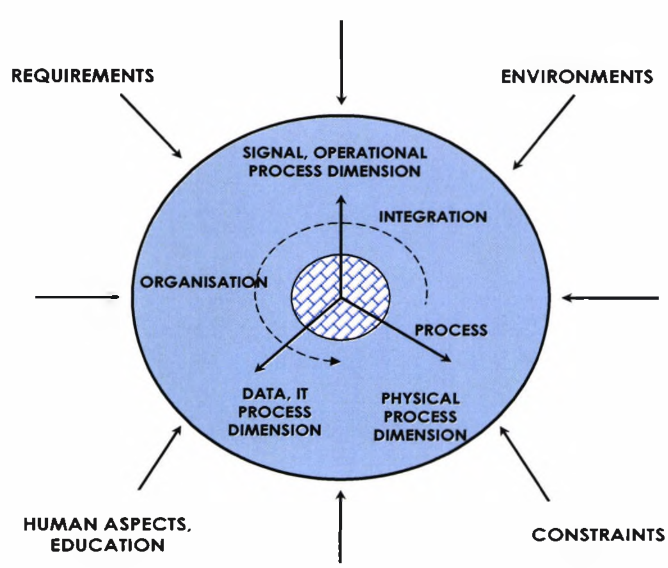


Figure 2.2: The multidimensional nature of systems [Karc, 1].

The figure above is an example of the multidimensionality of the systems. Obviously some aspects of the figure illustrated here are not relevant with every area (e.g. the IT process dimension is not relevant when the system is family), but all of the above aspects come together when there is a variety of interdisciplinary areas. The problem of system integration in the context of an "organisation" (the notion organisation here, as previously noted, is used in a very translucent way, meaning every kind and form of organisation from a system, to discipline, to a compilation of interdisciplinary areas) is a multidimensional problem with fundamental dimensions those of [Karc, 1]:

① Overall process operations,

② Overall System Design/Redesign, and

③ Information and Data flow, as seen on the diagram, and a number of other side activities.

- Every discipline has different *conceptualization* of what is a system; that is, definition of system, system boundaries, system environment, attributes, constraints and variables. That conceptualization will become the cornerstone for the definition of the system within the limits of that area. Each of the above three areas is of multidisciplinary nature and it is frequently considered by the respective groups as representing the entirety of the problem.
- The conceptualization of the system by each area is based upon the specific purpose each area gives to the system. That purpose is something evident within the limits of the discipline, since it is the driving force of existence for the area as well. The Physical Process Dimension deals with issues of design-redesign of the Engineering Process and here the issues are those related to integrated design. The Signals, Operations Dimension is concerned with the study of the different operations, functions based on the Physical Process and it is thus closely related to operations for production. In this area, signals, information extracted from the process are the fundamentals and the problem of integration is concerned with understanding the connectivities between the alternative operations, functionalities and having some means to regulate the overall behaviour.

- Any system is viewed for different considerations. Each time the system might change according to the purpose of the consideration.
- Every area finds different ways to define the system, according to the needs and the terminology available to that area. A system may have different definitions and related terminology ascribed to it, depending on the area of use. Both design and operations generate and rely on data and deploy software tools. Compatibility and consistency of the corresponding data structures and software tools expresses the problem of integration in this area and relies heavily on adopting common standards. The development of integration requires support from a number of other areas such as formation of multidisciplinary teams, relevant educational programs, etc. Understanding the relationships between problems on the same level, implies an ability to describe the links between models/problem solving techniques associated with the particular problems, as well as a capability to translate issues, requirements from one set up to another.
- The observer of the system contributes his or her own qualities when identifying a system. The outcome of the system is highly dependant from the observer.

Systems integration is a multi-task, multidisciplinary problem which is central in handling the major challenges in technology, economy, society, and environment. Bringing together people from different areas is clearly necessary, but not sufficient in producing solutions with acceptable performance. The key issue here is the lack of methodology that bridges disciplines and provides a framework for studying problems in the interface of particular tasks [Karc, 1].

2.5 Historical evolution of the notion of system

The Greek word *σύστημα* "system" comes from the ancient verb *'συνίστημι'* that means: Establish, combine, connect, form, construct something solid, solidify, and thicken. Maintain order, form party, engage, be related, be composed, happen, exist.

In ancient Greek written sources, the word 'system' appears frequently and the meaning of it varies according to the sources, such as geometry, music, and medicine. Let us cite some indicative instances:

- In geometry, mathematics, astronomy: Plato, *Epinomis*, 991e [Jow]: " ... *all geometric constructions, all systems of numbers, all duly constituted melodic progressions, the single ordered scheme of all celestial revolutions* ..."
- In politics: organised government, constitution, and polity (Plato, *Laws*, 686b)
- In music: harmony, system of musical intervals : Plato, *Philebus*, 17d [Jow]: "[Socrates speaks...] *when you have learned what sounds are high and what are low, and the number and nature of the intervals and their limits or proportions, and the **systems** compounded out of them ... you have technical skill*" [Jow] (vol. 3).
- In medicine [Heid]: the 'whole' of body, Hippocrates, *Aphorisms* [Adams]: "*When the throat is diseased, or tubercles form on the body, attention must be paid to the secretions; for if they be bilious, the disease affects the **general system**; but if they resemble those of a healthy person, it is safe to give nourishing food.*"

These references indicate some understanding of the "system" notion that comes to us mostly by Plato, Aristotle, and Hippocrates. However, the notion has its origin to the Presocratic philosophers and their effort to comprehend the unknown physical world which is seen as an ordered system composed of basic elements which interact and produce a variety of new phenomena. Their effort to develop a rational investigation for the understanding of such phenomena, contributes to the initial formation of the concept of system. They question the basic elements, their composition, the emerging relations and structure, and the overall operation of the system. They examine the number of basic elements, the essence of these elements, their individual behaviour, the processes that take place and they ask questions, such as:

- What is the nature of the system?
- What are the basic, fundamental, primary elements of system's creation?
- How the system is being generated and what is the prime cause for this generation?
- How does the system work and how does it progress?

Thales is the first to introduce a system analysis by investigating the basic elements, by introducing the notion of analogy and producing the first classification of abstract geometric shapes. Anaximander considers the system (world) as a dynamical entity that is a time-dependant unity with ceaseless transformations, consisting of opposite qualities. Anaximenes goes further and examines the functional relations amongst the components variables, such as motion, density, and temperature. For Heraclitus, the system is more identified with a colossal process of events, changes, and facts, and he focuses his view mostly on the dynamical evolution of it. Empedocles, Anaxagoras, and the atomists Leucippus and Democritus, regard the system as a complicated whole. In particular, Anaxagoras speaks of compositions and decompositions that characterise the world and rejects the possibility of coming into being or ceasing to be. This notion of combining together in various ways the smallest parts of world, the atoms, results in the formation of macrocosm according to Leucippus and Democritus. Later on, Plato aiming at the proper definition of specific concepts explores the decomposition notion and sees that things can naturally be compounded into one and divided into many and introduces a method that goes on from the general to the particular and vice-versa.

2.5.1 The First Definition of System

The first explicit definition of the system in the ancient times has been given by the Lakonian and Pythagorean Kallicratides, in his work *«Περί οίκων ευδαιμονίας»* (On the Happiness of Family) [Stob]. He defines what a 'system' is and then explains it in terms of three examples (Anthology of J.Stobaei, Economicos, 16, 485) [Stob]:

"Any system consists of contrary and dissimilar elements, which unite under one optimum and return to the common purpose".

Example 1: The dance: Any particular dance constitutes a system in the singing societies. This system has a common purpose and ends up in a common result that is the harmony, the concordance of sound and motion.

Example 2: The crew: In the ships, the system of the crew is composed of contrary and dissimilar elements that unite in an optimal manner, guided by the captain, and realise the common purpose, which is the good sailing.

Example 3: The family (the household): The family exists as a system in societies formed by relatives, because it is composed of different parts that unite in realising an optimum form, under the head of the family, and achieve a common target.

According to Kallicratides, a system exists as a subsystem of a larger entity and it is characterized by:

- The particular, specific components-elements, i.e., the parts that determine its material substance (for example, the dancers of the dance, the crew in the ship, the members of a family).
- The desired common purpose that governs the system and specifies its behaviour (i.e., the concordance of sound and motion in a dance, the good sailing and the right direction of a ship, the harmony in a family).

Thus, a system consists of its elements, its structure, i.e., the way these elements connect between themselves, in terms of specific relations and of the desired target, goal and whose fulfilment depends on the good function of the system. Furthermore, a system is always embedded in a larger whole. Some special features in the discussion of the examples [Stob] indicate that the system as it is described by Kallicratides could be considered not as a simple open system, but as a closed loop control system bearing the following characteristics:

- The system consists of opposite parts, i.e., it contains the concept of contradiction It unites to the one optimum, which in modern terms called the controller that aims at the optimisation
- It returns to the common target, i.e., it holds the potentiality of return, or in other words it embodies feedback, which results in the common objective and ensures the desired balance and harmony.

2.5.2 The Holistic Approach of Hippocrates and the Concept of System

The notion of system, which is widespread in the Greek bibliography, acquires an integrated form in the field of medicine on account of Hippocrates (*ca.* 460-377 B.C.). He regards the human body as the general system, in which the particular organs and their operations fit harmonically. He considers all diseases as natural phenomena governed by natural laws. The basic Hippocratic principle, which is also associated with the notion of system, is the precise observation. The very characteristic of Hippocratic method is that it deals with the individual but it aims at a total unified picture of a diseased state. The construction of such a picture is based on the diagnosis resulting from careful inquiry and examination of all the factors regarded as significant.

The Hippocratics use experience to determine what is and what is not significant, observe certain indications that are not typical, compare the different cases and classify them. The observation of similarities and differences of cases results in generalisations, whereas the whole procedure results in the description and definition of a disease as an entity with certain character. The whole procedure of the Hippocratics diagnosis follows a cycle. The clinical records of previous cases are necessary for the creation of a new diagnosis, and each diagnosis they end up with is used as prognosis to a new case. The process of taking advantage of the results of experiments so as to make the right decision, is similar to a feedback process.

Hippocratics evolve their system of medicine by taking advantage of the previous doctrines of the Presocratic philosophers and of the Pythagoreans and exploit the concept of analogy. The Presocratics Anaximander and Heraclitus develop the idea that the man is subject to the same law as the universe. The Hippocratics on the other hand, believe in an analogy between the cosmos and the microcosm. This analogy may be turned either way, i.e., the human bodies imitate the cosmos as that the parts of the cosmos imitate human organs [Burk]. Therefore, they assert that the four elements, earth, air, water, and fire, form the basis not only of all things but also of the human body. When a person is healthy, i.e., in a normal state, the fluids exist in his body in harmonious proportions. The notion of the four elements and of the bodily humours, as crucial elements in health and disease, was already familiar by the time of the Presocratics philosophers. However, the bringing together in a theory that was to be known as Humoralism occurred by Hippocrates in his work on the *Nature of Man*. The Hippocratics derive their theories either from reasoning by analogy, or from experiments.

2.5.3 Aristotle's System of Logic

Aristotle, by means of his Logic, introduces a methodology that enables the development of the basic concepts involved in the system notion. In his work of *Organon*, he analyses the nature of scientific explanation and of the methods for establishing scientific principles and introduces the logical tools for scientific work. Aristotle's Logic is itself meant to be an *organon*, a tool, for reconciling and bringing into active co-operation the Sciences of the Laws of Thinking and the Laws of Nature.

According to Aristotle the kernel of scientific inquiry and explanation is something he calls Demonstration. Each object or phenomenon, in addition to certain individual characteristics, possesses also some necessary properties, which cause it to be the kind of thing it is. The scientific knowledge is closely related to the ability to demonstrate that a necessary property is inherent in an object because the object belongs to a specific species, which is characterized by that essential property. In other words, the concept of scientific inquiry is based on a method of proof, on a procedure of demonstrating by deductive or syllogistic reasoning that certain conclusions follow certain premises. Aristotle's standpoint that an event is explained if we specify its cause is expressed now by his view that the demonstration is the cause of a conclusion, the mean to exhibit the reason for the conclusion being what it is. Aristotle examines a system by analysing its logical structure and hierarchy. He introduces for the first time the notion of complexity in systems and deals with issues related to hierarchical forms of organisation of concepts and objects. This analysis enables the depiction of a system or process in terms of its components, their interaction and relations between components.

Aristotle proposed a scientific method which was highly influential for many centuries. His method, in broad terms, consisted of making observations of phenomena, using inductive arguments to deduce general principles which would explain the observations, then deducing facts about the phenomena by logical argument from the general principles. He saw this as leading from observations of a fact to an explanation of that fact [Klir,1].

Aristotle had stated that "the whole is more than the sum of its parts". It is a definition of the basic system problem which is still valid. Aristotelian teleology was eliminated in the later development of Western science, but the problems contained in it, such as the order and the goals of the living systems were negated and by-passed instead of solved. Hence the basic system problem is still open [Klir,1].

2.5.4 Final thoughts

The contribution of the Presocratic philosophers to the development of science is accompanied by a contribution to the emergence of the concept of system, as well as to the introduction of the basic characteristics of a system. Presocratic system-analysis searches for the basic elements of the system, introduces the notion of dynamical behaviour, the examination of the functional relations among the system's variables, and the conception of a system as a colossal process of events. Thus, in Presocratics' thought, we find the origin of the concept of system and some prominent steps of its evolution, but in an indirect way. A precise definition of the system is given by Kallicratides who defines the system notion in terms of the different elements, the structure, and the objective related to its operation. The system concept further evolves with Hippocrates in the context of medicine. In his theory, he ascribes to the system a further feature to those of the Presocratic, that of totality. Of wider interest in Hippocratic speculation, is the further development of the analogies between different physical phenomena or systems, contributing in this way to the concept of analogy, which in turn is closely related to the development of modelling. Last but not least, Aristotle adds an additional characteristic to the system, that of its organisation and development from basic notions. Hierarchy is the dominant feature in the system of thought he develops. These early developments provide the basis for the developments that follow in the Hellenistic period when the scientific method is properly developed and the system notions affect technology and the construction of engines and automata.

2.6 The need for a unifying framework of system concepts

The development of Information Society has produced many new challenges for traditional Sciences, Mathematics, Engineering disciplines and many other subject areas by posing new problems of clear interdisciplinary nature. Significant activity emerges in the cross boundaries between disciplines and this leads to the formation of new subject areas. Such developments require strong co-operation between many and diverse areas and disciplines and the adoption of approaches and methodologies common now-days to engineering disciplines. The central feature of interdisciplinary work is the effort to bridge diverse areas and disciplines and eventually lead to some form of integration.

Integration is an area with its own dynamics and crucial for its development is the presence of a common language, concepts, and methodologies which may act as facilitators for integration. This task is of holistic nature and this makes the need for development of a general systems framework of paramount importance. Such a framework is essential for *modeling* and the shaping of relevant analysis and design tools.

A systems account of the observed world and a systems approach to its problems are found in many different areas. Overcoming the barriers to learning requires a synthesis of many methods and disciplines, from mathematics and computer science to psychology and organizational theory. Theoretical studies must be integrated with field work. The aim is to try to study the interaction from multiple perspectives, holistically, to provide a general language with which to tie together various areas in interdisciplinary communication, to join together the many splintered disciplines with a "law of laws", applicable to them all and integrating all scientific knowledge [Skytt].

The aim of General System Theory when first developed, was to encourage the development of "theoretical systems which are applicable to more that one of the traditional departments of knowledge".

The aims of General Systems Theory were to be [Check]:

- To investigate the isomorphy of concepts, laws and models in various fields and to help in useful transfers from one field to another.
- To eliminate the duplication of theoretical efforts in different fields.
- To improve the unity of science through improving the communication between specialists.
- To develop integrated models, that describe the whole and relate to the particular.

It is evident from the previous discussion that, a general systems framework is missing from todays research areas, but is urgently needed. There have been many efforts to introduce such a framework of a general interest. Some of them have fallen into the pitfall of introducing so general ideas that any application, other than that of an abstract Managerial type, would be condemned of missing important information and thus provide an ill fated system, or integration of systems. Others have made the common mistake of ignoring the surrounding disciplines and focusing on just one, thus deliberately shutting out any port of communication between areas. Although delivering a framework suitable for one area, is wise and efficient in its locality, a more unifying framework has all the advantages previously mentioned, of the uniform exchange of ideas and applications, as well as more possibilities of a successful integration.

The issue of integration must be broken down and seen as a step by step approach, and dealt with from several perspectives. It must be dealt as the *fundamental* rethinking and *radical* redesign of how the system is defined, to achieve *dramatic* improvements in performance, quality, service, and speed:

- The observer. The observer is an integral part of the system definition and afterwards the modelling process. He/She represents the area for which the system is defined, and therefore sets the priorities and the purpose of the system. To promote a platform for integration, there has to be an *evolution of the conceptual framework*. The identification and definition of the system has to have a domain independent architecture. This is somehow abstract, but has to be a parallel action with the following.

- The terminology. The terminology is the second important ingredient to a successful attempt for integration. It represents a kind of transparency; a level of access to information and exchange of ideas that allows areas to communicate and interact for maximum efficiency and adapt terms to meet the needs of unifying system framework. This will not mean that each discipline will lose its identity-anyway the diversity of the various disciplines today is down to the need for further specialisation. With the promotion of this integrated platform for system definition each discipline will use, tailor, extend and create their own system definitions, but under some generally accepted rules.

The idea of integration is to provide a frame of system identification, and thereafter system modelling, and not to obstruct and prevent any kind of imaginative creation or the freedom of expression either of the observer of the individual disciplines. It must be seen as the basis of constructing, capturing and evolving knowledge.

The following chapter provides an integration of system definition in all its levels and contexts. This following effort will *not* be a try to integrate systems! It will be a laying out of the simple rules that *any* system possesses. The integration involved here will be an integration of all the definitions and notions that are available today about the system, so as to create a framework of system notions, relevant to all contexts.

In recent years, an increasing attention has been paid to the development of domain-specific modeling languages. It is believed that these languages can lead to an increase in productivity in the modeling activity and contribute to the production of models that are more flexible, reusable and easier to maintain than models produced by using general-purpose modeling languages. Notwithstanding, in order to be effective, a domain specific modeling language must be defined taking into account the needs of its client users. From their perspective, the use of the language should be intuitive and satisfactory in the following terms:

- The semantics of the produced models should be clear, i.e., it should be easy for a model designer to recognize what language constructs mean in terms of domain concepts;
- The language should be sufficiently expressive to represent all domain concepts that should be captured by the intended models.

For these reasons, there is a demand for techniques that support the construction of explicit models of domain conceptualizations. Additionally, there is a need for concrete and precise guidelines for selecting which domain concepts should be represented as language constructs and how.

Chapter 3

System Description and System Notions

The development of Information Society has produced many new challenges for traditional Sciences, Mathematics, Engineering disciplines and many other subject areas by posing new problems of clear interdisciplinary nature. Significant activity emerges in the cross boundaries between disciplines and this leads to the formation of new subject areas. Such developments require strong co-operation between many and diverse areas and disciplines and the adoption of approaches and methodologies common now-days to engineering disciplines. The central feature of interdisciplinary work is the effort to bridge diverse areas and disciplines and eventually lead to some form of integration. Integration is an area with its own dynamics and crucial for its development is the presence of a common language, concepts, and methodologies which may act as facilitators for integration. This task is of holistic nature and this makes the need for development of a general systems framework of paramount importance. Such a framework is essential for modelling and the shaping of relevant analysis and design tools.

The development of a systems framework for general systems is not a new activity [Tak. & Takah., 2], [Mes., 2], [Gog., 1] and [Gog., 2]. However, such developments have been influenced predominantly by the standard engineering paradigm and as a result they failed to cope with new paradigms such as those of the business processes, data systems, biological systems, and so on. Our task here is to reconsider existing concepts and notions from the general Systems area, detach them from the influences of specific paradigms and generalise them appropriately to make them relevant for the new challenges. This is a first attempt to identify the areas and issues where abstract system theory requires extensions and modifications in order to cope with the challenges of the new paradigms, such as that of Business Processes (BP), Life-cycle issues in Engineering Problems (LCEP), Data Structures (DS) etc.

3.1 The System and its Structural Features

Dealing with systems coming from many and diverse disciplines requires defining the abstract notion of the system in a way such that:

- It encompasses the basic features of all classes of paradigms known at the moment.
- It has the potential to specialise and being capable to cover the special features of certain interesting classes.
- It provides the potential to build up concepts and properties in a progressive way from the general to the particular.

This section is an attempt to establish such a general conceptual framework for general systems which contains the basic features of the dynamical systems emerging in physical and engineering properties, which have been the main sources of motivation so far and are characterised by rich properties.

3.2 The notion of the System

The definition of a system, that is given here, is rather general and aims to encompass many paradigms (including the traditional engineering and business ones).

*"A **system** is an interconnection, organisation of objects that is embedded in a given environment."*

Such a definition is very general and uses as fundamental elements the primitive notions of: *objects*, *connectivities – relations* (topology), and *environment*. It can be symbolically denoted as:

SYSTEM $\Longleftrightarrow$ **OBJECTS + INTERCONNECTIONS+ BOUNDARIES – ENVIRONMENT**

This definition of a system is suitable for the study of "soft", as well as "hard" systems and it is based on a variety of paradigms coming from many and diverse disciplines. It does not make use of notions such as causality, input-output orientation, definition of goal, behaviour, and so on. Quite a few systems do not involve these features and thus they have to be introduced as additional properties of certain families.

The general family of systems which are considered may be classified according to their origin and properties to the following main classes:

- Natural Systems
- Designed Systems
- Influenced Natural Systems

The term of ***natural systems*** refers to those that appear in nature, biology, society etc which are products of an evolutionary process that is not under the control of human activity, although in certain cases human activity may have been part of their formation process. The exactly opposite class is that of ***designed systems,*** which are entirely the product of human activity and this includes all engineering systems. The huge recent developments in science and technology have created the conditions for the human intervention to certain natural processes and systems and this leads to the emergence of systems which are mainly natural, but with certain parts with a designed character (or influenced by the human activity in a controlled manner); it is this new and interesting class that will be referred to as ***influenced natural systems.***

It should be noted that the notion of a system is conceptual and involves fundamental notions that come from our experiences, previous knowledge. The concept of a system refers to the level of reality, i.e., we consider it as a physical or manmade construction, which is part of our sphere of our reality. This observation is essential, since later on we shall examine the notion of system model, which involves a level of understanding of what this reality is and thus introduces an abstraction in the form of a model. This definition has an element of abstraction itself, since the basic, primitive elements, on which the definition is based, have a generic, abstract character. We consider first the meaning of the basic ingredients of the system's definition.

3.3 The notion of Objects

In this study, *objects* are considered to be the most primitive element, allowing them to be almost anything. By not restricting the definition to any particular class, we allow freedom to construct any type of system that is of interest.

*An **object**, B, is a general unit (abstract, or physical) defined in terms of its attributes and the possible relations between them.*

As such, an object is defined as an entity in our world (reality, or conceptual); the more precise definition of this entity involves some additional primitive notions. Thus, we consider here objects to be the most primitive concept and we allow objects to be almost anything. By not restricting the definition to any particular class we allow the freedom to consider systems from any domain. Objects are defined in terms of a collection of observations of some selected characteristics, attributes and of the relations between these attributes, expressing in this way, a form of organisation or a degree of knowledge. The relations between the attributes may be functional, linguistic, structural, and so on. They express knowledge that stems from past history of the object or the environment in which it has been operating. Thus, a collection of observations and the possible relations between them are considered in order to define precisely an object. The general classification of systems into ***natural, designed,*** and ***influenced natural systems*** is also adopted for objects in the same way.

There are different types of objects, which may be classified using some basic properties as:

- atomic versus composite
- neutral versus relational
- determinate versus indeterminate

The consideration of objects, which are ***atomic***, implies our inability or lack of desire to decompose them into simpler elements. The emergence of objects, which are themselves interconnections of other more basic objects, leads to the notion of ***composite*** objects, or forms. ***Neutral*** objects are simply characterised by their attributes, whereas ***relational*** objects involve both attributes and the relations defined between them. Object with well defined attributes and possibly relations between them are called ***determinate***, whereas those for which there is lack of knowledge on the nature of their attributes and/or relations between them (partial or complete) are said to be ***indeterminate***. Further classifications will be subsequently introduced, when further primitive notions become more precise.

3.4 The notion of the Environment

The notion of an object, or a system involves some separation from other objects, or systems in terms of a boundary, that may be refereed as ***object's boundary*** and this introduces the notion of the environment:

For a given object, we define its ***environment*** as the set of objects, signals, events, structures, which are considered topologically external to the object, and are linked to the object in terms of a structure, relations between their attributes.

The essence of this definition is that for a given object, a boundary around the object is defined that includes all structures and attributes associated with it. The object under consideration may be related to other objects that are, however, considered to be external, since they are external with respect to the boundary. The existence of the objects environment implies crossings of the imaginary boundary and such crossings indicate the connectivities of the object to objects in its environment. This provides a local view of the environment and the interconnection of it with the other objects of the system. The previous statement may be indicated by the following Figure 3.1.

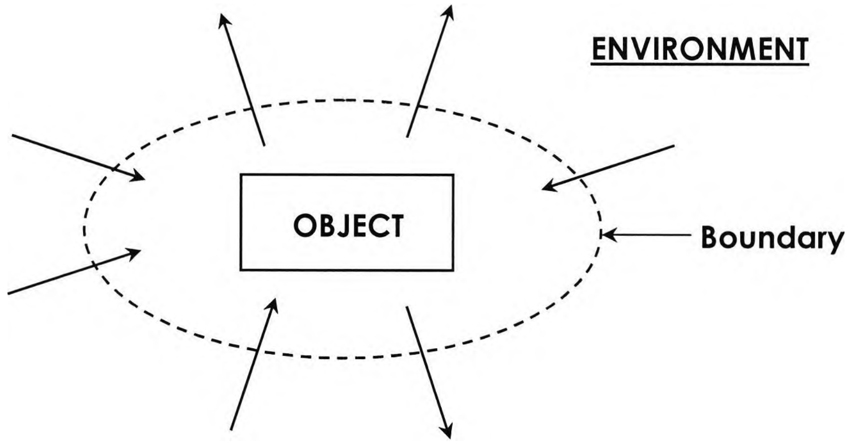


Figure 3.1: Object Environment Relation

The linking of objects to an environment provides a classification of objects into classes such as free and embedded objects. A ***free*** object has no links to its environment and thus its attributes are not affected by external stimuli and do not affect its environment. An ***embedded*** object has active links to its environment and thus its attributes are affected by external stimuli, events. In the above diagram, lines crossing the boundary indicate the embedding of the object to its environment. Free objects may be studied on their own without consideration of an environment. The linking of objects to their environment provides a further classification of them. An embedded object that has no links, or stimuli coming from the environment is called ***autonomous;*** an autonomous object that has no influences that feed to the environment will be called ***non-active***; otherwise, it will be called ***active***. Clearly, an object may function as free, or embedded, autonomous depending on external events, or on time; the term *mode* will be used to refer to the different ways one considers an object.

3.5 The notion of the Interconnection Topology

The set of objects in a system are related between themselves and their environment and this defines the third important ingredient of the notion of the system which depends on the primitive notion of relation and defined below:

The set of objects in a system are related between themselves and to the system environment in a specific way and these relationships are referred to as ***interconnection topology***. The part of topology expressing the internal linking between the objects of the system defines the ***internal interconnection structure***, whereas that part expressing the links of the objects to the system's environment will be called ***external interconnection topology***.

The interconnection structure may be fixed (time and event independent) or varying (time or event dependent); this introduces an additional classification for embedded objects and systems. An object, or a system for which its interconnection topology is fixed will be referred to as ***topologically fixed***, whereas those for which even part of its interconnection topology varies will be called ***topologically varying***. Alternative classifications may be introduced by distinguishing between the internal and external parts of the interconnection topology, or based on the nature of the interconnection topology. The latter leads to classes such as ***natural-, designed-, influenced natural- interconnection topologies***, which follow the classification previously given for systems.

3.6 Attributes, Variables and Parameters of an Object

The deeper characterisation of objects uses the notion of attributes associated with them and this is formally defined below:

*"An **attribute** for an object is an identifiable and possibly measurable characteristic of the object."*

In the current context, identifiability implies the ability to associate a label, tag with the object, whereas measurability requires ability to quantify it by associating a set of values, or functions. The notion of attribute allows a more detailed characterisation of an object, since it introduces elements that may be used to provide more concrete forms of description for the object. Describing objects (and systems) in terms of attributes lays the basis for making the transition from conceptual to quantifiable descriptions and forms. Using the notion of attributes some of the previous classifications may take a more specific form. In fact, objects may be in two distinct modes: An object is said to be in ***static mode***, when its attributes are not changing as time changes. The object will be said to be in a ***varying mode***, when its attributes are changing as a function of time, or events and it will be in a ***dynamic mode,*** when it is in a varying mode that depends on initial conditions. Note that the term dynamic refers to phenomena that produce time-changing patterns, the characteristic of the pattern at one time being interrelated with those of other times due to the dependency on initial conditions. This term is nearly synonymous with the ***time-evolution***, or ***pattern of change***. Thus, 'dynamic' refers to unfolding of events in a continuing evolutionary process that depends on the past, as this depends on the dependency on initial conditions.

Associating attributes to objects allows the description of the interconnection topology in a formal way as relations on the sets of object attributes. The introduction of ordering in these relations is a formal way for introducing causality in an abstract way and thus some direction in the described relations.

The family of embedded objects may be further classified by assigning direction to individual relations, thus introducing some abstract form of “causality” and represent it by the line connectivities and arrows. The assignment of direction is equivalent to introduction of ‘causality’ in the traditional way (cause and effect). In Figure 3.1, arrows directed from the environment to the object are referred to as ***object inputs***, or ***stimuli*** and those with direction from the object to the environment are called ***object outputs***, or ***influences***. Stimuli, inputs are attributes generated in the object’s environment and affect the overall behaviour of the object; influences, outputs on the other hand are attributes generated within the object which cross the boundary and thus feed to the environment. For autonomous objects the evolution of their attributes is not affected by external stimuli, but only from internal ones (i.e., initial conditions). An object that is non-active does not generate stimuli for objects in its environment. Objects are referred to as ***non-autonomous***, or ***forced***, when they have inputs; in this case the environment plays a crucial role in determining the evolution of the object variables. Similarly, the object will be called ***active***, when there are outputs, or influences; in this case the object generates stimuli for objects in its environment. Such classification stems from topological considerations and some basic understanding of causality, or notion of flow.

Attributes, which are determined on an object in the autonomous static mode, will be referred to as ***parameters*** and those that are defined in the active, or embedded mode will be called ***variables***. Note that such a classification may depend on the different stages of the lifecycle of the object. Note that such a classification may depend on the different stages of the lifecycle of the object. Determining the parameters of the object may be the outcome of direct observations, a priori knowledge, or the result of experiments at some time in past.

3.7 The notion of State

We shall denote an object by B and by $\boldsymbol{P_B}$, $\boldsymbol{V_B}$ the corresponding set of parameters and variables. Assuming the object B to be in the general embedded (or autonomous) mode, we may define ***relations***, with the standard mathematical meaning of the term, on the set V_B; these may be trivial (identity mappings), or nontrivial and this set will be denoted by $R(B)$ and referred as the ***full relations set*** of B. A subset of $R(B)$ may be defined when the object is considered in the autonomous active mode; this is denoted by $R_a(B)$ and referred as ***autonomous relations*** set of B. The subset of nontrivial relations will be denoted by $\tilde{R}(B)$ and provide some form of identity for the object B. The absence of trivial relations guarantees some form of independence for the resulting set of variables in V_B. We may further classify $\Im(B)$ as follows:

"The subset of V_B on which there are no trivial relations will be referred as ***implicit states*** *and denoted by $\Im(B)$. All variables of V_B, which have nontrivial relations and are defined when the object is in the autonomous mode, will be called* ***extended states*** *and the set is denoted by $\hat{\Im}(B)$. Any subset of $\hat{\Im}(B)$, with the additional property that the elements are independent (in some algebraic sense) and describe completely $\tilde{R}(B)$ will be referred as* ***states*** *of B. The latter are denoted by $\tilde{\Im}(B)$."*

The notion of the state is fundamental for dynamic objects and it is defined on the free mode, as well as the embedded mode. Defining the notion of the state permits complete knowledge of all attributes defined on the object, and provides ways for describing the time, event evolution of all object attributes. For embedded objects, described as in Figure 3.1, the totality of variables associated with the object may be classified as shown below:

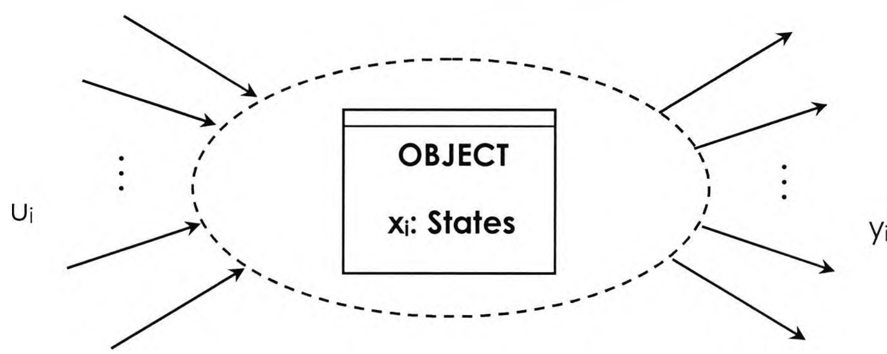


Figure 3.2: Object and Fundamental Variables

In Figure 3.2 we have: x_i are the elements of the state set $\tilde{\Im}(B)$, u_i expresses the stimuli (originating from the environment), and y_i the influences of the object to its environment. We may arrange all such variables in the form of vectors and this is represented by: $\underline{x} = [..., x_i, ...]'$, $\underline{u} = [..., u_i, ...]'$, $\underline{y} = [..., y_j, ...]'$; $\underline{x}$, $\underline{u}$, $\underline{y}$ are referred to correspondingly as ***state-, input-, output-vectors***; the vector $\underline{\xi} = [\underline{y}', \underline{x}', \underline{u}']'$ made up from the three sub-vectors is referred to as the ***composite vector*** of the object.

Such a definition is intuitive and rather conceptual, and for different classes of objects, an appropriate mathematical formalism is required in order to provide answers to issues such as:

- Proper classification of parameters and variables.
- Definition of independence of variables and relations.
- Expression of dependent variables in terms of an independent set of variables.

Clearly, for general dynamical systems and linear systems in particular, the vector space algebraic set up is adequate for addressing the above. For general classes of objects a more general framework is needed.

The set of implicit states, extended states and states associated with an object may be ordered into a vector form. Such representations may be denoted as $\underline{\xi}_B$, $\hat{\underline{x}}_B$, $\tilde{\underline{x}}_B$ respectively and $\Im(B) = \{\underline{\xi}_B\}$, $\hat{\Im}(B) = \{\hat{\underline{x}}_B\}$, $\tilde{\Im}(B) = \{\tilde{\underline{x}}_B\}$ provide an explicit representation of the corresponding elements. To every variable in $\boldsymbol{V_B}$ there is an associated set of values, indexed by time or events, and this is called the *range set* of the particular variable. The ordered set of all range sets of $\boldsymbol{V_B}$ is denoted by $\boldsymbol{T_B}$ and it is called the ***space*** of $\boldsymbol{V_B}$. An important number associated with objects is the cardinality μ_B of the set $\boldsymbol{V_B}$. The corresponding spaces for $\underline{\xi}_B$, $\hat{\underline{x}}_B$, $\tilde{\underline{x}}_B$ are denoted respectively by Z_B, $\hat{X}_B$, $\tilde{X}_B$ and shall be referred as ***implicit-, extended-, regular-state space*** respectively of *B*. The space

$\mathcal{Z}_B$ expresses all knowledge one has about the object at a given time and for a set of events that have occurred. Such a space, however, does not necessarily imply minimality, independence, as far as the information contained. Note that for the above spaces the following set relationship holds:

$\tilde{\mathcal{X}}_B \subseteq \hat{\mathcal{X}}_B \subseteq \mathcal{Z}_B \subseteq \boldsymbol{T_B}$

During the object's lifecycle, certain variables may vanish and new ones may appear together with the possible changing of values of the variables that are preserved. There are two extreme cases associated with the object: The first corresponds to the disappearance of all variables at a certain instance and it is referred as the ***death*** of the object and the second characterises the ***birth*** of an object characterised by the emergence of a number of variables associated with it. During the lifecycle of a process, objects may have variables disappearing, or emerging and this leads to the following definition.

*"An object B, for which its cardinality μ_B of $\boldsymbol{V_B}$ remains fixed throughout the object's lifecycle, will be called **fixed**; otherwise, it will be called **changing**."*

The above definition is motivated by the need to consider the system's evolution throughout its lifecycle. Indeed, objects in systems may have variable cardinality and some of them may be born, or die throughout the system's lifecycle. The fact that we refer to $\boldsymbol{V_B}$ varying cardinality implies that variables linked to the interconnection topology of objects may also emerge or disappear. The latter implies that the interconnection topology may be also subject to variability.

The cardinalities of the state sets $\hat{\mathcal{X}}_B$, $\tilde{\mathcal{X}}_B$ provide measures of complexity of the object, they will be denoted by $\hat{\mu}$, $\tilde{\mu}$ and they will be referred to as ***extended-, regular-dimension*** of the object respectively. Clearly, from the definition we have:

$$\tilde{\mu} \leq \hat{\mu} \leq \mu_B$$

3.8 Relations and Object Properties

The relations defined on V_B of an object B provide a characterisation of it in an explicit way. Objects for which the full relations set $R(B)$ contains at least one nontrivial element will be called ***relational*** objects, otherwise, they will be called ***neutral.*** The class of relational objects has richer properties and they will be examined further. If $\underline{\xi}_B$ is the implicit vector (denoted whenever there is no ambiguity by $\underline{\xi}$) of a relational object, then the existence of nontrivial relations amongst the components of $\underline{\xi}_B$ may be expressed in a functional form as f_B:

$$f_B(\underline{\xi}_B) = 0 \qquad (3.1a)$$

or alternatively as

$$h : \{\underline{\xi}\} \to \{\underline{\xi}\} \qquad (3.1b)$$

In this relation $\{\underline{\xi}\}$ denotes the set of all values of the composite vector for the complete lifecycle of the object; $\{\underline{\xi}\} \equiv Z_B$, is referred to as the ***composite space*** of the object. A description like that of (3.1) does not make a distinction between the input, output, state components of $\underline{\xi}$ and thus we say that lacks orientation; for this reason the description in (3.1) will be referred to as ***non-oriented relational*** description. For relational objects where $\underline{\xi}$ may be partitioned in the $\underline{\xi} = [\underline{y}', \underline{x}', \underline{u}']'$ form, where $\underline{x}$, $\underline{u}$, $\underline{y}$ correspond to state, input, output vectors it may be possible to express the relation in (3.1) in the form:

$$f : \{\underline{x}\} \times \{\underline{u}\} \to \{\underline{x}\}, \quad r : \{\underline{x}\} \times \{\underline{u}\} \to \{\underline{y}\} \qquad (3.2)$$

Whenever we define such relational descriptions, we say that the specific partitioning of $\underline{\xi}$ introduces an orientation and (3.2) introduces an ***oriented relational*** description. Note that $\{\underline{x}\}, \{\underline{u}\}, \{\underline{y}\}$ denote the sets of all values of the $\underline{x}$, $\underline{u}$, $\underline{y}$ vectors for the complete lifecycle of the object and shall be referred to in turn as ***state-, input-, output-spaces*** associated with the given orientation of the object. The orientation of the object may be interpreted as a partitioning of $\underline{\xi}$, such as h produces a corresponding pair of relations (f,r), which define (3.2). For an object having an oriented relational description as in (3.2) it may be possible to have a relationship:

$$g:\{\underline{u}\}\rightarrow\{\underline{y}\} \qquad (3.3)$$

This relationship links in a direct way inputs (stimuli, causes) to outputs (influences, effects); whenever such relations may be established, the object will be called ***input-output relational***. For traditional dynamical systems, the above expression has a standard functional interpretation. For more general objects, there is a need to develop an appropriate algebraic set up within which such relations may be interpreted. The specification of exact nature of such relations is intimately linked to the subject of modelling.

3.9 Embedding of Objects in a System

The classification of implicit variables to inputs, outputs and states (extended states) requires a more refined consideration when we consider an object as part of a system.

The embedding of an object to its environment (case of embedded objects) implies that the implicit vector $\underline{\xi}_B$ may be partitioned as:

$$\underline{\xi}_B=\left[\underline{w}_B^t;\hat{\underline{x}}_B^t;\underline{v}_B^t\right]^t \qquad (3.4)$$

where $\underline{w}_B$, $\underline{v}_B$ denote vectors expressing ***generalised outputs***, ***generalised inputs*** respectively, that is variables feeding through to other objects, associated with external objects respectively that express the interactions of the object with its environments locally; $\hat{\underline{x}}_B^t$ denote the internal variables or extended states of the object. The vectors $\underline{w}_B$, $\underline{v}_B$ are manifestations of the ***local structure*** of the object, that is the way the given object is embedded in its environment. The vector $\underline{w}_B$ may be partitioned as

$$\underline{w}_B = \left[\underline{y}_B^t, \underline{\hat{w}}_B^t\right]^t \qquad (3.5)$$

where $\underline{y}_B$ denotes the vector of ***measurements,*** observations and $\underline{\hat{w}}_B^t$ are the variables, which are not necessarily measured, but become inputs to other objects and are referred to as ***output connections***, or ***output influences.*** Similarly, $\underline{v}_B$ may be partitioned as

$$\underline{v}_B = \left[\underline{u}_B^t, \underline{\hat{v}}_B^t\right]^t \qquad (3.6)$$

where $\underline{\hat{v}}_B$ is a vector with variables associated with other objects and referred to as ***input connections***, or ***input influences;*** $\underline{u}_B$ is a vector with arbitrarily assignable variables and are referred to as ***inputs***. The vectors $\underline{\hat{v}}_B$, $\underline{\hat{w}}_B$ express the way the object interacts with its environment. The definition of partitioning of the implicit vector $\underline{\xi}_B$, as defined above, describes explicitly the embedding of the object to the environment, specifies the local structure and it is referred as the ***orientation*** of the object with respect to its environment. This orientation is described diagrammatically as

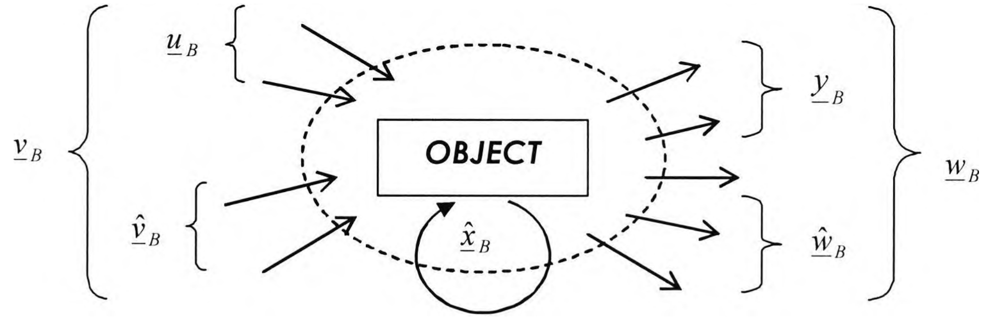


Figure 3.3: Embedding of an Object and Classification of Variables

The spaces corresponding to $\underline{v}_B$, $\underline{\hat{v}}_B$, $\underline{u}_B$, $\underline{w}_B$, $\underline{y}_B$ vectors will be denoted by V_B, $\hat{V}_B$, U_B, W_B, Y_B respectively and referred as:

V_B : ***input influence space*** , W_B : ***output influence space***

U_B : ***control input space,*** Y_B : ***output measurement space***

and characterise the nature of the object's embedding during its lifecycle. An interesting class of relational objects are those for which (3.1) expression may lead to explicit oriented relations as:

$$\underline{y}_B = \Phi_B(\underline{u}_B, \hat{\underline{v}}_B) \qquad (3.7)$$

$$\hat{\underline{w}}_B = \hat{\Phi}_B(\underline{u}_B, \hat{\underline{v}}_B) \qquad (3.8)$$

A relational object B for which the implicit relation (3.1) can be expressed in the form (3.5), (3.6) will be called ***strongly oriented*** *and* Φ_B, $\hat{\Phi}_B$ *will be referred as* ***transfer-*** *and* ***co-transfer-relations****.*

It is worth noting that if the object environment is structured, then the object orientation is partially structured (some freedom may exist in the selection of $\underline{u}_B, \underline{y}_B$). If, however, the environment is not explicitly defined, then the orientation of the object may be the result of specific design (of input, output structure) and thus, not uniquely defined. In the case where only (3.7) may be explicitly defined, then the object will be called ***weakly oriented***.

Objects have been identified with the set of variables and relations associated with them. For a given variable of an object, x, its values evolve as a function of time and/or events. The time and possibly event evolution of a variable is called ***behaviour*** of the variable and it is a notion that extends to the implicit and to other vectors associated with the object. Relations on an object, which involve behaviours of variables, will be called ***behavioural relations*** and, in general, they comprise a proper subset of the relations defined on an object; such relations are crucial in describing the lifecycle evolution of an object.

3.10 Composite Objects: Structure and Basic Properties

The essence of a system is the organisation of objects in a specific way and in general terms this is what describes the notion of a composite structure. Oriented objects, in the form described by Figure 3.3, interact in a certain manner via a given topology of influences, and this produces composite objects exhibiting properties that in general are different to those of constituent objects. Composition of objects has two main aspects:

- Definition of an object-based, local, interconnection structure
- Rules for interconnecting objects

The definition of an orientation, as described in the previous section, involves a partitioning of the implicit vector $\underline{\xi}_B$ as:

$$\underline{\xi}_B = \left[\underline{y}'_B, \underline{\hat{w}}'_B; \underline{\hat{x}}'_B; \underline{\hat{v}}'_B, \underline{u}'_B\right] \tag{3.9}$$

and such a partitioning leads to the following definition:

Given an orientation of the implicit vector of an object B, as in (3.9), we define:

① *The pair* $(\underline{y}_B, \underline{u}_B)$ *together with the relations* R_y, R_u *of* $\underline{y}_B$, $\underline{u}_B$ *with the implicit vector* $\underline{\xi}_B$ *as the* ***decision structure*** *of the object;* $(\underline{y}_B, R_y)$ *defines the* ***measurement*** *and* $(\underline{u}_B, R_u)$ *the* ***control structure*** *of the object.*

② *The pair* $(\underline{\hat{w}}_B, \underline{\hat{v}}_B)$ *together with the relations* $R_{\hat{w}}$, $R_{\hat{v}}$ *of* $\underline{\hat{w}}_B, \underline{\hat{v}}_B$ *with the implicit vector* $\underline{\xi}_B$ *as the* ***interconnection structure*** *of the object;* $(\underline{\hat{w}}_B, R_{\hat{w}})$ *defines the* ***output connection structure*** *and* $(\underline{\hat{v}}_B, R_{\hat{v}})$ *the* ***input connection structure****.*

③ *The vector* $\underline{\hat{x}}_B$ *together with its relations* $R_{\hat{x}}$ *with the implicit vector* $\underline{\xi}_B$ *as the* ***internal structure*** *of the object.*

An oriented object enters relations with other objects and the embedding of the object in the general environment is described in Figure (3.4). The vectors $\underline{u}_B$, $\underline{y}_B$ are elements of the ***global control*** and ***measurement structure***, that will be defined later, whereas the pair $(\underline{\hat{w}}_B, \underline{\hat{v}}_B)$ define the available variables that may enter relations with other objects and thus are elements of the interconnection ***structure*** that is defined below:

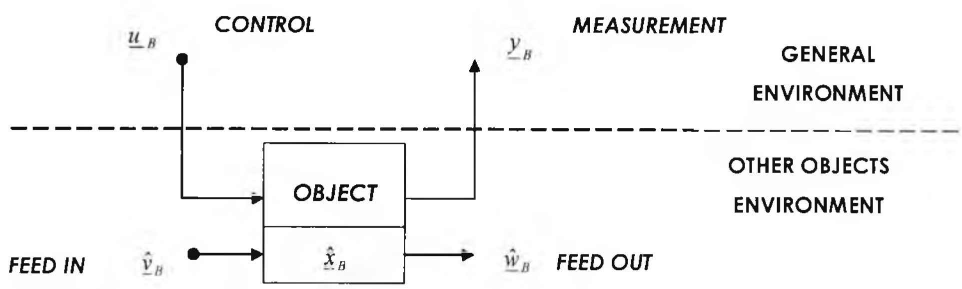


Figure 3.4: Objects interaction with environment

Consider a set of objects {B_i, $i \in p$} with interconnection vectors $(\underline{\hat{v}}_i, \underline{\hat{w}}_i)$, $i \in p$, *and let* $\hat{V}_i$, $\hat{W}_i$ *be the corresponding spaces of* $\underline{\hat{v}}_i, \underline{\hat{w}}_i$ *respectively. If* $\hat{V} = \hat{V}_1 \oplus \cdots \oplus \hat{V}_p$, $\hat{W} = \hat{W}_1 \oplus \cdots \oplus \hat{W}_p$ *denote the aggregate spaces, that is* $\forall \underline{\hat{v}} \in \hat{V}$ *and* $\underline{\hat{w}} \in \hat{W}$:

$$\underline{\hat{v}} = \left[\underline{\hat{v}}_1^t, \cdots, \underline{\hat{v}}_p^t\right] = \underline{\hat{v}}_1 \oplus \cdots \oplus \underline{\hat{v}}_p,$$
$$\underline{\hat{w}} = \left[\underline{\hat{w}}_1^t, \cdots, \underline{\hat{w}}_o^t\right] = \underline{\hat{w}}_1 \oplus \cdots \oplus \underline{\hat{w}}_p \qquad (3.10)$$

then we define as an ***interconnection structure*** *any map* F *such that* $F: \hat{W} \to \hat{V}$.

It is clear that the nature of objects and their associate variables define the nature of the interconnection structure. Such structures may not necessarily be fixed but they may change as a result of discrete events. Interconnection structures where some objects are evolving and/or the structure may change as a result of events will be called ***changing***; otherwise, they will be called ***fixed***. Evolving interconnection structures appear in many areas of engineering and non-engineering type systems and correspond to faults, redesign of processes, changes of requirements, etc. Typical forms of interconnection structure are those defined by graphs. Having defined the composition structure we may now define a composite object or a process as follows:

Let {B_I, i∈p} be an ordered set of objects and F *be an interconnection structure defined on them. The object defined by the given set* $B_a = \{B_1, \ldots, B_\rho\}$ *will be also denoted by* $B_a = B_1 \oplus \cdots \oplus B_\rho$ *and will be called the* ***aggregate object*** *of the set. If* $*$ *denotes the action of the interconnection structure on* B_a *, then the resulting object is called* ***composite object****, or* ***process*** *and it is represented by*

$$P = [\{B_i\}; F] = B_a * F \qquad (3.11)$$

A composite object may be represented in a similar way to that of an object and it is indicated in Figure 3.5. The extended state vector $\hat{\underline{x}}_\rho$ is the aggregation of the corresponding vectors $\{\hat{\underline{x}}_{B_i}, i \in\}$, i.e., $\hat{\underline{x}}_\rho = \hat{\underline{x}}_{B_1} \oplus \cdots \oplus \hat{\underline{x}}_{B_\rho}$, but the relation $R_{\hat{\underline{x}}_\rho}$ is in general different than those defined by the set $\{R_{\hat{\underline{x}}_{B1}}, \ldots, R_{\hat{\underline{x}}_{B\rho}}\}$, since it is affected by the composition rule F. Regarding control and measurements, we have that

$$\underline{u}_\rho = u_{B_1} \oplus \cdots \oplus u_{B_\rho}, \qquad \underline{y}_\rho = \underline{y}_{B_1} \oplus \cdots \oplus \underline{y}_{B_\rho} \qquad (3.12)$$

The set of variables in $\{\hat{\underline{v}}_{B_i}, i \in \rho\}$, $\{\hat{\underline{w}}_{B_i}, i \in \rho\}$ are external for the constituent objects but they may, or may not enter the relations implied by F and thus they may become internal for the composite object P; the variables entering the relations defined by F are called ***active***, whereas those not affected by F are referred a ***neutral.*** Active variables enter relations and thus become internal for the object, whereas neutral variables remain unaffected and thus external for the composite object. By collecting all neutral components from the sets $\{\hat{\underline{v}}_{B_i}, i \in \rho\}$, $\{\hat{\underline{w}}_{B_i}, i \in \rho\}$ respectively we define vectors $\hat{\underline{v}}_\rho$, $\hat{\underline{w}}_\rho$ which are external to the composite object. In fact, the components of $\hat{\underline{v}}_\rho$ express variables that are externally specified (either as control variables, or outputs from another process) and those in $\hat{\underline{w}}_\rho$ express products of some processing which may, or may not be measurable completely. The composite system is thus represented as follows:

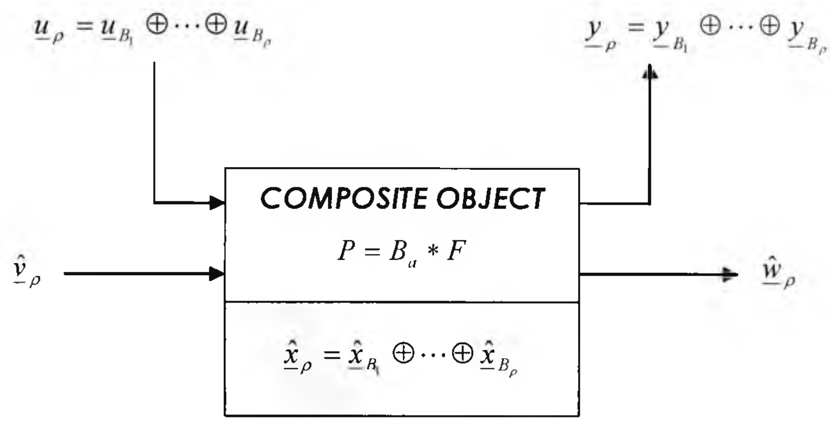


Figure 3.5: Composite object representation

The relations R_y, R_u , $R_{\hat{w}}$, $R_{\hat{v}}$, $R_{\hat{x}}$ associated with the composite object are shaped by the interaction of F with the aggregate object B_a and thus the resulting behaviours of the corresponding vectors are in general different to those simply defined on the aggregate object. The role of the interaction structure map F is thus crucial in determining the nature of the composite object P. The nature of the process we are dealing with determines the specific properties of F and leads to the following general classification of them:

- Natural maps
- Design maps
- Influenced natural maps

Such a classification is not related to the mathematical structure of F, but to the process that leads to the formation of the composite object. Composition of objects occurring as a result of phenomena beyond the control of a 'designer', or which emerge as unique solutions to design problems have a map F which is ***natural maps.*** In such cases, the properties of F are beyond the control of any designer and thus issues of modification of them cannot be addressed due to lack of control mechanisms. This case includes physical, chemical, biological processes that occur in nature and cannot be influenced by human interaction. The crucial feature of this case is the lack of accessibility of the structure, that is ability to modify parts of it and thus its properties. A number of processes are entirely manmade and therefore, the corresponding map F is the subject of a design process; such maps are referred as ***design maps*** and issues related to their synthesis are significant. Many processes exist, either natural or designed, where there is some ability to modify them to improve properties without altering the original functionality, purpose of the process; we shall refer to them as ***influenced natural maps*** and for such processes redesigning F within a set of constraints is an important issue under consideration.

3.11 The system and its environment

Composite objects may be combined, according to some rule defined by an overall interconnection topology F', and this leads to the definition of a system that may be represented in a general form as shown in Figure 3.6. If $S_a = \{p_1, \ldots, p_v\}$ denotes the aggregate of processes interconnected under F', then the system S is the interconnection of S_a and the F' topology and $\underline{x} = \hat{\underline{x}}_{p_1} \oplus \ldots \oplus \hat{\underline{x}}_{p_v}$ is the system ***state vector***, whereas vectors $\underline{u} = \underline{u}_{p_1} \oplus \ldots \oplus \underline{u}_{p_v}$, $\underline{y} = \underline{y}_{p_1} \oplus \ldots \oplus \underline{y}_{p_v}$ are respectively the ***input*** and ***output vectors***, which are expressed as aggregates of those of the P_i sub processes. The vector $\underline{x}$ represents the complete knowledge on the internal mechanism of the system and its components express variables which may be identified down to the object level.

The interconnection topology is instrumental in determining the behaviour of $\underline{x}$. The space of all values of $\underline{x}$ is referred as ***state-space*** of the system and is denoted by $\mathcal{X}$. There exist relations between the states of the system and this map is denoted by r and referred as the *internal* ***relations map***. Such a map describes the evolution of the internal behaviour and it is affected by the behaviour of the processes in the aggregate system, the topology of interconnections F' and the coupling of the internal mechanism to the environment. The vector $\underline{u}$ represents the set of all external variables which affect the objects and processes in the system through some interface and which can be arbitrarily assigned externally. This vector is the input vector and the space of all values of $\underline{u}$ will be denoted by U and referred as ***input space***. Similarly, $\underline{y}$ represents all measurements performed on objects of the system and will be called ***output vector***. The space of all values of $\underline{y}$ is denoted by Y and is referred as ***output space***. The coupling of U to $\mathcal{X}$ and $\mathcal{X}$ to Y is expressed respectively and referred as ***input-, measurement-maps*** distinctively. The nature and properties of g,h express part of the interaction of internal mechanism and the environment and manifest the desire of the system designer to control, influence the system behaviour, as well as measure it.

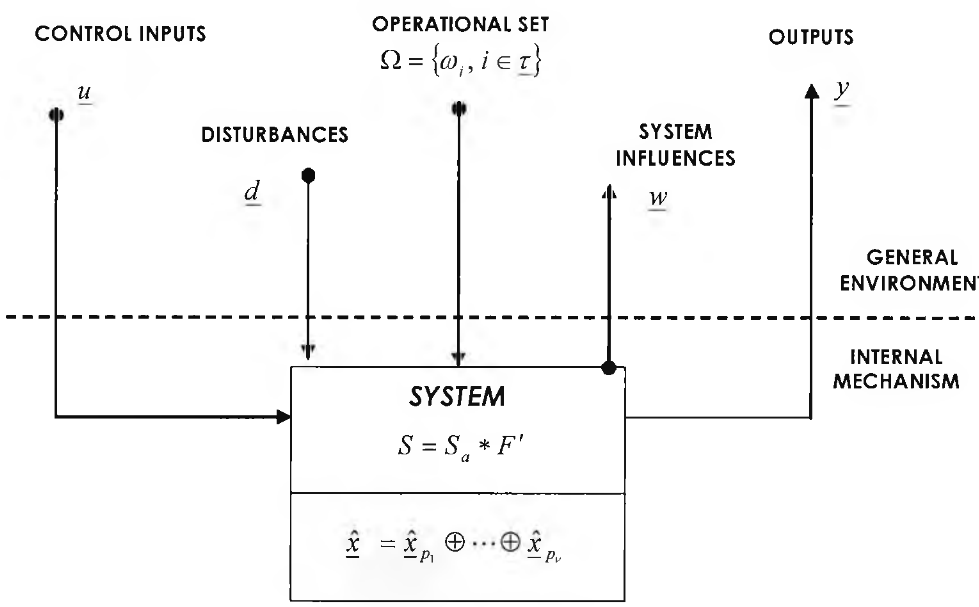


Figure 3.6: The System and its Environment

The interaction of internal mechanism with the environment has also two other signal, event components. The external vector $\underline{d}$ represents inputs, which affect the behaviour of the system states. However, the values of the components of $\underline{d}$ cannot be arbitrarily assigned but are predetermined by other external processes. These variables take values from a set of signals, events, which are generated by some process, or may be unknown. We shall use the term ***disturbances*** for such vectors and the set of all values of $\underline{d}$ is denoted by D and referred as ***disturbance space***. The coupling of disturbances to the internal mechanism is denoted by relations, or map δ and it is referred as ***disturbance map.*** Note that the disturbances generated by known processes express the embedding of the given system in a wider context of interconnected systems and they will be referred as ***loading disturbances***.

The variables that are measured express the knowledge extracted from the system. There exists, however, a vector of variables, which may be measured, or not measured and which affect other systems (they may become loading disturbances for other systems). These are denoted by a vector $\underline{w}$, which will be referred as ***influence vector*** and the set of values is a space denoted by W and called the ***space of system influences***. The coupling of W to $\mathcal{X}$ is expressed by a relation, map μ and it is called the ***influence map***. Determining the nature of the map μ requires knowledge of the system, as well as of the environment within which the system is embedded.

A set of externally imposed rules, which represents objectives, performance indices, constraints, operational instructions is represented by $\Omega = \{\omega_i, i \in \underline{\tau}\}$ and will be referred as the ***operational set*** of the system. This represents higher-level functionalities, which affect the system behaviour, but not in a direct signal, or event way. The functionality of the system, as this is represented by higher-level goals, crucially depends on the nature of Ω set. The set Ω may contain rules, which affect the behaviour of individual processes, may alter the topology of interconnection to guarantee an alternative operational scheme, or may change to objectives, goals of the system operation. This set is linked to the lifecycle aspects of the system and its elements and their functionality are defined at a higher level. In general Ω may be seen as the goal-setting governor of the system, which introduces the lifecycle aspects and impose alternative operational modes, goals and stimulate needs for redesigning the system. An overview of this system is given in Figure 3.7.

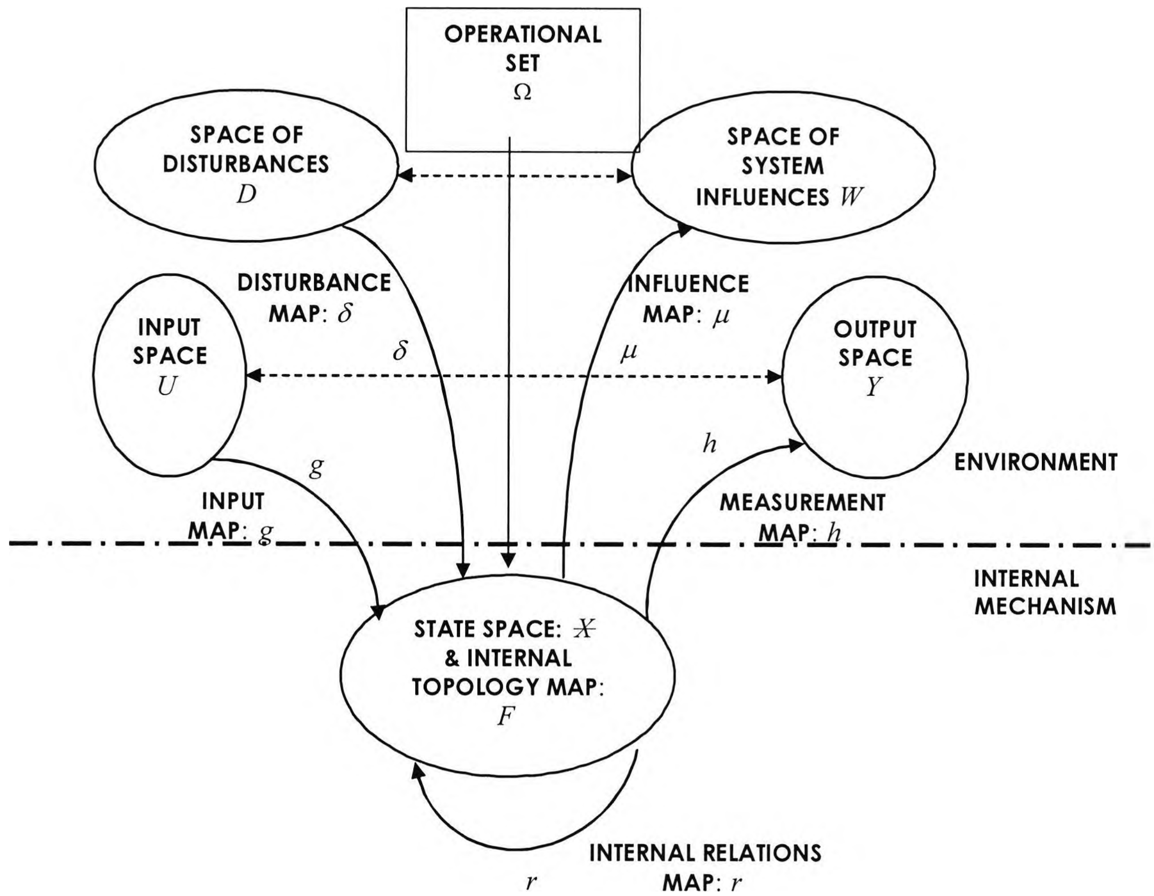


Figure 3.7: System Overview

3.11 System Behaviour, Dynamics and Qualitative System Properties

We consider the simplified system description of Figure (3.7), where we restrict our attention to the spaces $(U, \mathbb{X}, Y)$ and the corresponding maps $g: U \to \mathbb{X}$, $r: \mathbb{X} \to \mathbb{X}$, $h: Y \to \mathbb{X}$. Our interest here is to examine a richer subclass of the general family of systems, which is characterised by properties of the behaviour of the states, inputs and outputs associated with the system. For a given variable of an object, x, its values evolve as a function of time and/or events. The time and possibly event evolution of a variable has been defined as the ***behaviour*** of the variable. Using properties of the behaviour, considered for the totality of variables, enables the introduction of important system properties, which in turn allow the introduction of richer classifications of systems.

The behaviour of a variable is the set of all possible values of the variable obtained under a given initial value and expresses the time evolution of the values of the variables. The behaviour of the implicit states of an object is referred as ***implicit behaviour***, or ***implicit trajectory***. For relational objects the function f_B for which

$$f_B(\underline{\xi}_B) = 0 \qquad (3.1)$$

expresses a relation that holds true for all implicit behaviours that may be associated with the object. An orientation on the relational object associated with the partitioning $\underline{\xi}_B = \left[\underline{y}_B^t; \underline{x}_B^t; \underline{u}_B^t\right]^t$ may lead to reduction of (3.1) into

$$r_B(\underline{x}_B, \underline{u}_B) = 0 \qquad (3.2)$$

$$\underline{y}_B = h_B(\underline{x}_B, \underline{u}_B) \qquad (3.3)$$

Such orientations will be referred as ***solvable*** and for them the notion of ***state***, ***input*** **and** ***output trajectory*** are introduced as partitionings of implicit trajectory. If conditions (3.2) and (3.3) can be solved and produce the expressions

$$\underline{x}_B = \rho_B(\underline{u}_B, \underline{x}_B(0)) \qquad (3.4)$$

$$\underline{y}_B = \sigma_B(\underline{u}_B, \underline{x}_B(0)) \qquad (3.5)$$

where ρ_B, σ_B are uniquely defined and $\underline{x}_B(0)$ denotes the value of $\underline{x}_B$ at some initial time, then the solvable orientation will be called ***regular***. The relation ρ_B will be called ***state transition*** and σ_B as the ***output transition***, or $\underline{x}_B(0)$***-transfer function***.

In the following, we shall consider objects or systems for which there exists at least a regular solvable orientation. Such systems may be classified as:

An object B, or system Σ with a regular solvable partitioning, will be called ***dynamic*** *if relations (3.4), (3.5) contain explicitly $\underline{x}_B(0)$; otherwise, they will be called* ***static****.*

The meaning of a dynamical object, or system is that the time evolution of the trajectories $\underline{x}_B$, $\underline{y}_B$ are dependent not only on the $\hat{\underline{u}}_B$ trajectory, but also on the past history of the object state, as this is expressed by $\underline{x}_B(0)$. On the other hand, static objects, or systems are characterised by relations (3.4), (3.5), which are independent from the initial condition $\hat{\underline{x}}_B(0)$ and thus they are of instantaneous nature. With dynamical objects, systems, a crucial notion is that of equilibrium, which is defined below:

For a dynamic object B, or system Σ, a pair of constant vectors $(\underline{x}^*_B, \underline{u}^*_B)$ defines an ***equilibrium****, if for $\underline{x}_B(0) = \underline{x}^*_B$*

$$\underline{x}^*_B = \rho_B(\underline{u}^*_B, \underline{x}^*_B) \qquad (3.6)$$

Furthermore, if $\underline{u}^*_B = \underline{0}$, the equilibrium will be called free, and if $\underline{u}^*_B \neq \underline{0}$ it will be called forced.

The above definition characterises the equilibria as the fixed points of the state transition. This implies that equilibria correspond to pairs $(\underline{x}^*_B, \underline{u}^*_B)$ which when they are considered as inputs to state transition result in no movement at all. In the following, the trajectory that results from zero input, i.e., $\underline{u}_B = \underline{0}$, will be called ***free motion***, whereas when $\underline{u}_B \neq \underline{0}$, the resulting motion as described by (3.15), (3.16) will be called ***forced motion.*** An important qualitative property of trajectories with respect to equilibria is that of stability, which characterises the behaviour of trajectories with respect to an equilibrium point. Introducing this notion requires the introduction of some topology on the spaces $\mathcal{X}, U, Y$, which may be general sets. Boundedness is reduced to a distance problem for each of the variables with a domain of values that may be a general set (signals, sequences, events, general sets).

Similarly, defining regions for a point requires the definition of a distance function. Objects, systems, such that a metric topology may be defined on $\mathcal{X}, U, Y$ will be called ***metric*** objects, systems. For such systems, we define:

For a metric object B, or system Σ with a static equilibrium point $\underline{x}_B^*$ we consider two spheres centred at $\underline{x}_B^*$ with radii r, R, $\Omega(\underline{x}_B^*, r)$, $\Omega(\underline{x}_B^*, R)$ such that $r < R$. We may classify equilibria in the following way:

① *$\underline{x}_B^*$ will be called **state bounded**, if for any $\underline{x}_B(0) \in \Omega(\underline{x}_B^*, r)$ the free trajectory $\underline{x}_B = \rho_B(\underline{0}, \underline{x}_B(0)) \in \Omega(\underline{x}_B^*, R)$ for all time.*

② *$\underline{x}_B^*$ will be called **asymptotically stable** if it is state-bounded and $\underline{x}_B = \rho_B(\underline{0}, \underline{x}_B(0)) \to \underline{x}_B^*$ as $t \to \infty$.*

③ *$\underline{x}_B^*$ will be called **unstable** if no matter how small r is selected, there exists at least one $\underline{x}_B(0) \in \Omega(\underline{x}_B^*, r)$ such that for some time $t > \tau$ the free trajectory $\underline{x}_B = \rho_B(\underline{0}, \underline{x}_B(0))$ escapes $\Omega(\underline{x}_B^*, R)$.*

The above notions are expressions of the standard definitions of internal stability of dynamical systems expressing notions of Lyapunov stability, instability. In this more general set up, however, the selection of appropriate metric topology is crucial in defining the notions. Such topologies have to be natural and be linked to the specific characteristics of the object under consideration. The definition above may be readily extended to a characterisation of stability of forced motion when $\underline{u}_B$ is a fixed input and we consider variations in the initial conditions $\underline{x}_B(0)$. An alternative notion of stability based on input-output properties is defined below:

For a metric object B, or system Σ, we may define alternative notions of stability as shown below:

① *The system is **Bounded-Input, Bounded-Output stable**, or simply **BIBO-stable**, if for all bounded inputs $\underline{u}_B$ and for $\underline{x}_B(0) = \underline{0}$, the forced output trajectory $\underline{y}_B = \sigma_B(\underline{u}_B, \underline{0})$ is bounded.*

② *The system is **totally stable**, or simply **T-stable**, if for any bounded input* $\underline{u}_B$ *and any* $\underline{x}_B(0)$ *bounded (within a given set), the state and output trajectories* $\rho_B(\underline{u}_B, \underline{x}_B(0))$, $\sigma_B(\underline{u}_B, \underline{x}_B(0))$ *are bounded.*

The characterisation of such properties depends on the nature of the system and the selected metric topology. Different types of criteria may be derived for classes of models representing families of systems. The notions of stability have been presented in an abstract way aiming to cover all families of metric objects. Characterising such properties in terms of criteria requires use of models for the different types of systems and it is beyond the scope of this work.

For families of solvable partitionings of (3.4), (3.5) type, it may be possible to eliminate $\underline{x}_B$ from (3.4) and derive a uniquely defined relationship between $\underline{y}_B$ and $\underline{u}_B$, that does not involve $\underline{x}_B$. Then (3.4), (3.5) may be represented in an equivalent manner as

$$r_B(\underline{x}_B, \underline{u}_B) = 0 \qquad (3.7)$$

$$\underline{y}_B = \Phi_B(\underline{u}_B) \qquad (3.8)$$

and this description has been referred as *strongly oriented* in the input-output sense and Φ_B is defined as a *transfer relation*. The description (3.8) on its own does not necessarily provide a complete representation of the object behaviour. We may classify such descriptions in the following way:

*A strongly oriented description will be called **complete**, if there is a procedure of reconstructing the relationship (3.7) from (3.8). Otherwise, it will be called **incomplete**.*

Completeness thus refers to the ability to recover a relationship between internal variables, states, from input-output or transfer relationship. Assessment of presence of such property requires use of specific features of the particular objects. For complete objects Φ_B is adequate to describe the object in the input-output sense. Objects for which the initial state

$\underline{x}_B(0) = \underline{0}$ are referred as ***relaxed***. An important feature of the dynamic behaviour that can be discussed in terms of the transfer relation is defined below:

*Consider the strongly oriented object, or system represented by the transfer relation Φ_B. It will be called **causal** or **non-anticipatory**, if the output of the system at time t does not depend on the input applied after time t; it depends only on the input applied before and at time t. Otherwise, it will be called **non-causal** or **anticipatory**.*

Causality, in short, implies that the past affects the future, but not conversely. Hence, if a relaxed object is causal, its transfer relation can be written as

$$\underline{y}_B(t) = \Phi_B(\underline{u}_B(-\infty, t)), \quad t \in (-\infty, \infty) \tag{3.9}$$

The output of a non-causal system depends not only on the past input, but also on the future value of the input. This implies that a non-causal system is able to predict the input that will be applied in the future. For real physical systems, this is impossible. However, for processes involving human operators, or some form of intelligence, non-causality may be a naturally observed property.

Two important properties related to the family of state, output and input trajectories in relation to the spaces $\mathfrak{X}, Y$ and U, are those expressing ability to transfer the object state between two points of $\mathfrak{X}$ by some appropriate input, and the ability to reconstruct the initial state of the object by knowledge of the input and output trajectory. These properties are defined below.

Consider a dynamic object B, or system Σ, with state, output trajectory families as in (3.4), (3.5) defined for all possible inputs $\underline{u}_B$. We define:

① *The object or system as **reachable**, if given any two points $\underline{x}_B^1, \underline{x}_B^2 \in \mathfrak{X}$, there exists an input trajectory $\underline{u}_B \in U$ such that*

$$\underline{x}_B^2 = \rho_B(\underline{u}_B, \underline{x}_B(0) = \underline{x}_B^1)$$

and this occurs in finite time. If there exists a pair $(\underline{x}_B^1, \underline{x}_B^2)$ *for which this property does not hold true, then the system will be called* ***non-reachable****.*

② *The object or system is* ***reconstructable*** *if knowledge of the input and output trajectories over a finite time allows the reconstruction of the state trajectory and thus, also the initial state of the system; otherwise, the system is called* ***unreconstructable****.*

The characterisation of these properties in terms of specific criteria requires the consideration of particular classes represented by specific families of models.

3.12 Final thoughts

The work here uses the existing methodologies, but aims at redefining notions, concepts and introduces new ones reflecting the needs of the new paradigms. In this thesis, the emphasis is on introducing a conceptual framework, rather than developing a formal mathematical set up that supports it. The development of a relevant mathematical framework is a future task that can be undertaken when the relevant concepts have been appropriately defined. The current approach may be seen as a hybrid of the 'Soft Systems' [Check], [Check. & Scho], [Klir,2] and the 'Abstract Systems' approaches [Mes,1], [Gin &Gog,1], [Tak & Takah,1,2]. Soft Systems approaches are flexible, use heuristic and linguistic expressions and they are close to the formulation of problems which are difficult to describe. However, they are using vague statements, lack rigor and they do not provide formal tools for analysis. Abstract Systems, on the other hand have rigor in the description of concepts and properties and they are linked to formal methods which enable the systematic study of analysis and design/synthesis problems. However, they are not very adaptable, frequently sacrifice flexibility for the sake of rigor and the formalism becomes 'esoteric' and difficult to use. The aim here is to develop an approach that combines the advantages of both, whereas at the same time avoids their disadvantages. We are heavily influenced by the major paradigm of Dynamical Systems and by the range of their applications to conventional and non-conventional engineering applications. However, we also aim at expanding the framework by generalising the fundamental notions and concepts involved to suit other paradigms of interest and where it is needed to introduce additional new concepts and notions to serve the new paradigms.

Chapter 4

Modelling and the System notion

In the previous chapter fundamental descriptions and definitions were given related to the notion of the *system, its parts* and *its environment.*

Modelling, as will be seen, is strongly connected to how the modeller conceives the "world", and by "world" it is meant the union of the system and the environment, independently of how small or large that union is. This action of trying to observe systems, comprehend their behaviour and then describe and communicate this knowledge expresses the art of modelling. The model is the interpretation of the system as seen by the modeller-observer. The dependence of modelling on the observer, however, does not negate the existence of a concrete process of modelling, characterised by formal stages and steps. Some of the steps may sometimes be merged, depending on the area, or even non-existent, because they are thought self explanatory and self-evident.

Trying to understand the process of modelling, there are issues to consider before the actual activity of the modelling begins and this involves clarifying issues such as:

- What is a model?
- How do you actually begin to model?

- What are the incentives to create a model?
- What are the first things you have to consider in the process?
- How detailed or how simplified does the model have to be in order to satisfy specifications and what are those specifications if you are trying to create a general systems framework?
- How does the purpose of modelling affect modelling?
- How does ones personal knowledge of things affect the modeling process?

All of the questions above might seem naïve until the moment of starting the actual modelling process, where a combination of brainstorming, creative thinking and structure need to be combined to create a model.

This chapter will try to give a structure to the process of modelling and the "pre-modelling" stages, with an open mind, but always treasuring the central role of the idea of the system. First, a general discussion regarding scientific methods is given and then an introduction to general concepts and issues concerning modelling and how modelling relates to systems. The importance of the observer in the modelling process and his or hers irreplaceability is emphasised. The issue of mapping the system to the model will emerge, along with how data mining and knowledge management help in that process and the further stages of modelling. A review of some soft approaches to problem solving will take place and finally the attention will be focused on the conceptual modelling; the definition and its role in the subsequent stages of modelling.

4.1 Background on scientific method and the act of modelling

Scientific theories are distinguished from non-scientific simply by the fact that they rely on evidence, experiments, data and they provide the means to test them. They can be criticised and are open to change in the light of criticism and new emerging evidence. Now, on the basis of these criticisms, most developed working *methods* for discovering the truth, range from mere description of phenomena through observation and both qualitative and quantitative experiments, to the creation of models.

The act of modelling has to be seen within the framework of the scientific method. In fact, scientific methods are intimately linked to modelling and modelling emerges as the result of such methods. The usefulness of the term "scientific" is evidently linked with the effort to distinguish scientific knowledge from other types of knowledge, such as historical knowledge or empirical knowledge. The definition of scientific methods involves the following essential features [Rus]:

1. *Their statements are not about concrete objects, but about specific theoretical entities.* For example, Euclidean geometry makes statements about angles, thermodynamics about temperature or entropy of a system, but in nature there are no angles, temperature or entropy.
2. *The theory has a rigorously deductive structure;* it consists of a few fundamental statements (called axioms, principles) about its own theoretical entities and it gives a unified and universally accepted means for deducing from them an infinite numbers of consequences. In other words, the theory provides general methods for solving an unlimited number of problems.
3. *Applications to the real world are based on correspondence rules between the entities of the theory and concrete objects.* Unlike the internal assertions of the theory, the correspondence rules carry no absolute guarantee. The fundamental method for checking their validity-which is the applicability of the theory-, is the experimental method. In any case, the range of validity of correspondence rules is always limited.

The immense usefulness of science consists in providing models of the real world within which there is a guaranteed method for telling false statements from true. Whereas natural philosophy failed in the goal of producing absolutely true statements about the world, science succeeds in guaranteeing the truth of its own assertions. Such models, of course, allow one to describe and predict natural phenomena, by translating them to the theoretical level via correspondence rules, then solving the "exercises" that are obtained and translating the solutions obtained back to the real world. There is, however, another possibility: moving freely within the theory and so reaching points not associated to anything concrete by the corresponding rules. From such point in the theoretical model one can often construct the corresponding reality, thus enhancing the knowledge for the existing world [Rus] (See Figure 4.1). Scientific theories even if created for the purpose of describing natural phenomena are able to enlarge themselves by means of the deductive method. As a consequence they provide means for the development of models of many areas of human activity. *Scientific methodology* is intrinsically connected to the methodological structure of science.

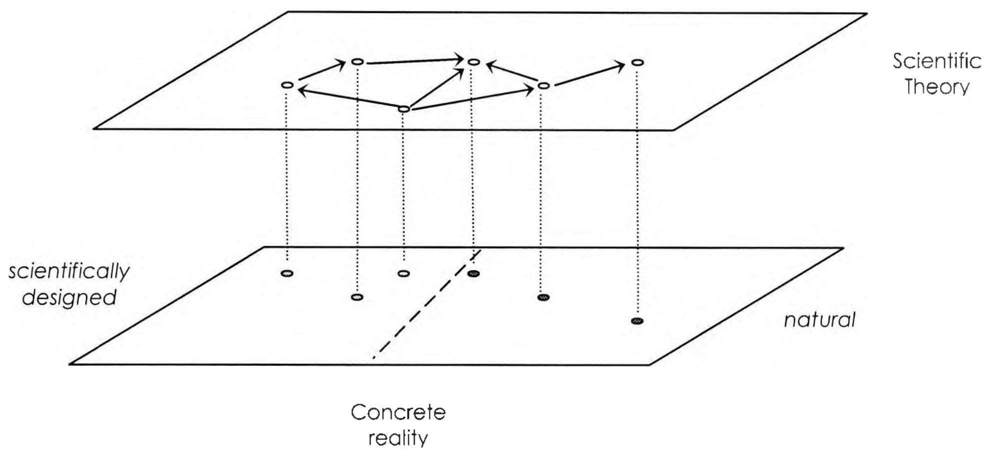


Figure 4.1: The role of scientific methodology [Rus].

In the figure above dark shaded circles on the reality level (lower plane) represent objects from nature. Their counterparts on the theoretical level (upper plane) are linked via logical deductions (arrows) to many other constructs, which may or may not have concrete counterparts. Some of these theoretical constructs give rise, via correspondence rules (dashed lines) to new concrete objects (lightly shaded circles on the lower plane) [Rus]. The effort, therefore, is concentrated in using tried and tested concepts and approaches developed within various fields (this thesis uses the engineering field), to more general problems not identified with the specific nature of those various fields. It is important to understand that scientific "truth" descends from observations and experiments, which provide the starting point for the construction of a theory which hopefully corresponds, interprets and unifies those observations. The theory itself is considered as an instrument to handle a system and the concept of the model is closely related to the theory and can be considered a link between the theory and the reality on which it is based.

4.2 Systems and Modelling: Concepts, Issues and Relations

At this point a description of issues related to modelling, systems and systems thinking will be made, from the very abstract to the very specific level and the connections between the system and its models will be made explicit. Modeling has been the focus of human attention from the early years of human effort to understand the world and has been considered in various areas. Numerous ideas, methods as well as determined efforts to show an understanding of the subject, led to arguments about the philosophical side of modeling first, as well as the scientific and practical side of it. Questions such as:

- How can you model a system?
- Are there any formal steps in the modelling process?
- Why are we modelling a system?
- What are the preconditions for initiating modelling?
- What criteria will judge success/lack of success, completion?

In our effort to understand the various issues of modelling, all the different conflicting ideas regarding the subject made it difficult to draw a conclusion on a framework that would ensure a formal methodology for modelling. As a result, a more open minded approach towards the relevant theories and methodologies-if any-was kept. The purpose was to determine how other areas deal with the idea of modelling, and try to find common ground. This does not necessarily mean that the different ideas will be used in our effort to create a modelling framework; on the other hand they will be taken into consideration. Such a decision, makes the end result that is the modelling framework that will be created, have more chances to be concise and practical, since it will incorporate knowledge from various disciplines.

4.2.1 Model Definition and issues related

4.2.1.1 The concept of a Model

A model is a representation of a system, and thus provides an abstraction of the reality, tries to mimic the system that it represents and can give a grouping of snapshots of the dynamic system behaviour. A model, as a representation of the system, is linked to the theory used to interpret and capture the knowledge about it. Thus to develop a model requires a theory and tools that would include verbal or symbolic description, descriptive sketches or drawings or representational physical constructions to express this knowledge [Brown]. A more useable concept of a model is that of an abstraction, of a real or mental system, involving key variables and relationships. These are abstracted in order to simplify the problem itself. Modeling allows the user the means to better understand a problem and provides tools for manipulating the description in order to analyze the results of various inputs by subjecting them to a changing set of assumptions.

A model is a theoretical construct that may represent physical, engineering biological, social and other systems, with a set of variables and a set of logical and quantitative relationships between them. As a theoretic construct it fits the known, available facts into a neat, purposeful package. Models in this sense are constructed to enable reasoning within an *idealized* logical framework about these processes and are an important component of scientific theories. Idealized here means that the model may make explicit assumptions that are known to be false in some detail. Such assumptions may be justified on the grounds that they simplify the model while, at the same time, allowing the derivation of acceptably accurate solutions [Wik]. By representing the reality with something that can be interpreted with mathematics, a modeling language or even a drawing, the user can begin to understand how the system works and try to manipulate it. This is so broad that it includes not only scientific models but also all sorts of everyday descriptions and representations and many artistic activities. In a scientific context, the word 'model' is used much more narrowly than this. A *scientific* model is one that deliberately sets out to represent particular, detectable properties of something under describable conditions in an unambiguous and publicly accessible way [Brown].

Classical science seeks to find "laws" as the ultimate objective, in regard to the prescriptions it provides to the modeler. So the modeler is instructed to write down the laws (physical or logical) that are relevant and govern the system. Then he/she must obtain "parameters" with the aid of which these laws are combined into equations and eventually fed to computers which will then presumably provide the answers to the original questions.

In fact the success of simulation is very much dependent on the quality of prior information, in other words, the scientific accuracy of the laws which have been put into the model [Kalm 2]. In modern modelling the difference is simply that modelers do not have to rely exclusively on prior laws-but can make use of them-because the prescription is simple: DATA→Model. It is not intuitively obvious how this can be done, and a major part of contemporary system theory is concerned precisely with this question, where it is technically known as *realization theory*. The essential point about realization theory is that it replaces intuitive requirements such as we have listed above by a much more precise formulation so that the problem is put on higher scientific level. There are really two research directions [Kalm 2]:

- Study a class of models (which are candidates for concrete models) in such depth that most general questions arising in modelling can be answered for the whole class.
- Develop methods for representing given data by a specific model from a class.

A central issue in the construction of a model is its “purpose”. Since models are abstractions of reality, the purpose serves as the crucial notion that defines the projection of the system to smaller spaces and thus decides which features of the system are preserved and which are either neglected or represented in an aggregate or simplified way.

Figure 4.2 represents the projection of reality through the act of modelling in the case of representation of a geometrical shape and it is a graphical way to demonstrate that modelling only captures certain aspects of the system, depending on the specified purpose and ignores a number of other features. Every projection changes depending on the point of view that is used to describe the system. Any model is an imitation, a projection of the real world, based on the modeler’s problem area of interest. In this simplified version of the real world, the model brings out certain characteristic features in the object of study and at the same time excludes others. Since the model is just an effort to represent the system, many factors have been left out, either because there were not needed for the particular model, or because there were just not taken into consideration, as they were not useful for the purpose of the model. The notion of a perfect model is a fiction and all models are functions of the purpose and the capabilities of the modeler. The notion of model uncertainty is natural and should always be considered as part of the modelling exercise. Description of uncertainty is linked to the assumption we have made and the nature of experiments used to generate the data and extraction of knowledge.

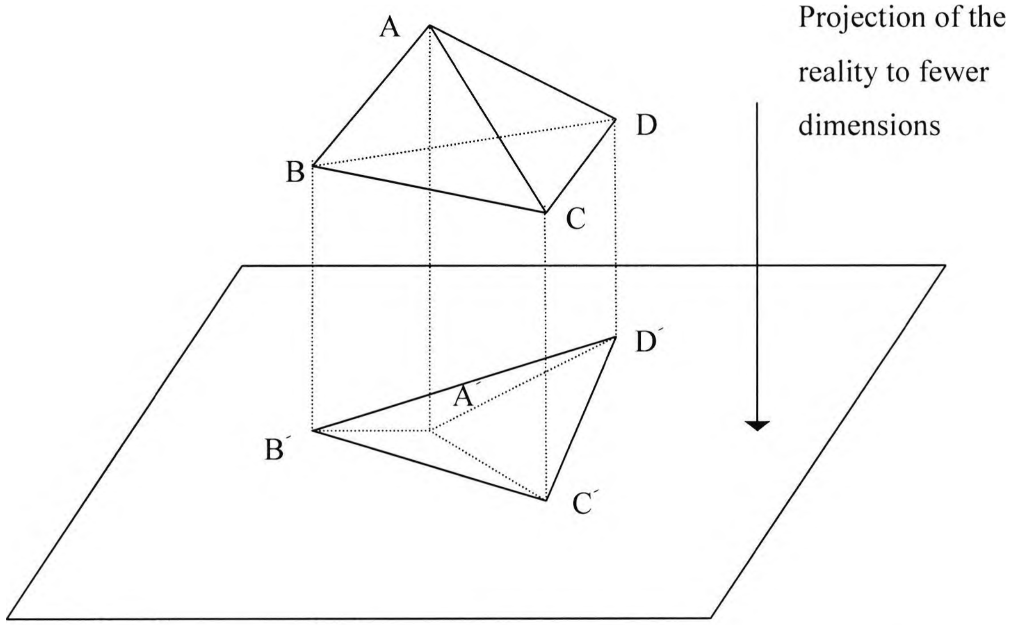
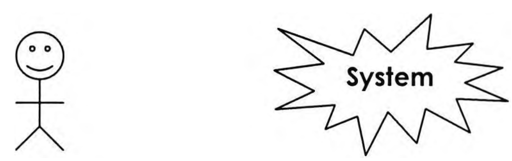


Figure 4.2: Projection of reality to fewer dimensions

In our concrete geometric example the "system" will remain a pyramid but the projection-model can take various forms. Typically a model will refer only to some aspects of the system in question, and two models of the same system may be fundamentally different. This may be due to differing requirements of the model's end users or to conceptual or esthetic differences by the modellers, difference in skills and decisions made during the modeling process. Esthetic considerations that may influence the structure of a model might be the modeller's preference for a reduced ontology, preferences regarding probabilistic models vis-a-vis deterministic ones, discrete vs continuous time etc. For this reason users of a model need to understand the model's original purpose and the assumptions made during its construction. However, the quality of the model, as well as its significance, can only be judged against its purpose. That is, the aspects of the system that need to be captured, in the form of a model, and the way they are captured by it-the detail, the focus, the functionality, etc-will eventually be the specifics that will judge whether a model is a good representation of the system.

4.1.2.2 The definition of the Model

The process of modelling starts with the observer and the system



The observer associates with the system:

1. A set of data generated by input-output type observations
2. A set of object relations stemming from past knowledge, experience (knowledge extracted from past observations)
3. A mixture of data and object relations.

As such the two distinct approaches that emerge in trying to understand the system and thus model it are:

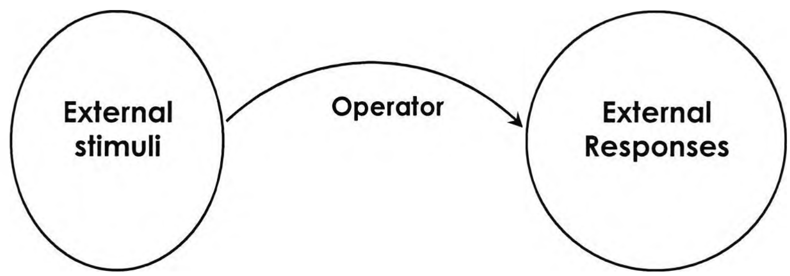
(a) *External or input-output approach* where the system is treated as a "black box" generating responses and data to input stimuli.

(b) *Internal description* where some knowledge of the objects-relations insight the "black box" is assumed, in addition to stimuli-response information.

The fundamental difference between the two views is that the first considers the system as an operator acting on data/signals only, whereas the second elaborates this view by enlarging the domain and co domain and introducing some structure on the corresponding operator by the specification of object/relations description. The above view of the internal description does not assume orientation, that is, separation of variables to input and outputs and it is thus an implicit view of the system.

The emerging two alternative views are these described as:

Input-Output system view



Internal system view

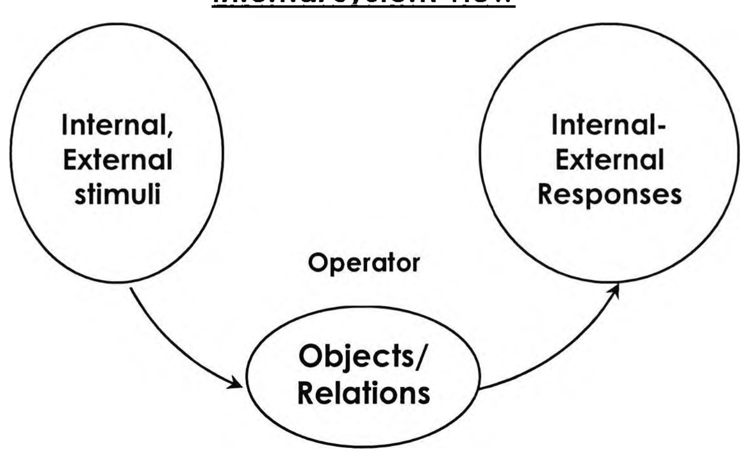


Figure 4.3: The two distinct approaches to understanding the system.

Given that the basic components of the system are:

(i) Stimuli, responses, data

(ii) Objects, relationships.

any modelling activity has to deal with the above basic building blocks of the system notion. The rules for interconnecting the elements into a system is given to the modeller, by previous knowledge. The physical or logic rules governing the elements, and their quantitative description are known or assumed to be known. With such data the user builds the model, draws conclusions concerning the system by performing experiments-observations by deploying the model and testing its validity [Kalm].

The adopted view of the system here is based on some form of knowledge for it. Thus viewing the system as a set of stimuli-responses-data-objects-relations set, gives to modelling a specific character and task which is the definition of an "operator" that interprets the link of stimuli-responses subject to the constraints provided by the objects/relations (a priori knowledge). The search for such operators is the main task of modelling. Given that there may be more than one operator describing a system (according to the given information), what is important is the definition of operators that interpret the given evidence and "nothing else". This expresses a notion of minimality and in Popper's language it is referred to as the "minimal unfalsified model". Kalman [Kalm] claims that given any precisely describable behaviour, it is possible to construct an underlying model which explains the behaviour minimally, that is, the behaviour of the model is identical with the given behaviour of the real system and the model cannot be simplified further without any loss of properties. The existence of such a model always depends on the given set:

$$\mathbf{M_s=\{stimuli;\ responses;\ data;\ objects;\ relations\}}$$

which will be referred to as a ***modelling set*** for the system. Such a statement expresses the problem of realization, which however has to be addressed within certain families of behaviour and classes of models.

Although the modelling process assumes the existence of basic components (stimuli, responses, data, objects, relations), there are some fundamental questions relating to the nature of these building blocks which have some key role in the modelling exercise and some of them have a philosophical character. Thus, relevant questions to ask are:

- (MQ1)How do we acquire information and then knowledge from sets of data associated with a system?
- (MQ2)How do we acquire the knowledge on objects, relations on a system and how is this related to the previous question?

- (MQ3)How do we select appropriate stimuli to excite responses that are "rich", as far as providing knowledge on the system?
- (MQ4)How do we define a model that fits the existing data and objects/relations knowledge?
- (MQ5)How do we guarantee model minimality on the basis of the given modelling set M_s?
- (MQ6)How do we expand the range of a given modelling set M_s and thus also enhance the range of corresponding models?
- (MQ7)How do we simplify a given model or derive focused submodels on aspects of its behaviour?

The above range of problems is critical in the overall modelling exercise and each one of them has many more sub problems that require attention. In particular:

- **(MQ1)** is the classical problem of extracting information from data and then creating knowledge from information. Techniques for Data Mining, Pattern Recognition methodologies and in general approaches for "Knowledge Discovery" are important contributors as far as developing a systems approach to such a difficult and open problem. We may refer to MQ1 as the ***Knowledge Extraction Problem*** **(KEP)**.
- **(MQ2)** refers to the problem of identifying objects, components, in a system as well as relations between them. Such a process is the result of familiarization of the observer with the system which is the result of a long process of observations, interactions with the system, use of previous knowledge and analogies. The results depend clearly on the observer and his/her experience and knowledge and the way the engagement with the system has taken place. We may refer to MQ2 as a ***Systems Conceptualisation Problem*** **(SCP)**.
- **(MQ3)** deals with the problem of defining appropriate ways of exciting the system to get responses, which are rich enough to provide useful knowledge on the system and its parts. This involves designing a sequence of experiment-observations which are independent from each other and each one of them highlights a different aspect

of the system features. Clearly, such a design of experiments-observations requires some form of knowledge about the system and it is thus linked to SCP. We shall refer to MQ3 as ***Design of Experiments Problem*** **(DEP)**.

- **(MQ4)** is the mainstream modelling question which deals with the definition of an appropriate machinery that can be either conceptual or formal and describes the given Modelling set. Within this area we have the problems of physical modelling, system identification, hybrid versions between the two, etc. Usually we assume certain abstract classes parametrised by generic features (linked to SCP) and then we use heuristic or mathematical formulations to select the free parameters. This problem is referred to as the ***Model Construction Problem*** **(MCP)**.
- **(MQ5)** deals with the characterization of minimality for a model that corresponds to M_s. This is a problem at the level of the class of models considered and its study involves the parametrisation of all available solutions and then development of criteria that permit the testing of minimality. We may refer to this as the ***Model Minimality Problem*** **(MMP)**.
- **(MQ6)** deals with the problem of expanding the modelling set M_s to a larger M_s' that contains M_s. This involves study of a number of DEPs, guaranteeing their independence from the experiments generating M_s and then finding ways to expand the model by solving MCPs under some initial constraints (unfalsified model for M_s). Such problems may be referred to as ***Model Expansion Problems*** **(MEP)**.
- **(MQ7)** contains a number of particular problems aiming to simplify the derived model associated with a set M_s ($M_s \rightarrow M$). Simplification may occur in two alternative ways as it is suggested by the following diagram:

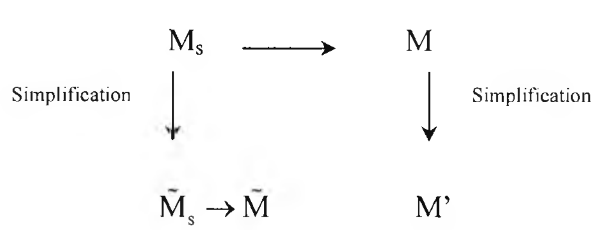


Figure 4.4: Derivation of simplified models

In the above diagram the derivation of an approximation (simplification) of the model M to the model M' is referred to as ***Model Reduction Problem*** **(MPR)** and involves the derivation techniques that allow the approximation of M within the same class to a simpler model M'. The second class of problems deals with the simplification of the set M_s itself and not of the corresponding model. This simplification may occur in every element of M_s, such as stimuli, response, data and/or description of the objects and their relationships. The simplified set $\tilde{M}_s$ is affected by defining criteria that emphasize certain aspects of the system behaviour at the expense of others. As such the models $\tilde{M}$ corresponding to the simplified set $\tilde{M}_s$ may belong to a different class than that of M'. The latter class will be referred to as ***Model Simplification Problem*** **(MSP)**.

Developing modelling approaches requires tackling problems of the above classes in a rather concrete form. In summary, a scientific formalization of modelling involves the following main features [Kalm]:

- The relevant behaviour of a real system must be isolated and described in a formal mathematical way.
- A class of models must be given, constructed on the basis of laws suspected to govern the relevant system and containing sufficiently many parameters so that the behaviour of the model can be fitted to the behaviour of the system.
- The class of the models must include the behaviour in question-now interpreted as a model- so that the modeller is assured a priori of a realization of the behaviour within the given class of models.

4.1.2.3 Purpose and benefits of Modelling

Generally, the purpose of a model is to provide a framework for applying logic and mathematics, that can be applied for reasoning in a range of situations. In many cases, a model is created or chosen to try and best predict the probability of an outcome. Models are used primarily as a reusable tool for discovering new facts, for providing systematic logical arguments as explicatory aids, for evaluating hypotheses theoretically, and for devising experimental procedures to test them. Reasoning within models is determined by a set of logical principles, although the reasoning used may not be completely formal. They are employed to develop new knowledge, to modify existing knowledge or to obtain knowledge from systems. They provide a basis for simulation and testing as a counterpart to deductive arguing. They can also be used to interpret behaviour of the system, to predict the outcome of actions, future behaviour or to devise decision making/control strategies that can be applied on the real system.

The goal of modeling is to adequately portray realistic phenomena. Once developed, a plethora of information can be available about the real-life counterpart by manipulating a model's variables and observing the results. Real-world decisions involve an overwhelming amount of detail, much of which may be irrelevant for a particular problem. Models allow the elimination of unimportant details so that the user can concentrate on the relevant decision variables that are present in a situation.

Models provide the most effective means developed for predicting performance. It is hard to conceive a prediction for a system, if a model of that system is not first produced. To construct a model of a real process or system, careful consideration of the system elements that must be abstracted is required. This in itself is a profitable activity, for it develops insights into the problem. When building a model, the modeler is immediately struck with the magnitude of ignorance; ignorance not necessarily for the subject being modeled, but for the amount of information available for that subject. What does the modeler really know? Where are the gaps in available data? It is often impractical or impossible to manipulate the real world system in order to determine the effect of certain variables. An example of how important is the use of models can be seen in business systems; to use any business system itself as a laboratory could be disastrous and very costly. The dangers in using models lie in the possibility of oversimplifying problems to keep models in a form that can be handled more easily. To summarise, the purpose of modelling comes as a threefold:

- Deductive reasoning. Using the model as a platform, properties of the system are deduced, behaviour is predicted and further analysis, design and decision making can be carried out.

- Simulation. Simulation is the counterpart of deductive reasoning. The model is used and through experiments (simulations), knowledge is deduced. This knowledge is based on the model but it is not deductive; it is the outcome of the analysis of the results obtained from the experiments made on the model.
- Emulation. Emulation is a bridge that brings the system and the model together, so that knowledge can be gathered from the interaction of the two. An example would be a simulator/emulator of an airplane; a model of the airplane is provided and given the interaction with the pilot (who is part of the real system) information can be gathered that is useful in the real life situation of flying a plane.

It is important for the model user to realise that model development and model based solution are not completely separable. While the most accurate representation possible may seem desirable, the user still must be able to find a solution to the modeled problem. Model users need to remember that they are attempting to simplify complex problems so that they may be analyzed easily, quickly, and inexpensively without actually having to perform the task. Also desirable is a model that allows the user to manipulate the variables so that "what if" questions can be answered. Models come in many varieties and forms, ranging from the simple and crude to the elegant and exotic. Whatever category they are in, all models share the distinction of being simplifications of more complex realities that should, with proper use, result in a useful decision-making aid.

To summarise, some of the most important benefits of models are listed below:

- Models generally are easy to use and less expensive than dealing with the actual situation. By using models a test area is established where more data can be collected regarding the possible states of a system, its behaviour and its nature, without having any implications on the system per se. This way the modeler can analyse "what if" questions and scenarios and analyse the system in a standardized format, provided by the model
- They require users to organize and sometimes quantify information and, in the process, often indicate areas where additional information is needed. This way the

modeler becomes far more specific about objectives and the understanding of the problem becomes much clearer for the modeler.

- They provide a systematic approach to problem solving.
- They serve as consistent tools for evaluation.
- They provide a basis for simulation.
- They enable users to bring the power of mathematics to bear on a problem.
- They provide a standardized format for analyzing a problem.

4.2.1.3 Types of Models

Models are indispensable tools for reperesenting a reality that sometimes can be too complex to be described as it is and needs to be simplified in a form of a model. Models allow the formalization of existing knowledge and may set up a mathematical framework that allows the development of deductive reasoning, analysis or experiential knowledge development in term of simulation. Some models are replicas of the physical properties (relative shape, form, and weight) of the object they represent. Others are physical models but do not have the same physical appearance as the object of their representation. A third type of model deals with symbols and numerical relationships and expressions. Each of these fits within an overall classification of four main categories:

- Physical models
- Schematic models
- Verbal models
- Conceptual models
- Mathematical models
- Data structures.

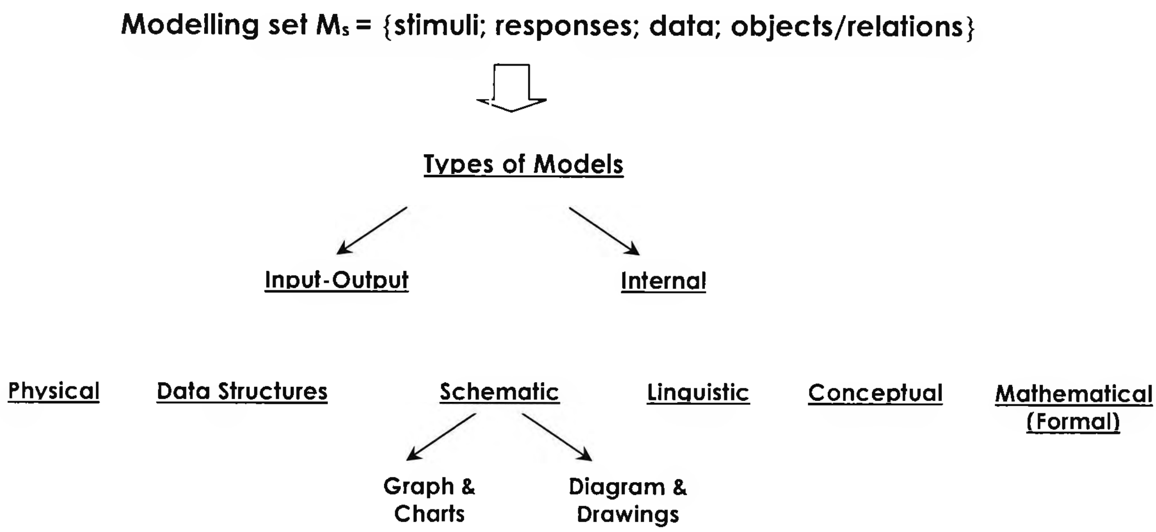


Figure 4.5: Types of models.

4.2.1.3.1 Physical Models-Abstract Functional Models

Physical models are representations of natural, or engineered (man made) systems. Their subsystems are physical processes described by physical variables; the relations between such variables are expressed in terms of physics, chemistry, biology laws. For such models the topology of interconnections is expressed as relations between the physical variables; inputs, outputs and frequently the internal variables for the subsystems are physical variables and thus constraints on their performance may be introduced in a natural way. The two main clusters of physical models that can be distinguished are:

- Models of physical world around us that is *natural models.*
- Models of Artificial systems, that is systems produced as the result of human activity and referred to as *engineered systems models.*

The fundamental difference between those two types is that in the first case the subsystems, processes are the results of a natural evolution process, whereas in the second they are results of a design process under the control of the man designer. For physical systems and associated models there are cases where although many variables can be measured, there may be variables that cannot be controlled (typical example is the case of planetary systems, where we can measure, but there is no variable that can be controlled). For engineered systems, modeling is an essential part of the design process and design now becomes an evolutionary process where the final system model is shaped successively at the different design stages. Such differences give to the modeling of those two different classes different objectives. In the first, modeling provides a representation of a system and its properties and has to serve the activities of measurement/diagnostics and control. For the second class, modeling serves the previous tasks, but it is also an instrument to describe the evolutionary process of system design. The latter dimension is new and it is referred to as *modeling for integrated design* [Karc, 1].

In the modeling of Physical Systems there are some important issues which are critical in the derivation of models where functionality of the components and key fundamental issues are preserved. These are the notions of:

- Analogy
- Duality
- Scaled models.

An *analogue* [Shear] is a model that comes from a different physical domain, but its behaviour, as far as relations between variables is quite similar to that of the original system. An analogue retains the original properties of the topology, the properties of "through" and "across" variables [MacF, 1] and simply uses the correspondence between the physical elements coming from different physical domains. Thus, for instance, a mechanical system may be represented in terms of an electrical analogue by using the correspondence:

Mechanical domain ⇔ Electrical domain

Force ⇔ Current
Velocity ⇔ Voltage
Mass ⇔ Capacitance
Spring ⇔ Inductance
Friction Element ⇔ Resistance

The advantage of analogues is to transfer methods of analysis of system models from one domain to another. Given that many disciplines develop independently from others and model analysts may come from a given domain, analogy becomes a very useful conceptual and analytic tool for the study of physical systems. The deeper essence of analogy stems from the fact that processes from different physical domains may behave in a similar way. Analogy is an instrument at the level of conceptualization, as well as the development of formal models.

The notion of *duality* [Shear] is used to produce models within the same physical domain and for which the resulting formal model has a similar mathematical structure to that of the original physical model. As such *dual models* may be useful instruments in the analysis stage based on formal models. Their basic principle for construction is the use of inversion between "through" and "across" variables (current becomes voltage and vice-versa, force becomes velocity and vice-versa etc) [Karc, 3]. As a result of such a correspondence, the interconnection topology is not preserved but it is "dualised", in the sense that series connection becomes parallel and vice-versa. The advantages of duality is at the behaviour study level, where a dual can be used to describe for instance properties related to infinite frequencies behaviour, in terms of finite frequency behaviour of the dual [Karc & Hay]. The construction of duals [Karc, 3] is not a straightforward problem and duals can be constructed as a vehicle for studying different types of properties.

A notion related to analogy is that of *Scaled Models.* These are defined as representations of the original system that retain in a scaled form certain features (such as geometry of objects), but do not necessarily retain all information on behavior and properties that may be associated with the original systems (a car model preserves the geometry of the shape and some other properties, but by no means represents the behaviour of a real car). There are advantages of working with scaled models, but there are also a number of important questions that arise as a result of scaling of the physical dimensions of a given system. By intervening to the size of subsystems, the effects on relations between the variables and the validity of physical laws are not necessarily issues which are understood well. Scaling may lead to a model that has only a rough resemblance to the original system, from the functionality of the system viewpoint.

Alternatively, some physical models may not look exactly like their object of representation but represent important aspects of reality through similarity in relations between entities expressed in forms that are easier to handle. Such models behave like the reality they represent without looking like it. An example is a mathematical graph. Another example is the use of cardboard cutouts to represent the machinery being utilized within a manufacturing facility. This allows planners to move the shapes around enough to determine an optimal plant layout. Furthermore, physical models are not necessarily scaled versions of reality, but they may be component based abstractions of reality with appropriate representation of topology (example, a graphical representation with a functional description of sub processes).

So far the discussion was led by the need to describe systems of the physical world, focusing on physical and structural aspects of the model. This however does not describe systems that are *abstract*, such as Organisational, Socio-Economic and Political. In these cases the systems involve humans and thus they do not obey the rules of Physics and therefore Physical models cannot describe such systems. The models used in these cases are called *Abstract Functional models*. These models identify functions and concepts and use relevant means (such as *Schematic models, Linguistic models*) to create a basis for understanding the system. It is worth mentioning however, that notions of analogy may be used for such systems since notions of "generalized flow" and "generalized potential" may be identified together with corresponding variables. Thus, although in the strict sense such systems do not correspond to the category of physical systems, they may be represented in terms of physical system models. A fundamental distinction has to be made however and this relates to whether we refer to the human as an individual, or humans as group/society. Physical models cannot be used for the first case, but they have some potential for the latter.

4.2.1.3.2 Schematic Models

Schematic models use graphs, charts and diagrams to represent aspects of the system at a given level of abstraction. We distinguish the structural representations provided by the graphs, which provide description of the natural topology at some level, and the chart, diagram representations of the behaviour of certain variables of the system. Schematic models are more abstract than physical models. Although the above two classes both use graphical means, they provide different forms of information. Normally general and abstract, they are often more difficult to construct but easier to use than others models. Graphs and charts are schematic models that provide pictorial representations of variables and frequently indicate mathematical relationships between them. Plotting a line on a graph indicates a mathematical linear relationship between two variables. Two such lines can meet at one exact location on a graph to indicate an equilibrium, or the break-even point, for instance. Pie charts, bar charts, and histograms can all model the behaviour (time evolution) of some variables, but really bear no physical resemblance to anything.

A schematic model reduces a state or event to a diagram or chart. A circuit diagram of an electronic amplifier, exemplifies a schematic model of the actual hardware. Diagrams, drawings, and blueprints also are versions of schematic models. These are pictorial representations of conceptual relationships. This means that the model depicts a concept such as chronology or sequence. A flow chart describing a computer program is a good example. The precedence diagrams used in project management or in assembly-line balancing show the sequence of activities that must be maintained in order to achieve a desired result.

4.2.1.3.3 Linguistic Models

Linguistic models have the characteristics of many classes and use as means of communication the language instead of schemata, maths or diagrams. They depict reality throuh the use of verbal statements that set the relationships between the concepts. In accordance with their purpose these models are prescriptive, descriptive and explanatory. Verbal models establish the meaning of both symbols and equations, if applicable, through words to metaphorically explain the meaning of the symbols and equations. Verbal models are always slightly ambiguous, inaccurate and use words to represent some object or situation that exist, or could exist, in reality. Verbal models may range from a simple word presentation of scenery described in a book to a complex business decision problem (described in words and numbers). A firm's mission statement is a model of its beliefs about what business it is in and sets the stage for the firm's determination of goals and objectives.

Verbal models are early state models where language is used as means for describing the system. Such models frequently provide the scenario necessary to indicate that a problem is present and provide all the relevant and necessary information to solve the problem, make recommendations, or at least determine feasible alternatives. Most of the cases presented in management textbooks are really verbal models that represent the workings of a business without having to take the student to the firm's actual premises. Sometimes, these verbal models provide enough information to later depict this problem in mathematical form. In other words, verbal models frequently are converted into mathematical models so that an optimal, or at least functional, solution may be found utilizing some mathematical technique. A look in any mathematics book, operations management book, or management science text generally provides some problems that appear in word form. The job of the student is to convert the word problem into a mathematical problem and seek a solution.

4.2.1.3.4 Conceptual Models

A Conceptual model will be defined in this thesis as a collection of properties of all the previously discussed models to create the basic model that provides the fundamentals on which any other model can be build. The conceptual model uses a combination of verbal, schematic and simple mathematical descriptions that best specify, describe and represent the problem at an early stage in the modeling process. In accordance with its final purpose, this model is *prescriptive*, *descriptive* and *explanatory* and contains the minimal information required to describe the fundamentals of the system. In this sense, a conceptual model acts as the “progenitor”, “parent” of all models that may be developed and which may describe system behaviour in a more detailed way. Conceptual models may thus appear at the beginning of chain of models where as we progress we have more details and respective complexity. One way to view a conceptual model is as the least complexity model that may allow development of further formal models. The importance of such models and how they blend with the rest of the other models will be discussed further in this chapter, where also their construction is elaborated.

4.2.1.3.5 Mathematical or Formal Models

A model is a useful working tool if it is linked to a language that allows deductive reasoning or carrying out of experiments which are both seen as means to understand, on one hand, and change the behaviour of the system, on the other. A general family of models which allow the study of behaviour, structure, experimentation, design and decision making, are those enabled by formal mathematical languages (mathematical tools). Models described by a mathematical language are called formal/mathematical models. Such models evolve from those dealing with the conceptualization of the system by enhancing those aspects dealing with aspects of behaviour of key system variables.

Formal models usually describe a system by a set of variables and operators that establish relationships between these variables. The values of the variables can be practically anything; real or integer numbers, boolean values or strings, or functions of time. The variables represent some properties of the system, for example, measured system outputs often in the form of signals, time, data, counters, event occurrence (yes/no). The actual model is the set of functions that describe the relations between the different variables. These models may not look like their real-life counterparts. Mathematical models are built using numbers and symbols that can be transformed into functions, equations, and formulas. They also can be used to build much more complex models such as matrices or general operators. The user can then solve the mathematical model by investigating the form and type of trajectories, seek an optimal solution by utilizing control/decision making enabled by the setup provided by the formal model. Since mathematical models frequently are easy to manipulate, they are appropriate for use with calculators and computer programs. Within the family of formal models we may have classifications which are linked to the nature of the behavioral variables, or the modeling languages which are used.

4.2.1.3.6 Data structures-models

Data Modelling is the process of structuring and organising data. A data model describes the structure of the data within a given domain and, by implication, the underlying structure of that domain itself. A data model may thus represent classes of data about which the modeller wishes to hold information, the attributes of that information, and relationships. The model describes the organization irrespective of how data might be represented in a data management system (e.g database). This structure plays an important role in the process of modelling; it gives a formal representation of the data needed to create a complete idea of the model, the objects, their relationships. Data models may be seen as the tracing of all information regarding the system sub-processes (parameters characterizing them), their associated variables and their values as functions of time, the possible history of the system in terms of events, experiments that have been carried out, changes etc. This information is a projection of the "living system" in the data domain, that is the data base. There is a need to link the structure of the data base with the conceptualization of the system. When such links exist, the data set, represented in a system structured data base may be seen as a primitive data model; otherwise, it may be seen as just a source that may assist the process of formal modeling. In following sections, we address issues of data and data mining as well as explain how a data description of the model fits into the modelling process.

4.2.1.4 The triad: System-Observer-Model

At this point and after having described the notion of the system and its relation to that of a model, the need to introduce the observer-modeller becomes essential. So far the model was considered a pure abstraction of the system. The missing link, however, defined by the modeller-observer relationship, was not properly discussed.

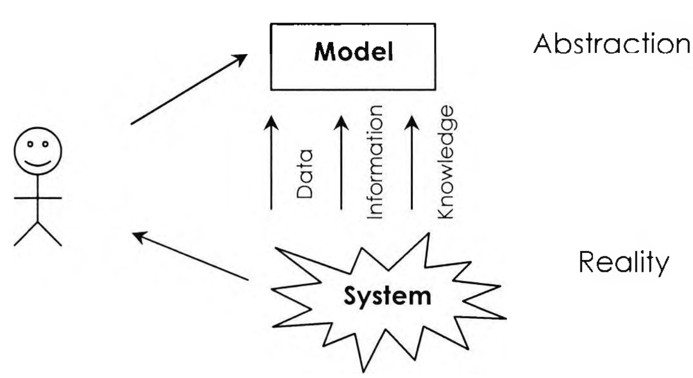


Figure 4.6: The triad system-observer-model.

A concrete idea of the system and its parts, how they communicate, the environment, and how they all work and compliment each other, gives the observer-modeler hints of what to look for when modeling. The observer is the person that "sees" the system and tries to create a model based on his understanding and interpretation of that reality. Central to the modelling problem is the crucial issue of the relationships between data, information, knowledge and how these combine in the derivation of models of the system under study. As simple as it may seem this is how someone would captivate the whole process of the projection of reality to a model. Charles Sanders Peirce defines the Peircean Sign, as the irreducible combination of Object (system), Representamen (model), and Interpretant (modeller). Charles Sanders Peirce (1839 – 1914) was a career scientist and engineer at the U.S. Coast Survey [Noz]. Figure 4.6 gives a graphic representation of the triad system-observer-model and Figure 4.7 is a representation of the Peircean sign.

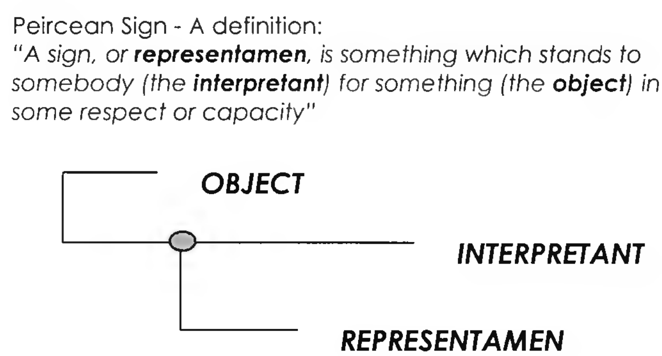


Figure 4.7: Peircean sign [Noz].

The tasks of the observer may be classified into specific groups as shown below:

- (OT1) *Specification*. The modeller is responsible for specifying the problem area; the problem area will be the system as visualised for the needs of the observer and the corresponding environment. Since he or she will be the one identifying the system, it is easily understood that the purpose of the modelling process will be closely dependent on the modeller. The observer will therefore produce the objectives of the modelling process, requirements and specifications.

- (OT2) *Conceptualisation.* The first step of identifying the system, its objects, and its environment. This task incorporates the production of the first picture of the model: a mixture of graphical, linguistic, structural and mathematical descriptions.
- (OT3) *Measurements, observations, experimentations and data collection.* As the link between reality and abstraction, one of the main duties of the modeller is to collect all the available information for the problem area-the system- so as to have a complete idea of the main features that need to be modeled.
- (OT4) *Interpretation, formal modelling.* An inseparable "component" of the modelling process, the observer will be the cornerstone associating data, observations and measurements with knowledge; this extraction process is inevitably strongly related to the observer. This task also incorporates the further steps from conceptualisation to formal modelling; from definition of key variables, attributes and identification of the relations between objects and the causality of those relations, to the formulation of a more formal form of modelling. Here the introduction to the formal language, is made, which gives the specific character to the system.

It is evident from the tasks of the observer that his importance lies in the linkage of the system to the model, through his interpretation. In his effort to model the system he/she will have to address many fundamental premodelling and modelling questions as described previously by (MQ1)-(MQ7), use several approaches, that will be discussed later on, either specific and systematic or abstract to move forward. This path leading to the modelling process, with result to complete the observer's tasks and produce a complete picture of the system that is, its parts, the environment of the system and the relations involved, as seen through the observer.

The outcome of (OT1), (OT2) and (OT3) define the premodelling stages and lead to the definition, construction of the modelling set M_s that has been previously defined. When interpretation of the elements of M_s starts we have the beginning of the main modelling activity. The identification of the above stages (premodelling stages (OT1)-(OT3)) may be further elaborated by considering the cluster of problems (MQ1)-(MQ7) previously discussed. The process of understanding how the projection of the reality yields a model uses the following "analogy" that summarises some of the key links and correspondences.

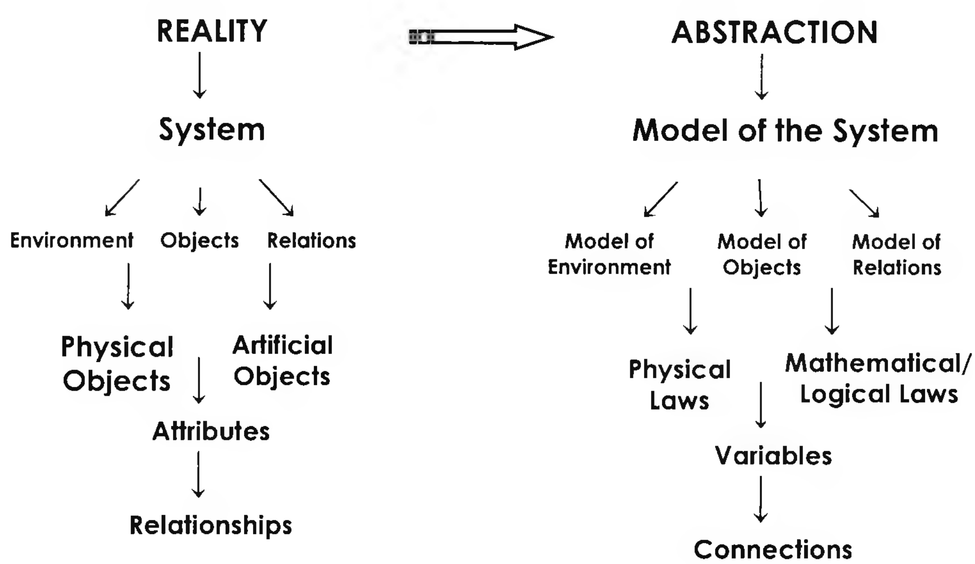


Figure 4.8: Correspondence between reality and abstraction.

The distance of the model from the reality judges the usefulness of the model. For a given purpose, the quality of the model depends on the "observer-modeller" capabilities, which involves the skills on how well to answer the issues in the key modelling problems in the list (MQ1)-(MQ7), as well as the purpose of the defined model. The modelling task is a problem that has fundamental conflict between accuracy and complexity and this is a compromise the modeller has to resolve. Defining the minimal, or so called "unfalsified model" is not always an easy task. The aim is always for models which satisfy the goals and requirements and have the simplest form (the least complexity model that interprets the current data and knowledge about the system and nothing more).

4.2.2 General issues in the process of modelling a system

At this point a discussion of various issues emerging when the modelling of a system begins. It will be noted that most of those issues are multidisciplinary and influence also other areas of interaction with the system.

4.2.2.1 Data and Data Mining

As previously described the process of modelling revolves around obtaining a useful description of the system. This involves the exploitation of large amounts of data from the system and its environment and some premodelling stages such as:

- Specifications of requirements
- Definition of Constraints
- Identification of Variables
- Specification of the Boundaries of the environment
- Processing the acquired data to extract information and knowledge.

This thesis does not regard data as a function irrelevant and independent from the modelling process. There is a role for data in the process of modelling and it is very important. However, its contribution cannot be exactly pinpointed in one stage of the process or another. Data flows throughout the modelling activity, but questions arise regarding the stages at which its use is much needed.

- Has the main structure of the model been decided upon and then data "comes" to enrich that structure? So basically, does the main structure, the variables, the functions, the behaviour have been resolved and then data infiltrates to provide more insight on all of the above aspects?
- Is data used in determining the main structure of the model? That is, does data play important role from the beginning of the modelling process as to how to construct the main skeleton of the model?
- Is the overall main structure of the model not affected by data?
- Does data only have significance by providing measures for the variables, constraints, limitations, etc? That is, is it only the quantitative aspect of data that is needed in the process of modelling?
- Or its significance comes as a combination of the above?

Coming up with answers for these questions would provide an insight in the process of *data mining*. Data mining is the process of discovering meaningful correlations, patterns and trends by shifting through large amounts of data, using pattern recognition technologies as well as statistical and mathematical techniques. It has been equated with the extraction of implicit, previously unknown, and potentially useful information from data. For that reason, data mining must be seen as part of the modelling process. During the data mining process the following areas have been identified as possible pitfalls, if not given the relevant attention:

- **Data Management.** Data Resource Management is the development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an enterprise. This definition is fairly broad and encompasses a number of professions which may not have direct technical contact with lower-level aspects of data management, such as relational database management.
- **Data Clustering**. Data clustering is a common technique for statistical data analysis, which is used in many fields, including machine learning, data mining, pattern recognition, image analysis and bioinformatics. Clustering is the classification of similar objects into different groups, or more precisely, the partitioning of a data set into subsets (clusters), so that the data in each subset (ideally) share some common trait-often proximity according to some defined distance measure [Wik]. Many data mining applications involve partitioning data items into related subsets; in image segmentation, clustering can be used to divide a digital image into distinct regions for border detection or object recognition.
- **Data quality**. There are a number of theoretical frameworks for understanding data quality. One framework seeks to integrate the product perspective (conformance to specifications) and the service perspective (meeting consumers' expectations) [Kahn]. Another framework is based in semiotics to evaluate the quality of the form, meaning and use of the data [Price]. One highly theoretical approach analyzes the ontological nature of information systems to define data quality rigorously [Wand].A considerable amount of data quality research involves investigating and describing various categories of desirable attributes (or dimensions) of data. These lists commonly include accuracy, correctness, completeness and relevance. **Data quality assurance** is the process of profiling the data to discover inconsistencies, and other anomalies in the data and performing Data cleansing activities to improve the data quality.

- **Data cleansing**. Data cleansing is the act of detecting and correcting (or removing) corrupt or inaccurate records from a record set. After cleansing, a data set will be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by different data dictionary definitions of similar entities in different stores, may have been caused by user entry errors, or may have been corrupted in transmission or storage. Preprocessing the data will also guarantee that it is unambiguous, correct, and complete [Wik].

The data acquisition/mining process is closely related with the past knowledge and experience of the modeller and in this process the modeller uses his or her experience, to judge, balance and decide on the information to be used. Trying to separate the observer from the data acquisition process is somehow impossible, as the observer is the one discovering data, transforming it to information and then using its output. The modeller must first develop an understanding of the system, the application of the model and its goals. He then needs to create a dataset for study; this dataset will be a sort of database which has been formed after processing the available data and structuring it into convenient sets. The term "processing the available data" suggests that some form of sorting of the available data has taken place, clearing out any irrelevant sets. For this reason one of the main concerns with data mining is that the questions the modeller needs to ask to get the information need to be accurate, relevant and representative of the specific system and the context in which this system is viewed. The exploration of any hypothesis regarding the system requires that the data from different sources and modalities be analysed in an integrated fashion such that the results of information obtained from the different data can be simultaneously interpreted and acted upon.

In this effort to collect and structure data, a more complete set arises, that of a *data model* a term will be explicitly discussed later on. From large amounts of data, a structured and representative set of data arises that can be used in the modelling process to enhance, clarify, measure aspect of the model.

4.2.2.2 The role of the observer: Knowledge transformers and Heuristics

The observer, as seen earlier, plays an important role in the process of modelling, by bringing into it certain attributes. The presence of these attributes can be seen by carefully observing the first stages of modelling. One of the first and main functionalities of the observer is to identify the problem area, from which the system will emerge. Since the universe of discourse is identified, the observer will then need to define the objectives of the model. By defining the universe of discourse and the set of objectives, the observer is using his own experiences to generate and transform knowledge and therefore all his definitions are biased by his own history. Here, some of the key procedures used in this process by the observer will be discussed.

The human mind is constantly gathering data and fusing useful information to support planning, decision making, and management of processes that consciously and unconsciously guide our every move and action throughout our daily lives. There has been some research done in this area, but the quantitative and qualitative measures of effectiveness and figures of merit for the human information processing system, i.e., the human cognition, have yet to be developed. Without a reasonable set of measures it becomes a difficult task to attempt to specify and to properly evaluate a user-centered, open-system information processing system. Therefore when consciously or even unconsciously a model has to be created, there exists a process through which this goal is achieved. In this report, an attempt is made to identify this process of modeling, giving it a more concrete framework if possible. Obviously, this will not be a try to tame someone's mind, perception and free will, by limiting his thoughts into framework, but it will give a general idea of how a model is created. The process of modeling begins with an idea, a concept; something that may defy a complete description and it ends with the delivery of a formal model.

- **Knowledge transformers**. Complex problems are meaningfully characterised in terms of multiple description levels. The human cognitive system – the mind-brain entity is arguably the most complex natural entity we are aware of. [Das] suggests that one of the ways in which we commonly attempt to grasp the complexity for the purpose of describing, explaining, or understanding mental processes is by recognising that it can be abstracted at multiple levels of descriptions. A knowledge transformer is an operator that derives a piece of new knowledge from a given input or an existing piece of knowledge.

Knowledge transformers	Description of how knowledge is transformed
Abstraction/ Detailing	Abstraction generates a new version of the knowledge with less detail than the original through the use of representation of abstract concepts or operators. Detailing is the opposite, in which the new knowledge is generated with more details.
Association/ Disassociation	Association determines a dependency between given entities or descriptions based on some logical, causal or statistical relationships. The opposite is disassociation, which asserts a lack of dependency.
Derivations (Reformulation)/ Randomisation	Derivations are transformations that derive one knowledge from another piece of knowledge (based on some dependency between them). Randomisation transforms one knowledge segment into another by making random changes.
Explanation/ Discovery	Explanation derives additional knowledge based on domain knowledge. Discovery derives new knowledge without an underlying domain knowledge.
Group Rationalisation (or Clustering)/ Decomposition	Group rationalisation involves the grouping of past designs according to their similarities when considering particular perspective(s) or criteria. Decomposition removes the groupings.
Generalisation/ Specialisation	Generalisation generates a description that characterises the entire concept based on a conjunction of all the specialisations of that concept. Typically, the underlying inference is inductive. Specialisation increases the specificity of the description
Similarity comparison/ Dissimilarity comparison	Similarity comparison derives new knowledge about a design on the basis of similarity between the design and similar past design(s). The similarity comparison is based on analogical inference. The opposite is dissimilarity comparison, which derives the new knowledge on the basis of lack of similarity between the two or more past designs.

Table 4.1: Knowledge transformers and their descriptions (adapted from Sim and Duffy, 1998) [Sim].

Table 4.1 describes the manner in which knowledge is transformed in the seven pairs of knowledge transformers. The process described in this table provides an explanation (in terms of generic mental steps) of how the basic elements of the modelling set M_s dealing with cognition are identified, that is, the objects/relations variables cognition.

- **Heuristics**. People solve problems in new domains by applying heuristics to declarative knowledge they have about the domain within which the problem resides. A heuristic is defined as a strategy that increases the chance that a problem will be solved, but does not guarantee success. The design process, with all its variations, can be seen as a heuristic. Declarative knowledge is defined as knowledge about something, for example, to know that a ball is round [Midl]. Declarative knowledge is described as "knowledge that". New domains are taken to be domains where the problem-solver may have knowledge about the domain in a declarative form but no specific procedural knowledge of how the problem is to be solved. Procedural knowledge is defined as knowledge of how to do something. For example, to know how to play basketball with the ball. Procedural knowledge is described as "knowledge how"[Midl].

  Experts in a domain have a large store of procedural knowledge that they can apply to any task within that domain. Novices in a domain, on the other hand, have some declarative knowledge, but a much smaller store of procedural knowledge. A consequence of this is that experts and novices solve problems differently. Experts in a domain simply apply their declarative and procedural knowledge to the problem and achieve a result. They draw on their understanding and apply a forward reasoning heuristic in solving the problem. A novice uses trial and error, and in the process, converts declarative knowledge to procedural knowledge. However, when solving problems requiring innovation such as in design and invention, problem-solvers are required to generate new ideas.

4.3 Final Thoughts

One of the important issues that needs to be stressed out is that the process of modelling is iterative and that should remain a focal point. Modelling is a feedback process, not a linear sequence of steps. Models go through constant iteration, continual questioning, testing and refinement. The initial purpose-the horizon of the model-dictates the boundary and scope of the modelling effort but what is learned from the process itself can be fed back to alter the basic understanding of the problem. The following figure provides a pictorial representation into the iterative process of modelling.

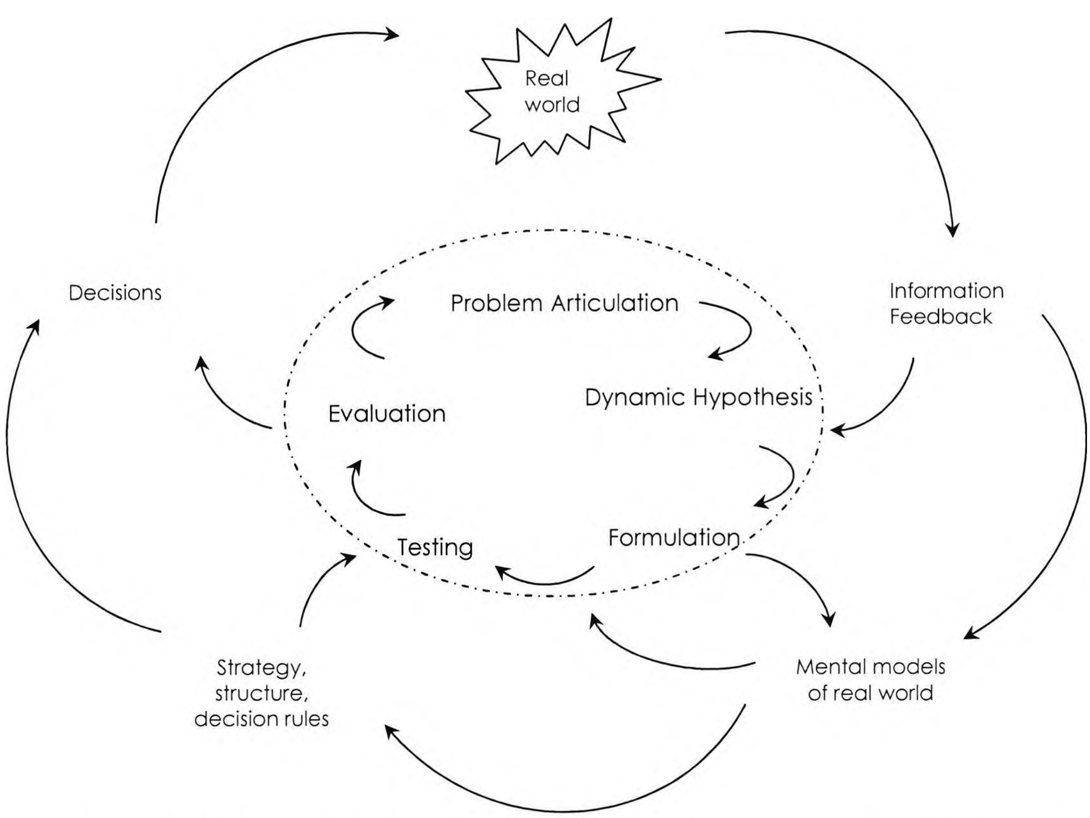


Figure 4.9: The modeling process embedded in loop learning feedbacks [Ster].

The learning feedback operates in the context of existing decision rules, strategies, culture and institutions which in turn are derived from the mental models of the observer. Mental models are described as collections of routines or standard operating procedures, typologies for categorizing experience, logical structures for the interpretation of language or attributions about every day life [Ster]. Most of the people do not understand the ubiquity and invisibility of mental models. Information fed back about the real world not only alters our decisions within the context of existing frames and decision rules but also feeds back to alter the mental model of the observer. As the mental model changes so does the structure of the system (how the observer sees the system), creating new decision rules and strategies.

From the figure the following steps can be derived in this circle of modelling:

- Problem articulation (System-Problem Identification area, System Conceptualisation)
  - What is the problem? Why?
  - What are the key variables and concepts that must be taken into consideration?
- Formulation of dynamic hypothesis (Knowledge extraction)
  - What are current theories of the behaviour being modelled? (Knowledge management, Design of experiments)
  - Explain the dynamics as consequences of the feedbacks
- Formulation of a model (Model construction)
  - Specification of structure and decision rules
  - Identification of parameters, behavioural relationships and initial conditions
- Testing (Design of experiments, Data Mining)
  - Test model for consistency with the purpose
  - Sensitivity testing
- Policy design and evaluation (Design of experiments, Data Mining)
  - What environmental conditions might arise?
  - What new design rules and strategies arise and how can they be represented in the model?

A number of issues regarding the process of modelling have been considered so far. This chapter was an effort to define a model in relation to the system being modeled, introduce the different questions that the modeller sets in the process of modelling. By identifying the questions of the modeller during the modelling process, we clarify and distinguish any grey areas in the different stages of the process and identify open problem area, for the questions that a straight forward answer cannot be given. The following are just few of the areas identified for which thorough understanding and clarification is needed.

- Question regarding the nature of the building blocks of a model aroused; questions we defined in this chapter as Modelling Questions (MQ). These relate with the knowledge we have for the objects, their relationships, their attributes. The knowledge we acquire and the way this knowledge is acquired is one of the focuses of those questions. Model definition that fits the existing data and the knowledge of the objects/relations is another issue. From model definition different kinds of questions come into the surface regarding model minimality as well as model simplification and expansion and how these are achieved. The questions formed provide a framework of problem areas; areas for which little has been done to provide concrete solutions or step-by-step approaches, and thus form open problem areas. These open problem areas have been identified as follows: *Knowledge extraction, Systems conceptualisation, Design of experiments, Model construction, Model minimality, Model expansion, Model reduction* and *Model simplification*. Developing modelling approaches requires tackling problems of the above classes in a rather substantial way.
- The importance of data was also stressed out throughout this chapter. Certain questions, though, were set regarding how its contribution affects the different stages of modelling. Those questions have yet to be satisfactorily answered in a way that is not specified in one area (e.g., Information systems, Artificial intelligence), but can be applied in a more general framework. This generic framework should not be followed by a lack of substantiality, but provide a basis for an understanding.

  Specifically:
    - How is data transformed into information? Issues of data mining and the knowledge brought by the modeller related with the modelling process. A question closely related with the Knowledge extraction problem previously identified, as well as the observer and his previous knowledge.

- What is the role of data in the structure of the model? Here again, we identify a question that if clearly answered could provide an insight in the Model expansion, reduction, as well as simplification problem.
- Is the role of data purely for quantitative reasons, that is, for providing measures for the variables, constraints, limitations, etc?

Coming up with answers for these questions would provide an insight in the process of data mining.

By clarifying these grey areas and obtaining more information, a unifying modelling approach can be built; one that can be understood and used by every area regardless of different terminologies and provide a concrete basis for any kind of modelling exercise.

A step that precedes the modelling process, which however involves some manipulation of forms of knowledge is the definition of a problem. It is perhaps redundant and trivial to ask what a problem is, for the simple reason that everyone knows the answer. However it would not be useless to define what is generally meant by the term *problem*. A problem exists when there is a need for something but the actions or series of actions are not immediately known. The desired objective may be tangible, how to open a can, or abstract, prove a theorem. The actions involved in getting the desired results include physical, perceptive and other purely mental activities, such as, judging similarity, remembering etc [Ares]. At this point the problem awareness stage has been reached and a philosophical analysis begins, regarding the specifications, the requirements, and the relevant domains, the assumptions about the applications, the relevance and interested parties. This stage of analytical argumentation, that can be thought as an early design of the system (problem), will be called in this report *conceptualization*; a term that will be discussed in the following chapter.

Chapter 5

Conceptual Modelling and the Stages of Modelling

The main purpose of this chapter is to identify the different stages in modelling, from early design to formal representation, giving special attention to the most critical phase of the modelling process, the conceptualisation. At present it appears that the opportunity to gain clarity within the design process is being wasted, with the process of design being generally poorly understood, which in turn leads to designers having no real structure or common focus [Mac et al]. Furthermore, in the previous chapter, general aspects of modelling were discussed; whereas in this chapter we will go into much more detail regarding the different stages of modelling.

5.1 Conceptual Modelling

5.1.1 A short bibliography of conceptual modelling

The conceptual phase of any modelling process is potentially the most creative stage of the overall process. However, it is at present the least understood. It is at this stage that modellers from all disciplines need to interact freely in a bid to achieve optimal design solutions that eliminate or reduce the need of compromise of design at a later, more critical phase of the process. There could be many causes for the poor problem understanding, one of which appears to be a lack of shared understanding of what processes should be followed. There can be little doubt that during the conceptual phase of a model, there exists great potential for taking decisions that can result in significant reductions in costs and increased overall model needs satisfaction [Mac et al].

In the definition of the conceptual model an essential concept that is involved, which relates to cognition, is the notion of conceptualisation, which deals with the specification of the system internally and externally. By internally it is meant the objects, the relations between the objects and by externally the environment the purpose of the model and the criteria, constraints. Conceptualisation is a first stage of understanding the system and it is a glance at how the system is and works without getting into much detail. It can also be perceived as a preliminary phase of problem definition. A *conceptual model* should reflect knowledge about the application domain rather than about the implementation of the system. The notion of *Conceptual Model* has emerged in many and diverse areas and such views are briefly summarized below.

Wilson [Wil] in his definition of a conceptual model, in the field of Systems and Information management, says that the models corresponding to his definition are pictorial/symbolic, fall in the category of *conceptual models,* which cover the qualitative aspects of the situation and building them. Such models precede any other kind of modelling (iconic, analogic, mathematic and logic, etc), as well as being a modelling form of its own. In his definition he gives four kinds of use for conceptual models [Wil]:

- As an aid to clarify thinking about an area of concern
- As an illustration of a concept
- As an aid to define structure and logic
- As a prerequisite to design.

In the area of Information Systems *Conceptual modelling* has been defined in the context of business applications as “an abstract model of an enterprise” and as a formal description of "some aspects of the physical or social reality for the purpose of understanding and communicating" [Wand]. Wand has proposed, that there are four roles for conceptual models:

- Provide a way for developers and users to communicate
- Increase analysts understanding
- Serve as the basis for the design
- And serve as documentation of the original requirements of the system for maintenance purposes.

The importance of conceptualisation has been considered to more depth in the area of General Systems and this was perhaps best expressed by Chesterton [Chest]. He claimed that the worst thing to not being unable to find the solution to a problem, was not being able to see the problem. Indeed, an intelligent being's understanding of reality is determined by two systems, whose interaction is obvious and so requires no further explanation. The first is the sensory system and its amplifications, equipped with systems to collect, directly or indirectly, information from the environment. The second is the conceptual system which extracts from all the above information the relevant concepts for solving the problem at hand, their internal relations and necessary or sufficient reasoning to arrive at the right conclusions [Ares]. So, conceptualisation is modelling by the problem solver. This means that there are many ways of conceptualising, and the components (concepts, relations and functions) of conceptualisations are not necessarily related. So, conceptualisation determines system validity, it represents the problem solver's view of the problem, and is, as a result, declarative or explanatory [Ares].

The main concern with the above viewpoints is the lack of any kind of definition of a conceptual model, abstract or concrete. All of the opinions regarding conceptual modelling come as expressions of why a conceptual model is needed, what are its roles, in what situations it is helpful, how significant and essential its use is, etc. These expressions try to describe what is conceptual modelling but fail to give any kind of definition. This is confusing and creates questions regarding the quality of the research done so far for conceptual modelling, and the results produced; results that are portrayed in the previous statements-descriptions of a conceptual model.

Not all of the viewpoints, though, were a poor description of the uses of a conceptual model. Some made an effort to vaguely "bound" a definition of a conceptual model. A conceptual model is composed of [Vem]:

- the universe of discourse, containing all the concepts output by conceptualisation;
- the basic relational set, encompassing all the relevant relations;
- and the basic functional set, which contains all the functions and procedures needed to reason correctly.

There are important conceptual difficulties that one has to surmount in building conceptual models which involve the visualization of the unfamiliar. This requires imagination and intuition [Vem]. Yet intuition is something difficult to define and even more so to pass on by formal instruction, but it usually amounts to an instinctive ability to make sensible guesses and judgments in the absence of adequate supporting evidence. A good guess is all that is needed for a major breakthrough. While intuition cannot possibly be taught, one can be helped in making inspired guesses by developing a mental framework. Vemouri [Vem] here identifies a problem area (universe of discourse), objects (outputs of conceptualisation), relational and functional sets, but fails to give the reason of existence, the role of building such a model.

What is evident from the discussion so far coming from different areas on conceptual modelling, is that although everyone might have a different idea of what the notion really means their notions have many common elements. One of the things that might change from area to area is the application of it and the terminology used to describe it. Fundamentally though they are not giving any definition on the subject and subsequently, their approach lacks a framework on which to base any further work, including the further stages of modelling.

5.1.2 Definition of Conceptual Modelling

Conceptual modelling is fundamentally the first attempt to model the system. It is the realisation of the "definition" of the system at a very early stage, meaning the specification of the main and basic notions that comprise a system and the relations of them. This implies the identification of the objects (sub processes), the most distinctive variables that describe them and their interconnection topology, as well as their environment and most importantly the goal/objective of the modelling exercise and the constraints. This stage is considered the most important in the modeling process, since it gives the basis for any other assumption made later on. At the conceptual modelling stage we specify the basic components:

- Goals and purpose of the modelling process
- Basic objects/sub processes
- Distinctive variables of the objects
- Structural relationships between objects
- System boundary
- Environment

The information that comprises the conceptual model is thus concerned with, the objects of the system, the structural relations between them, the boundaries with the environment, and these have to be addressed in a way that is as comprehensible as possible. A conceptual model can be a combination of verbal, schematic and simple mathematical descriptions that best specify, describe and represent the problem at an early stage in the modeling process. It will be a theoretical explanation; in accordance with its final purpose this model is *prescriptive*, *descriptive* and *explanatory*. For a system **S** with a modelling set $\mathbf{M_s}$, where

$$\mathbf{M_s=\{stimuli; responses; data; objects/relations\}}$$

the simplest form of the model that can be associated with M_S is the model that describes:

1. The objects of the system
2. The nature and type of relations between objects and the system boundaries
3. The listing of variables that appear as stimuli, responses.

A graphical/linguistic representation of the above elements is then referred to a *conceptual model* of the system.

In the definitions of conceptual modelling given earlier, there seems to be a unifying lack of definition of *the problem*. The case with the definitions about modelling and conceptual modelling, more specifically is that they all take for granted that there is a given system, but this is not always clear in every case. Real life does not provide modellers with paradigms and exemplary systems on which the modeller will work his or her techniques. Real life provides the world with problems, mishaps, omissions. Most of the times the modeller must look for the system, emerging from a problem. Most of the times the modeller does not only have to identify the objects, the environment, the specifications, the variables and the constraints, but most importantly he or she has to identify the problem and from that analyse how the system is emerging. Figure 5.1 is the representation of the process of formulation of the *system*, as it emerges from the problem area. Note however, that this process is entirely in the hands of the modeler and it is difficult to describe it in a systematic way. Extracting the relevant notion of a system that fits a given problem belongs to the premodelling set of activities and depends entirely on the skills of the modeler.

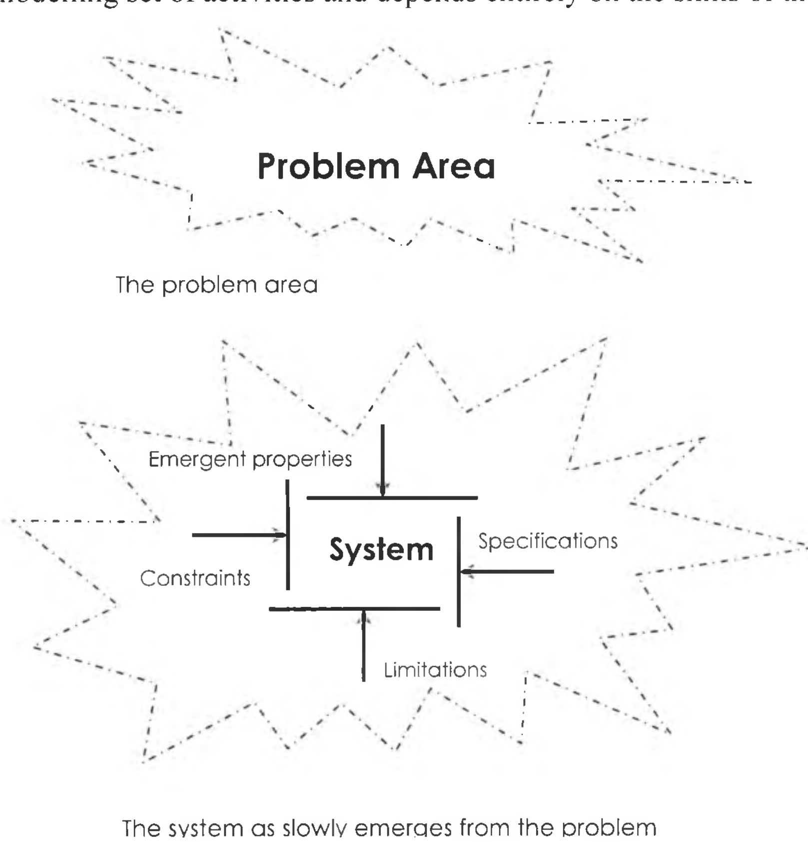


Figure 5.1: The problem lead to the emergence of an associated system.

In the context of Chemical Processes Synthesis we have the most systematic and constructive description of the formulation and development of early stages conceptual models, which are subsequently used as a basis for defining formal working models. The significance of the formulation of an early model of a process with inherent "good potential" for the final design, is the main driver for development of such models in the Process Engineering setup. Failure to do so, will result in an ill-defined model, which has inherent bad control properties, thus making the development of a control structure and control design very difficult or even impossible. The design of a process in an early stage involves as a fundamental stage the problem of conceptual modelling, which transforms requirements and objectives to sets of preliminary designs referred to here as *Conceptual Process Flowsheets*. The procedure of forming a conceptual model in the context of Chemical Process Systems is illustrated in Figure 5.2 and it has a generic character that allows its transference to other domains.

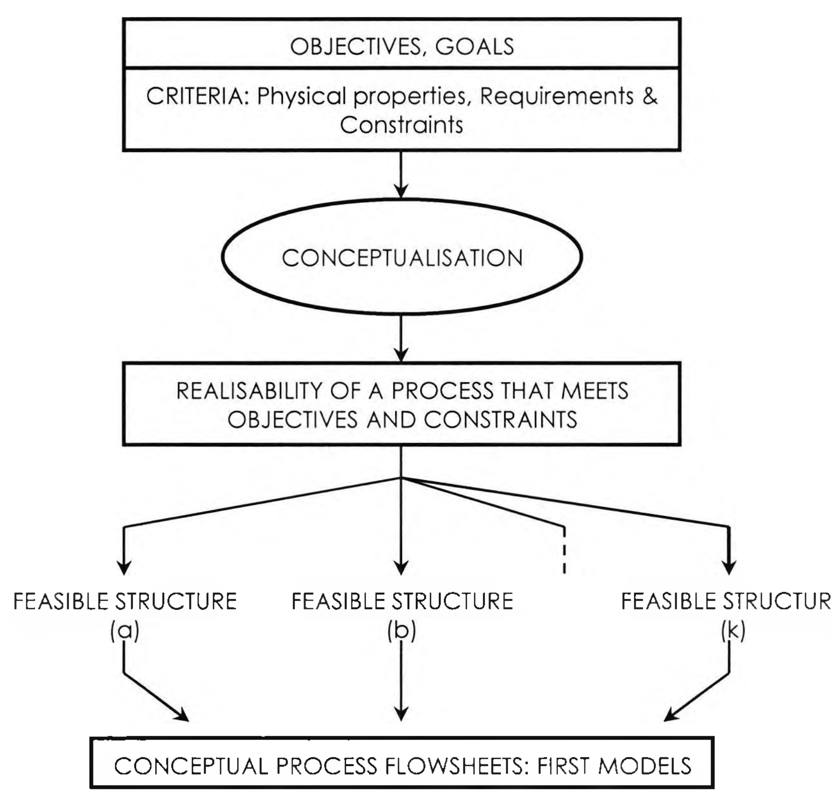


Figure 5.2: Summary of Development of Conceptual Process Flowsheets [Vaf].

The development of such models leads to a family of conceptual process flowsheets, which are the first models available and are denoted by M_i^c. The overall set of such models is denoted by $M \stackrel{\Delta}{=} \{M_i^c, i = 1,2,....,k\}$.The most important issue in this generation of process flowsheets is the identification of generic aspects, which may have some impact on other application domains, and simple ways of formulating such a conceptual family of models. Another very important issue is to characterise mathematically the structure of the resulting conceptual process flowsheet models M_i^c, in terms of the general interconnection rules and the associated graph, as well as the early description of subprocesses in terms of simple models [Steph], [Doug]. The resulting graph will contain the fundamental variables linked to the physical interconnection. As design progresses the dimensionality of physical interconnection streams may change, as the model becomes more detailed.

When using such models it is not possible to find accurate values for the parameters during the stages of early design and some of the assumptions lead to an oversimplification of the model. As design progresses, the model is the subject of continuous evolution. During this progression, choices have to be made constantly, for example on different types of equipment, operation modes etc. For this purpose, the most likely alternatives have to be elaborated in sufficient detail to enable a decision to be made. Quite often the requirements specification is not sufficiently detailed to allow making a rational choice, in which case the requirements specification is provided with more detail, in order to help the decision-making process. It is very common, for example, in chemical process practice, to make this decision in an informal way, primarily based on heuristics and industrial common practice. This decision making process is not restricted to the early stages of design, but applies to the whole design procedure.

The main mission of the pre-conceptual modelling stage can be summarized as follows:

- Investigate the domain and problem area. This is an extremely important but unstructured and difficult task. The modeller needs to identify clearly the problem area, from which the system will emerge.

- Specify the goals that drive the activity and the constraints that may restrict the outcomes.
- Gather and analyse the acquired information to identify strategic, tactical and factual information
- Identify concepts, objects and relations and record them in a concept framework, i.e. build a static model.
- Determine what concepts are important
- Create a formal representation of the objects and relations; formal considering the amount of information available at that stage of modelling.
- Determine the system structure and model information aspects.

The above activities are prerequisites to starting the formal process of conceptual modelling. The desirable features of a conceptual model should be:

- Expressive, concise, unambiguous, context-insensitive, and effective
- Clear and Correct
- Efficient
- Encode general knowledge about the domain
- Encode a description of specific problem instance

Conceptual modelling is the first stage in modelling and is an essential prerequisite for the modelling process. It is the building stone of the process; in fact that will bring together all the required information, construct the system and show the way for the solution of the problem. The whole of the modelling process relies on a solid conceptual model. It is obviously hard to construct such a "tool" but the advantages come with the concrete result. If information at this stage is faulty, incomplete, or non-existent, then the conceptual model will be carrying uncertainty that will be revealed later on in the modelling process. Obviously, uncertainty is a heavy burden in modelling and restricts the validity of the results. It is important that significant effort is dedicated at this stage due to the significance of the uncertainty at this early stage on the subsequent stages of model development.

5.2 The Modelling Stages

At this point it is useful to provide an overview of the sequencing of the various stages of the modelling process, and see where the stage of conceptual modelling fits in. As previously discussed, a model is a representation of the system. Any observer that has to model a system consciously or subconsciously passes through some stages in order to reach to the final formal model, that is either a model interpreted in mathematics or a model interpreted in another form of formal language. Hales [Hales] summarises the opinions of Bessant and McMahon [Bess & McMah] in suggesting that the way for designers and design researchers to gain improved understanding of the design process, is to move towards the development of flexible and adaptable models which take into account the dynamic nature of the design activity. Evidence suggests that the designer is better able to think on a particular problem when in possession of a general program of events through which the activity is likely to pass than when no such structural concept is held [Arch].

Obviously the complexity of the description and thus of the model is rising, as the stages move on. This is due to the fact that from the higher level where we consider the whole attention is now focused to the parts, concentrating into giving more details to what actually comprises the different aspects of behaviour of the system. Some of the stages that will be discussed may seem too simple to be considered a stage of their own, but the goal of this research is to break down the modeling process into small comprehensible parts. The idea behind this breaking down, apart from giving more information of the modeling process, is to investigate these subsequent stages and if possible, examine their formal relationships and how models from a given stage evolve to the following ones. This kind of research is crucial for integrating design and reengineering of processes, and can highlight the mechanisms of model evolution from early of design to late stages and thus may allow the development of systematic tools for intervening systematically in design. The various stages of the modelling process identified for the purposes of this thesis are:

- Linguistic Description
- Conceptual Modelling
- Structural Description

- Attribute Description
- Behavioral Description
- Data Structuring
- Formal Modelling

For each of the above stages of modelling we have the following dominant features:

- **Linguistic Description**

At the very beginning of the observation, when the observer acquaints himself with the system, the first description he or she can give is a Linguistic description, giving a general picture of the system. This description is a primitive picture of the system; it gives basic information about the objects of the system, the boundaries and the environment. This stage in the modeling process establishes the semantics of the problem area and therefore of the modeling stages that follow.

In many situations where a mathematical description is out of context, a linguistic description is always preferred. Documents, for example, that describe a scope of a business or a scope of a project, the management of a university, will use a linguistic description of the system, and that approach can be all that is needed to describe the system. If the purpose of modelling is to give a general description to an external observer, about, for example, the management of a university, the linguistic description-maybe along with some graphical description (see below)-would be adequate for that purpose. If however the purpose of the modelling is to identify risks and pitfalls, in the management of a university, and control them, then a linguistic description of the system is inadequate and further description is essential.

- **Conceptual Modelling**

This stage is considered the most important in the modeling process, since it gives the basis for any other constructions made later on. At this stage the problem area is defined and the system is identified. By system definition, it is meant the specification of:

  - Basic objects

- Structural relationships between objects
- System boundary
- Environment

This step can be a linguistic description or a schematic one. The information that comprises the conceptual model, that is, the objects of the system, the structural relations between them, the boundaries with the environment, need to be addressed in a way that is as comprehensible as possible. For the model to give a clear image of the system it is useful to have a graphical representation if this is possible.

- **Structural Description**

At a second stage a graphical aid of how the system is synthesized can be very helpful in understanding and further analysing the system. This description, that could just be a drawing of the system, is the first step in introducing the relationships between the objects of the system and the system and the environment. Visualisation is an important aspect of the process. The observer is directly involved in the data acquisition process as discussed earlier, and an effort to represent the system with something that is understandable is essential; that is the object of visualisation, for the model to be translated into a more natural representation for the observer. The structural description is a non dynamical description, focusing purely in the nature of the relationships of the system and thus assumes that the system conceptualisation has preceded.

In some cases a drawing of the system is non trivial if its parts that comprise the system are physical. In the case that the system was not physical, such as a center in a school of a university, the use of a visual aid in the form of block diagrams or something that would seem convenient to the specific observer can be helpful. Other times, when a graphical description of the system cannot be obtained-either because it is not helpful or because it is not possible and out of context- the use of an analogue description is more convenient. By comparing the given system with a system that has analogue characteristics, it is easier to identify the internal structure by comparison. Although trying to identify analogue characteristics requires a fundamental knowledge of the system from which the analogy is derived, simple characteristics that are dominant in the system can be used for that derivation.

- **Attribute description**

The description of knowledge for the objects as well as the relations between them in a way beyond the conceptual that is, the specification of their attributes will be called *attribute description.* By object attribute it is meant a quality regarded as a natural or typical part of the object. Thus, attribute modeling will be a way of introducing further characteristics and variables of the system description, as well as parameters. Furthermore, attribute modelling is setting a wide path for the mathematical relations to take place in the next stage, the formal modeling. Attribute modeling is a quality measure in the effort to describe the system. It is highly related with conceptual modelling, since some kind of identification of the attributes has already been seen in that first stage of modelling.

- **Behavioral Description**

This stage gives additional information about the relations between the objects, focusing on the causality of those relations. What is affected by who is the question that needs to be answered at this stage. This flow is what enhances this structural description and the system is seen to have a cause and effect. Given a representation of the behaviour of the components of a system and a representation of the structure of the system, that is the interconnection of the components, the ability to generate the behavioral description of the system as a whole is an important part of causal reasoning [Milne].

The time and possibly event evolution of a variable has been defined, in a previous chapter, as the ***behaviour*** of the variable. Using properties of the behaviour, considered for the totality of variables, enables the introduction of important system properties, which in turn allow the introduction of richer classifications of systems. The behaviour of a variable is the set of all possible values of the variable obtained under a given initial value and expresses the time evolution of the values of the variables. By behavioral description, it is meant a quantitative description of observations on the real system. The precise definition of the behavioral description is far from a trivial step; at this stage the scientific knowledge and know-how about observing, experimenting, measuring a system begins to take a more formal, mathematical way [Kalm].

At the stage of behavioral description all scientific information is incorporated in a model which has evolved by direct interaction with experiments or observations. Often behaviour may need to be abstracted to a level higher than that at which the component is specified. For example, in an electronic circuit the behaviour of the components such as a transistor and a resistor may be in terms of voltages and currents, while a device containing them may be described as an amplifier or oscillator. To go from the level of "currents" and "voltages" to one of "amplification" and "oscillation" requires an abstraction process [Milne].

- **Formal Description**

Formal modeling is the last stage of the modeling process and involves a class of models of variable complexity. The only real tool available for system-theoretic research today is mathematics. A formal model has a language associated with it that allows development of deductive reasoning as well as simulation. It gives a complete picture not only of the system and how it acts, but depending on the information available and the original specifications, can predict its reactions and give the necessary information of how to control it. Formal modeling makes use of the outcome of attribute modeling, in that it uses the parameters to construct the mathematical model.

- **Data Structuring**

It would be wise, at this point to make a distinction between the quality and quantity of the system. As already mentioned, attribute modeling is expressing several qualities, parameters of the system. It does not however, mirror the values of those parameters, which are essential for the calculation of the formal model. In this report all the relevant data of the parameters of the system will be "nested" in a different kind of model "full of numbers" that can be loosely called *Data model.* Data models are structures that enable data storage in a controlled and well thought out way. It is the link that connects data and the modelling process. Data modeling strives to bring the data structures of interest together into a cohesive, inseparable, whole by eliminating unnecessary data redundancies and by relating data structures with relationships [Wik]. From large amounts of data, a structured and representative set of data arises that can be used in the modelling process to enhance, clarify, measure aspect of the model. It is the projection of the existence and life of the system. A representative example of a data model is a database; sets of data collected and stored in such a way that for the specific system that it has been created it has a meaning and a purpose, and reflects the lifecycle of the system. The figure below shows how the various outputs of the modelling stages are interconnected.

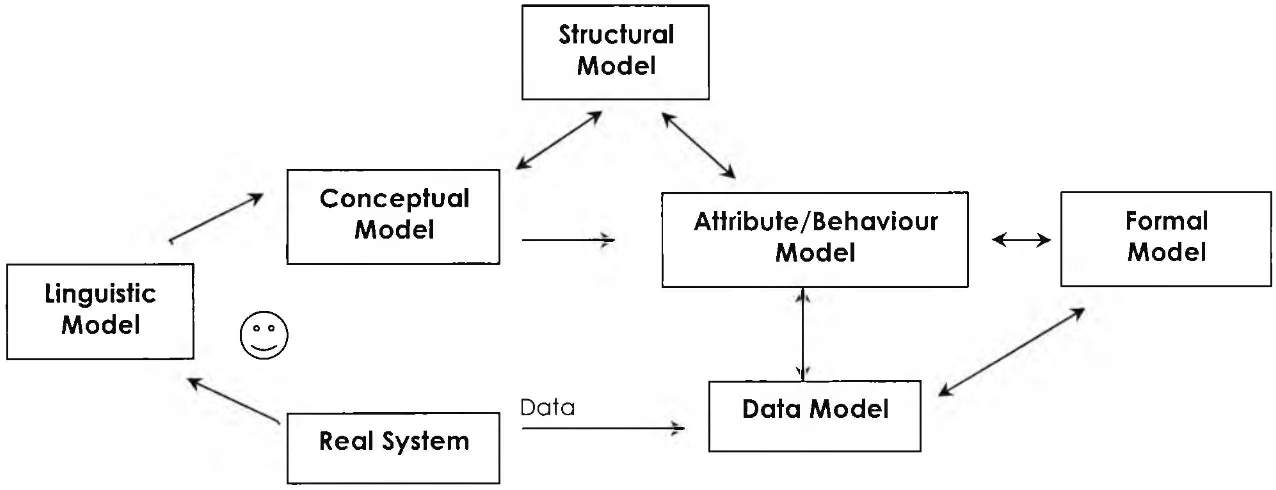


Figure 5.3: Relations between different stages of modeling

These various stages in the modelling process are part of a wider area called Qualitative research. Qualitative research is the non numerical examination and interpretation of the observations for the purpose of discovering underlying meanings and patterns of relationships. Qualitative research is generally considered to be exploratory and inductive in nature. It is used to get a general sense of what is happening and to form theories that can then be tested using quantitative research, which is viewed as confirmatory and deductive in nature [Wik]. It is designed to provide the researcher with the perspective of target the problem through immersion in the situation under study. Hypotheses are generated during data collection and analysis and measurement tends to be subjective.

Qualitative research in the area of modelling is:

- Exploratory. Explores the system and the universe of discourse, the description of the system and the processes, the functionalities, the responses of the system to stimuli.
- Explanatory. Finds answers to particular issues that arise out of the exploration of the system and its universe of discourse.
- Developmental. Uses ideas of the modeller for insight or inspiration and draws a path for a way forward.
- Evaluative. Measures responses to stimuli, evaluates processes.

In the qualitative paradigm the researcher becomes the instrument of data collection and results may vary greatly upon who conducts the research [Wein], an argument that has been discussed in the previous chapter of modeling. The advantage of using qualitative methods is that they generate rich, detailed data. The focus on processes and "reasons why" differs from that of Quantitative research (a term that will be described in the next chapter) which addresses correlations between variables.

5.3 Basic clusters of modelling methodologies

Through the basic steps of the modeling process, three terms were distinguished as a connecting chain between the stages;

- Structure. It indicates the relationships of the components that construct the system.
- Function. It is the intended purpose of the system and it is what is expected from the system.
- Behaviour. It is how the expected results from functions are obtained.

Throughout the process of modeling we have tried to identify and clearly define each of these characteristics. In the overall modelling process the above notions appear as fundamental building blocks of the information about the system. Furthermore, this information is used as a distinguishing feature in certain modelling methodologies; that is using as a main feature the functionality, or the structure or the behaviour of the system, to create several clusters of modelling methodologies that are moving inside an area that is trying to model the system using these characteristics.

At this point three examples of modelling methodologies that are using these characteristics will be discussed in an effort to understand the practical way these methodologies provide a richer, more detailed idea of the system and how it works, focusing on the information provided when analyzing these characteristics. These modelling methodologies are only examples of many and diverse modelling methodologies and by no means should they be considered as exemplary ones. In fact, it is the system in most situations, and the available information about it that will somehow decide about the modelling methodology to be used.

5.3.1 Petri-Nets

A Petri net is a graphical and mathematical modeling tool. Petri Nets theory allows the system to be modeled by a Petri Net, amenable both for correctness and efficiency. A Petri net, like a differential equation is a mathematical representation of the system. Despite the diversity of the systems that need to be modeled, there are several common points. One fundamental idea is that systems consist of separate interacting components, as previously mentioned. Each component may itself be a system but its behaviour can be described independently of other components. Each component has its own state of being. The state of a component is an abstraction of the relevant information necessary to describe its action. This state may change over time and the whole concept of "state" is very important to modelling that component. The components of a system exhibit *concurrency* or *parallelism*; activities of one component may occur simultaneously with other activities of other components. Petri nets are designed to model systems with interacting concurrent components. Mostly, Petri nets are suited to model and analyse discrete event dynamic systems; discrete event systems are dynamic systems, in which the state changes only at discrete points in time, in agreement with unexpected occurrence of physical events. Examples of such systems are communication networks, computer systems and discrete part manufacturing systems.

A Petri net is composed of four parts: a set of places P, a set of transitions T, an input function I, and an output function O. The input and output functions relate transitions and places (with *arcs*). *Input arcs* connect places with transitions, while *output arcs* start at a transition and end at a place. Places can contain *tokens*; the current state of the modeled system (the *marking*) is given by the number (and type if the tokens are distinguishable) of tokens in each place. Transitions are active components. They model activities which can occur (the transition *fires*), thus changing the state of the system (the marking of the Petri net). Transitions are only allowed to fire if they are *enabled*, which means that all the preconditions for the activity must be fulfilled (there are enough tokens available in the input places). When the transition fires, it removes tokens from its input places and adds some at all of its output places. The number of tokens removed / added depends on the cardinality of each arc. The interactive firing of transitions in subsequent markings is called **token game** [Des].

A graphical representation of a Petri net structure is more useful for illustrating the concepts of Petri net theory. A Petri net graph has two types of nodes:

- A circle, which represents a place and
- A bar, which represents a transition.

Directed arcs (arrows) connect the places and the transitions. An arc directed from a place P to a transition T defines the place to be an input of the transition. Multiple arcs from the input places to the transition indicate multiple input of the transition. An arc from the transition to the place indicates an output place. If an arc is directed from node i to node j (either from a place to a transition or a transition to a place), then I is an *input to j,* and j is an *output* of i. In the figure below for example, place p_1 is an input to transition t_2, while places p_2 and p_3 are outputs of transition t_2.

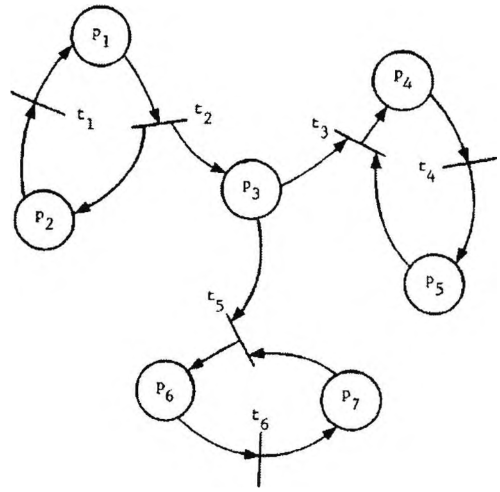


Figure 5.4: A simple example of a Petri net [Pet].

Petri nets model systems, and particularly two aspects of systems, *events* and *conditions,* and the relationships among them [Pet]. In this view, in a system, at any given time, certain conditions will hold. The fact that these conditions hold may cause the occurrence of certain events. The occurrence of these events may change the state of the system, causing some of the previous conditions to cease holding, and causing other conditions to begin to hold Consider for example the following description of a computer system [Pet]:

- Jobs appear and are put on an input list. When the processor is free, and there is a job on the input list, the processor starts to process the job.
- When the job is complete, it is placed on an output list, and if there are more jobs on the input list, the processor continues with another job; otherwise it waits for another job.

This is a very simple system composed of several elements: the processor, the input list, the output list, and the jobs. We can identify several conditions of interest:

- The processor is idle (inactive);
- A job is on the input list;
- A job is being processed;

- A job is on the output list;

And several events:

- A new job enters the system;
- Job processing is started;
- Job processing is completed;
- A job leaves the system.

The Petri net of Figure 5.5 illustrates the modeling of this system. The "job enters" transition in this illustration is a *source;* the "job leaves" transition is a *sink.*

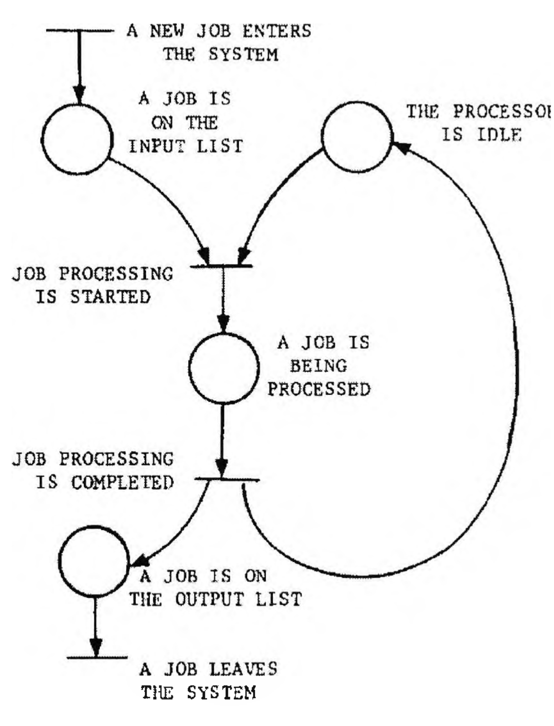


Figure 5.5: Modelling of a simple computer system [Pet].

Petri nets are a promising tool for describing and studying systems that are characterized as being concurrent, asynchronous, distributed, parallel, nondeterministic, and/or stochastic. As a graphical tool, Petri nets can be used as a visual-communication aid similar to flow charts, block diagrams, and networks. In addition, tokens are used in these nets to simulate the dynamic and concurrent activities of systems. As a mathematical tool, it is possible to set up state equations, algebraic equations, and other mathematical models governing the behavior of systems [Des].

In a Petri net model, the events which relate solely to one or the other can occur independently; there is no need to synchronize the actions of the jobs and the processor. Thus jobs may enter or leave the system at any time independent of the action of the processor. However, when synchronization is necessary, for instance when both a job and an idle processor must be available for processing to start, the situation is also easily modeled. Thus a Petri net would seem to be ideal for modeling systems of distributed control with multiple processes occurring concurrently. Another major feature of Petri nets is their asynchronous nature. There is no inherent measure of time or the flow of time in a Petri net. This reflects a philosophy of time which states that the only important property of time, from a logical point of view, is in defining a partial ordering of the occurrence of events. Events take variable amounts of time in real life; the Petri net model reflects this variability by not depending upon a notion of time to control the sequence of events. Therefore, the Petri net structure itself must contain all necessary information to define the possible sequences of events of a modeled system [Pet].

5.3.2 System Dynamics

System Dynamics have been defined by J.W.Forrester as "...the investigation of the information-feedback characteristics of managed systems and the use of models for the design of improved organizational form and guiding policy". It uses quantitative means to investigate the dynamic behaviour of socio-technical systems and their response to *policy*. Policy is the rationale that determines how a stream of decisions will be modulated in response to changing inputs of information [For]. The fundamental purpose of system dynamics is to achieve comparable quality of design and hence performance in managed systems. In systems dynamics state variables are called levels or stocks, the inflows and outflows are called rates or flows [Cell].

The concepts encompassed in System Dynamics include [Starr]:

- A focus upon descriptive behaviour over continuous time, such as growth, decay, oscillation, etc.
- Identification of a system boundary such that the behaviour of interest will be generated internally from feedback loops containing time delays and non-linear elements.
- A specified format for identifying and depicting feedback system elements, goals, discrepancies levels, rates and outflows.
- A model-building attitude that attempts to include and quantify all factors that are believed to influence the behaviour of interest whether or not such factors have foundation in accepted theory or empirical data.
- Identification of behaviour characteristics from the visual inspection of time response patterns generated by computer simulation.

A vital principle of System Dynamics is to incorporate all information believed to significantly influence behaviour into the model, leaving out unnecessary detail. System Dynamics is also problem-centered or better goal-oriented. The phase of problem definition is critically important. The first step is to define the rates and then formulate the set of state equations. The derivation of state equations could be managed by enumerating all the factors that influence the rate variables. Such an enumeration is called *laundry list* and the influencing factors may be levels, rates, or auxiliary variables called *converters*. After the design of the laundry list of all rate variables all the laundry lists can be connected in one flowchart, which is referred to as influence diagram or causal loop diagram. Causal loop diagrams give the information of which variable depends on which other variable, but they do not reveal the nature of this dependency. All paths in a causal loop diagram have signs which describe whether the influence is positive or negative. These signs allow the user to analyse the stability behaviour of the model in qualitative terms. By following the path around a closed loop, the number of positive negative or positive loops can be identified. If the total sum of the negative signs is even then, a positive feedback loop is identified. If the total sum of negative signs is odd then the feedback loop is negative. Positive feedback loops are always unstable, that is they are responsible for unbounded growth in the model.

Negative feedback loops are more difficult to access. Stability in these cases depends on the total open-loop gain. This observation is related to the Nyquist stability criterion for feedback control systems.

The strength of System Dynamics methodology is that it allows the modeler to blend deductive with inductive modelling techniques. With deductive modelling all models are created on the basis of physical understanding of the process that is being modeled. For more complex systems, though less knowledge is available that would support deductive modelling. Furthermore, uncertainties inherent in most parameters of such systems make it these models less accurate. In inductive modelling, the modeller makes observations about the system and then try to fit a model to the observed data. The structural and parametric assumptions behind inductive models are not based on physical intuition but on factual observation. Since inductive models are based on observation such models are difficult to validate beyond the observed facts.

The following example [Kirk] illustrates how a simple simulation model is developed for a business process, specifically, a model for a simple production distribution system. A basic stock and flow diagram for the system is shown in Figure 5.6.

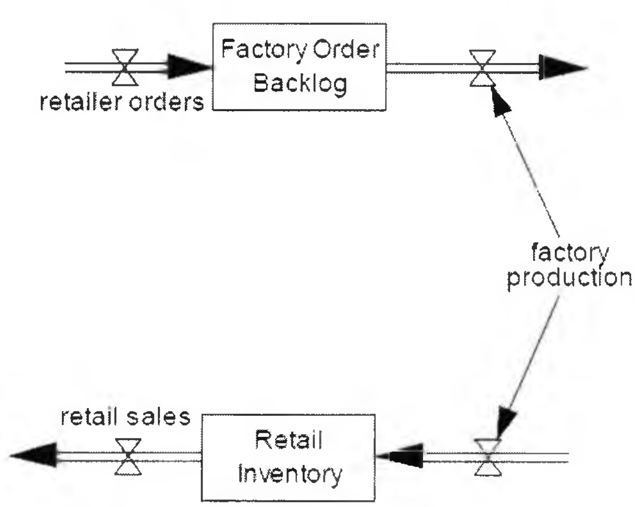


Figure 5.6: A simple production-distribution system [Kirk].

There are two flow processes:

- The production process shown at the top of the figure with a flow to the right, and
- The distribution system shown at the bottom of the figure with a flow to the left.

The production system is a flow of orders, while the distribution system is a flow of materials. The two processes are tied together by factory production, as shown at the right side of the figure. As items are produced, the orders for these items are removed from the Factory Order Backlog, and the items are placed into Retail Inventory.

The production-distribution system shown in Figure 5.6 is simpler than most real systems. These often involve multiple production stages, and also multiple distribution stages (for example, distributor, wholesaler, and retailer), each of which has an inventory of goods. By adding several information flows we have the following stock and flow diagram:

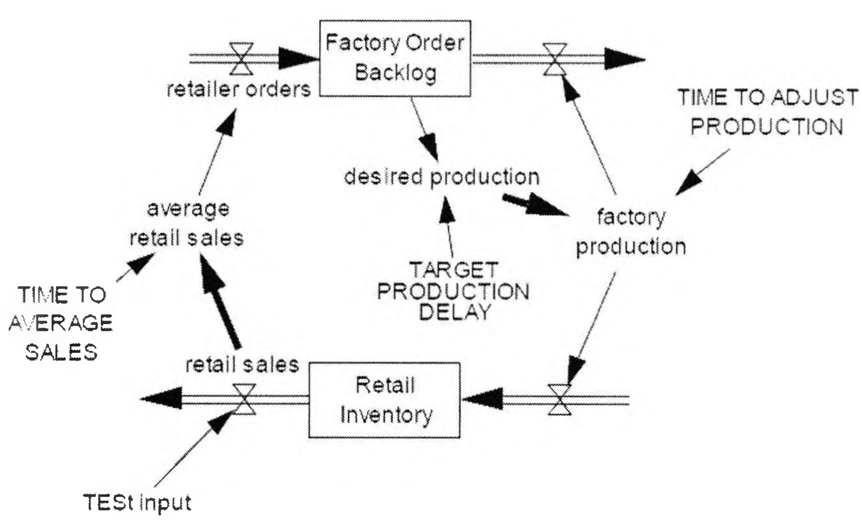


Figure 5.7: A stock and flow diagram of the model [Kirk].

From the information arrows shown there, we see that there is a "desired production" which depends on the Factory Order Backlog and the TARGET PRODUCTION DELAY. This desired production is then used to set the actual "factory production", but there is some delay in adjusting factory production, as shown by the constant TIME TO ADJUST PRODUCTION. In this diagram, a delay in an information flow is indicated by using a thicker arrow. Such thicker arrows are shown pointing from “retail sales” to “average retail sales” and from “desired production to “factory production”.

Based on these diagrams further analysis takes place, incorporating the system’s specifications as well as the values of the system variables, into equations that result from the diagrams and previous knowledge such as:

$$\text{desired production} = \frac{\text{Factory Order Backlog}}{\text{TARGET PRODUCTION DELAY}}.$$

The primary purpose in constructing this model is to determine ways to improve the performance of the production-distribution process. In particular, studying different possible retailer ordering policies and how these impact the performance of the entire production-distribution process.

5.3.3 Qualitative Physics

The behavior of a physical system can be described by the exact values of its variables (forces, velocities, positions, pressures, etc.) at each time instant. Such a description, although complete, fails to provide much insight into how the system functions. The insightful concepts and distinctions are usually qualitative, but they are embedded within the much more complex framework established by continuous real-valued variables and differential equations. Humans appear to use a qualitative causal calculus in reasoning about the behavior of their physical environment. Judging from the kinds of explanations humans give, this calculus is quite different from the classical physics taught in classrooms. This raises questions as to what this (naive) physics is like, and how it helps one to reason about the physical world [de Kleer]. Studying and formally representing this knowledge about the physical world is one of the motivations underlying research on qualitative reasoning. This motivation is illustrated by the Naïve Physics Manifesto ([Hayes], 1979) that proposed “the construction of a formalization of a sizeable portion of commonsense knowledge about the everyday physical world: about objects, shape space, movement, substance (liquids and solids), time, etc" [Bred].

In classical physics the crucial distinctions for characterizing physical change are defined within a non mechanistic framework and thus they are difficult to ground in the common-sense knowledge derived from interaction with the world. Qualitative physics provides an alternate and simpler way of arriving at the same conceptions and distinctions and thus provides a simpler basis for understanding the physical world; it provides an alternate way of describing physical phenomena [de Kleer]. Qualitative or Naïve physics is an alternative, far simpler physics, which helps understand how humans model the behaviour of a physical system or how they reason the functioning of a process [Cell]. This area of naïve physics can be found under different titles such as qualitative reasoning, common sense reasoning or knowledge, qualitative models and mechanistic mental models.

To summarise the main functions of Qualitative physics are to [Forb]:

- Formalize the intuitive knowledge of the physical world
- Develop reasoning methods that use such knowledge for interesting tasks
- Develop computational models of human commonsense reasoning.

The key ideas behind Qualitative Physics are [Forb]:

- Quantize the continuous for symbolic reasoning
  Example: Represent numbers via signs or ordinal relationships
  Example: Divide space up into meaningful regions
- Represent partial knowledge about the world
  Example: Is the melting temperature of aluminum higher than the temperature of an electric stove?
  Example: "We're on Rt 66" versus "We're at Exit 42 on Rt 66"
- Reason with partial knowledge about the world
  Example: Pulling the kettle off before all the water boils away will prevent it from melting.

Example: "We just passed Exit 42, and before that, was 41. We should see 43 soon."

To create the model it is necessary to not only identify the relevant objects and interactions, their important properties and quantities but also to determine what is irrelevant or negligible [Bred].

Qualitative physics modelling techniques requires a discretisation of continuous physical phenomena so every situation is regarded as a physical device or machine which consists of components and contributes to the behaviour of the whole device. The structure can be represented by a topology where nodes represent components and edges represent conduits. The variables used to describe the behaviour of the device are non continuous real valued but they are described qualitatively, that is they take only a small number of values. Each qualitatively value corresponds to some interval in the real number line. This qualitative approach sometimes causes loss of information but it is assumed that the potential presence of ambiguity does not affect the definitions and the concepts defined. The most important information of a quantity is whether it is increasing, decreasing or remaining the same. These states are described as +, - and 0 respectively.

Consider the case where we need to see what happens to humans when train accidents happen. The first thing that we need to do is to understand which parts of this modelling exercise are irrelevant for our purpose. So, things like the number of carriages, or the distance from one station to the other, are irrelevant and through reasoning we limit the modelling exercise to the acceleration of the train. What is therefore considered a huge modelling challenge-modelling a train on the move-is limited to analysing the different effects that the acceleration of the train has on humans and categorising those effects into spaces that have meaning for this modelling task. Therefore, by saying:

- Train acceleration 0-30 $m/s^2 \rightarrow$minor injuries
- Train acceleration 30-60 $m/s^2 \rightarrow$major injuries
- Train acceleration >60 $m/s^2 \rightarrow$fatalities

We construct a model that quantifies the information obtained from the analysis, divides it into spaces that are meaningful and represent partial knowledge of the whole problem-that being the moving train.

5.4 Final thoughts

An essential concept that is related with the first stages of modelling is that of conceptualisation. This first stage of understanding the system and how it works without getting into much detail is still an open problem area. An effort was made in this chapter to provide a basic understanding of what a conceptual model is, what it brings into the process of modelling. Further research, however is needed in the area of creating a conceptual model using a well defined step-by-step approach, that would take the modeller from the early stages of identifying the problem area, to the production of an early model that would show clearly objects, relations, environments, as well as purpose.

The different approaches and the resulting lack of a shared understanding of the individual processes, could explain the confrontational attitudes which are apparent between disciplines. An ordered approach to the design proves essential if people are to work together effectively towards common goals [Taylor]. To end this, it is apparent that an integrated design framework is essential; one that would leave no gaps in the modelling sages providing a process with constant flow, that is recognized by diverse areas.

Chapter 6

Conceptual Modeling and Development of Control Stuctures/Architectures

In this chapter a discussion regarding the use of general conceptualization and modelling in the development of control structures and architectures is considered. We use as a driving paradigm the integration of overall operations in an industrial enterprise [Karc, 1]. The area of overall process of operations involves processes of different nature expressing functionalities of the problem and specific relations are identified between the subprocesses. The issue of development of the generic features of the control and information architecture at such early stages is very important and expresses a new step in our effort to design the overall system that is part of the overall process of *conceptual design* of the system. This discussion will provide an introduction to the main issues related to the subsequent stages of conceptual modeling which is the conceptual design process. We consider first the general issues on the modeling of the Operational Hierarchy of Continuous processes [Karc, 1] and then we will consider the problem of developing the control and information structure architectures.

Integration in the process area is of paramount importance for improved profitability in a global market (flexibility in product portfolio and market variability), enhanced satisfaction of frequently conflicting and stricter requirements (environment, other legislation), as well as enhanced quality and reliability. A simple illustration of the overall enterprise level activity is given in the figure below where the following main areas are distinguished.

(a) Business Level Activities
(b) Production Related Operations
(c) Overall Systems Design
(d) Vertical Activities

The diagram below indicates a natural nesting of problem areas, where design issues provide the core, linked with the formation of the physical process that realises production.

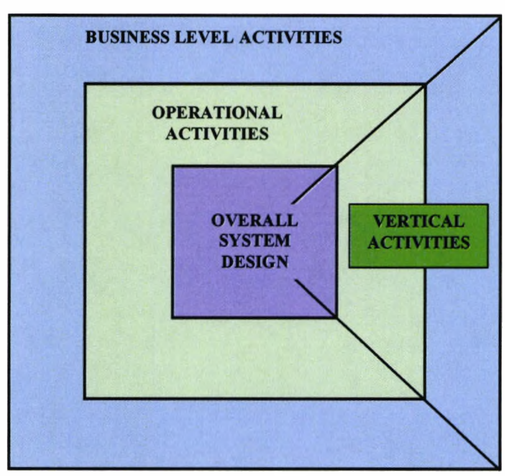


Figure 6.1: Nesting of Industrial Enterprise level activities

Production level activities take place on a given system, they are mostly organised in a hierarchical manner and they realise the higher level strategies decided at the business level. Vertical activities are issues going through the Business-Operations-Design hierarchy and they have different interpretation at the corresponding level. The problem of integration of Business level, Operation Issues and Design aspects is a multidisciplinary problem which is recognised as one of the major technological challenges. Understanding the relationships between problems on a horizontal (same level), as well as vertical (going through different levels) directions, implies an ability to describe the links between models associated with the particular problems, as well as a capability to translate issues, requirements from one set up to another. The study of the integrated industrial enterprise [Karc, 1], [Rijn] involves major tasks which are those related to:

- Design of Processes,
- Operation of Processes.

The operation of production of the types frequently found in the Process Industries relies on the functionalities, which are illustrated in Figure 6.2. Such general activities may be grouped according to certain criteria described below (see also [Mor, 2]):

(a) Enterprise Organisation Layers

(b) Monitoring functions (i.e. measurement, assessment) providing information to upper layers.

(c) Control functions setting goals to lower layers.

Note that the process unit with its associated Instrumentation (sensors and actuators) are the primary sources of information. However, processing of information (definition of diagnostics) can take place at the higher layer. Control actions of different nature are distributed along the different layers of the hierarchy (control and decision problems).

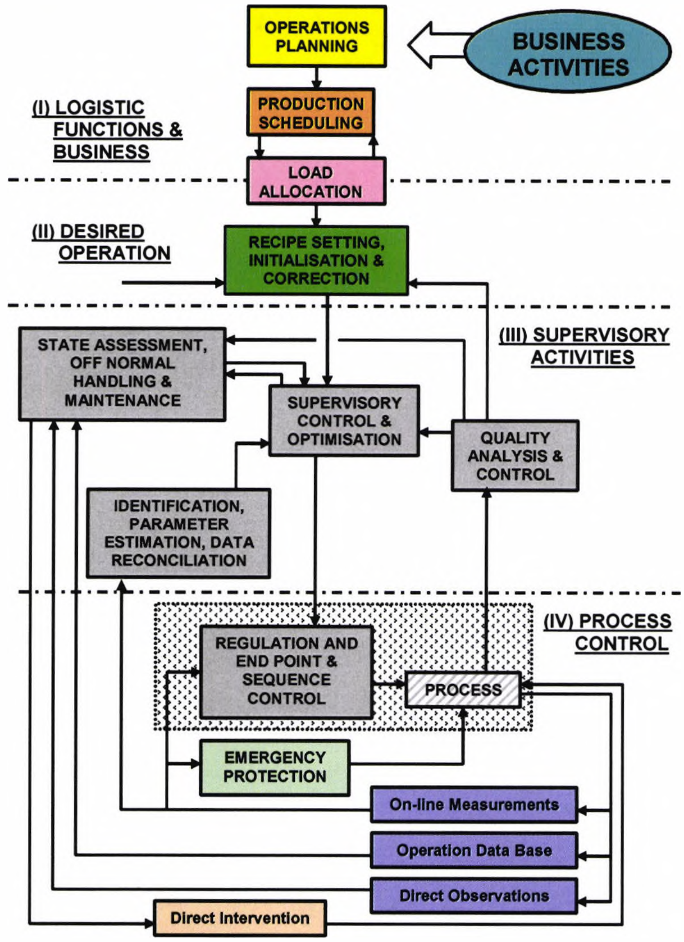


Figure 6.2: Functions for Operations of Process Plants

The functions shown in Figure 6.2 are of the following type [Rijn]:

(a) **Operations Planning:** This refers to activities such as feedstock negotiation and acquisition, customer orders, resource planning etc.

(b) **Production Scheduling:** This is concerned with the optimal timing of different operations runs and involves the combination of feedstock types and specification of the required type/quality of end products from all production locations

(c) **Load Allocation:** This involves the setting of the loads of the processing and utility plants of the overall production unit, such that they satisfy the production scheduling constraints.

(d) **Recipe Setting/Initialisation/Correction**: This is the higher layer of supervisory activities and deals with the co-ordination of the "mode" of operation defined as the set of conditions required for producing the desired products.

(e) **Quality Analysis and Control:** This involves the measurement, estimation of the important quality variables and attributes and then the initiation of corrective actions when product quality deviate from the set standards.

(f) **State Assessment, Off Normal Handling and Maintenance:** This set of activities is linked to the estimation of the actual "state" of the process based on all available information.

(g) **Supervisory control and Optimisation:** Integrating the results from desired operations, quality analysis, state assessment and the general business objectives (coming from the higher business layers of the hierarchy), as well as taking into account the operational constraints (physical) and regulatory constraints (safety, environment etc) to produce an optimal policy, is the aim of the current task. This activity produces as output the optimal set points for the physical operation of the process.

(h) **Identification, Parameter Estimation, Data Reconciliation:** The control activities require models and relevant data that can lead to the identification of such models. Part of the supervisory activity, in collaboration with the design team, is the selection of the data, their validation, and then the identification of model parameters. Such an activity provides links with design, as well as model based diagnostics. This area is part of a wider activity referred to as Data Management.

(i) **Regulation, End Point and Sequence Control**: This refers to the regulating control loops, usually embedded in the Process Control and Data Acquisition (PCDA) systems (i.e., DCS). Direct intervention on the process from the Control room is also included here.

(j) **Emergency Protection**: This refers to the Emergency Shut Down Systems.

(k) **Process Instrumentation and Information System**: This refers to the overall system for on-line Measurements, Creation of the System Data base and may involve direct Observations, Data Storing and Management.

It is apparent that the complexity of operating the production system is very high. A dominant approach as far as organising such activities is through a Hierarchical Structuring [Mes]. However, other forms of organisation are emerging at the moment, [MacF, 3], but their full potential has not yet been evaluated in the context of process systems. The study of systems and modelling issues depends on the organisational form that is adopted. Here we will restrict ourselves to the Hierarchical organisation paradigm [Mes]. This example is of an industrial process plant is an excellent paradigm to demonstrate issues of hierarchical type. The multilayer layout of the paradigm means that the output of every stage is fed to the next level as an input, and that layer is feeding its output as an input to the next layer. This type of architecture will be discussed later on as part of the global control and measurement issues in the hierarchy of operations.

The Clusters of Systems, Modelling, Measurement and Control problems in the Design and Operations of processes are considered below.

6.1 Modelling issues in the Operational Hierarchy

For the study of both areas we require models of different type, which are tuned to the respective goals of the particular function, design phase. The border lines between the families of Operational Models (OM) and Design Models (DM) are not always very clear and frequently the same model may be used for some functions.

A major classification of models is to those referred to as "black" and "white" models [Rijn]. White models are based on physical principles and their development requires a lot of process insight and knowledge of physical relationships. Such models can be applied to a wide range of conditions, contain a small number of parameters and are especially useful in the process design, when experimental data are not available. Black models on the other hand are based on standard relationships between input and output variables containing many parameters, require little knowledge of the process and are easy to formulate; however, such models require appropriate process data and they are only valid for the range, where data are available. Black models can be turned to grey ones [Rijn], if we know the ranges of process variables; hybrid, "White/black" models also may arise, when part of the model is white whereas difficult parts (such as chemical reactions etc.) are modeled as black models.

The overall problem of Process Operations is multidisciplinary and it is characterised by a high degree of complexity. The natural way of handling high complexity is through *aggregation, modularisation* and *hierarchisation.* To be able to lump a set of components, subsystems together and then treat the aggregate, composite structure as a single object with a specific function, the sub-systems must effectively interact. *Modularisation* refers to the composition of specific function units to achieve a composite function task. *Aggregation* and modularisation refer to physical composition of subsystems, set of subsystems through coupling, and it is essentially motivated by the needs of design of systems with dedicated operational function. *Hierarchisation* on the other hand, is related to the stratification of alternative behavioural aspects of the entire system and it is motivated by the need to manage the overall information complexity. The production system may be viewed as an information system and thus notions of complexity are naturally associated with it [MacF, 2].

Hierarchisation has to do with identification of design and operational tasks, as well as reduction of externally perceived complexity to manageable levels of the higher layers. At the highest level of the hierarchy, we perceive and describe the overall process as a conceptual activity; at this level we have the lowest complexity, as far as description of the process behaviour. At the next level down we perceive the process in terms of set of interacting systems sections, each performing functions which interact in such a way as to give an object -the conceptual model- at the higher level of description. At the next level of description we will use similar amounts of complexity in describing the particular functions of different operational functions. At the next level down we are concerned with specification of desired operational functions for each process and so on we can move down to operation of subprocesses with certain criteria and further down to dynamic performance etc. In an effectively functioning hierarchy, the interactions between systems, or subystems at lower level is such as to create a reduced level of complexity at the level perceived above [MacF, 2].The hierarchisation implies a reduction of externally perceived complexity successfully, as we proceed up the hierarchy till the top level.

At the level of the process we have the richest possible model in terms of signals, data, and full dynamic models. Then, as we move up in the hierarchy, the corresponding models become simpler, but also more general since they then refer not to a unit but to a section of the system etc. The operation of extraction of the simpler models is some form of projection, whereas wider scale models are obtained by using plant topology and aggregations. The mechanisms of projection depend on the particular function the model addresses and they are not always well understood. These models, although of different nature and scope, are related, since they describe sections, or aspects of the same process. Dynamic properties of subsystems are reflected on simpler, but wider area models, although not in a straight forward way. This is what we may refer to as *Embedding of Function Models* (EFM).

An agenda for long term research is to develop a systemic approach that aims at:

- Providing a conceptual framework that explains the interrelationships between the different aspects - problems of the integrated Technical Operations hierarchy,
- Select the appropriate modelling tools that describe the particular problems and provide qualitative and quantitative means enabling the understanding of hierarchical nesting and system properties emerging at different levels,
- Study control, optimisation and state assessment problems in the integrated overall operations set up; this involves top-down control and bottom-up diagnostics-prognostics issues,
- Understand the link between operational requirements and process design criteria,
- Develop criteria, modelling concepts and methodologies that explain the evolution of physical system structure through the different stages of the cascade design process,
- Formulate methodology, procedures which may guide design along paths, which guarantee the formulation of systems with desirable characteristics,
- Develop methodologies for redesigning existing systems to meet new operational requirements,

- Explore the system aspects of data merging and transformations which may provide useful tools that may support the operational and design aspects of integration.

6.2 Global Control and Measurement and their Architectures

The hierarchical model of the Overall Process Operations involves processes of different nature expressing functionalities of the problem. Such processes are interlinked and each one of them is characterised by a different nature model. We adopt an input-output description of each of the subprocesses, with an internal state expressing the variables involved in the particular process and inputs, outputs expressing the linking with other processes. Such a model is generic enough to be used for all functionalities and can take a specific form determined by the nature of the specific process. We may adopt a generic description for the various functions as shown in Figure 6.3, where $u_1,\ldots u_p$ denote independent manipulated variables of the function model, called system inputs; $y_1,\ldots, y_m$ are the independent controlled variables that can be measured and they are called the system outputs $d_1,\ldots, d_q$ are the exogenous variables which cannot be manipulated, but they express the influence of external to the particular function variables and they are called disturbances. A model describing the relationships between the vectors, $\underline{u}$, $\underline{d}$, $\underline{y}$ is expressed as $y = H(\underline{u};\ \underline{d})$ where H expresses a relationship between the relevant variables, and it is called an input – output model. The construction of such a model is a major problem and involves a number of issues which may be classified as:

(i) For the given function establish a conceptual model for its role in the operational hierarchy.

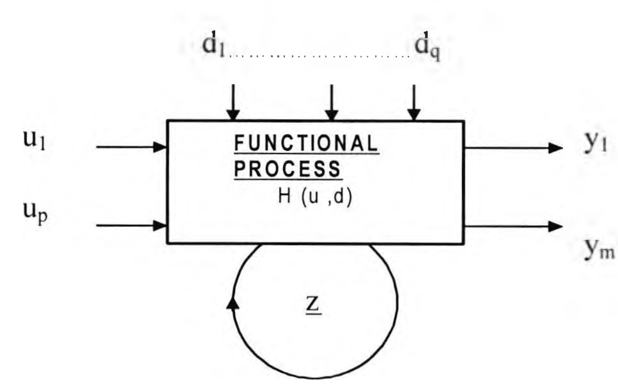


Figure 6.3: Generic Function Model with Internal Structure

(ii) Define the vector of internal variable $\underline{z}$ associated with a given problem and determine its relationships to input, output vectors by using any physical insight that we may possess about the functioning of the internal mechanism.

(iii) Establish the relationships that exist between the alternative vectors $\underline{z}$ associated with problems of the operational hierarchy.

(iv) Define the appropriate formal model (using inductive and/or deductive modelling methodologies), to provide an adequate description for the H functional model.

The above generic steps are providing an approach, which however, involves many detailed modelling tasks. Typical problems here are issues such as classification of variables to inputs, outputs, disturbances, internal variables [Karc, 4], specification of formal description for H, definition of performance indices etc. When the classification of internal variables is completed, the key issue is the establishment of relationships between such variables; such relationships may be classified to implicit and explicit (oriented) forms respectively as:

$$M(u,y,d;z)\begin{cases} F(\underline{z},\underline{u},\underline{d})=0 & \text{Implicit} \quad (1) \\ y=G(\underline{z},\underline{u},\underline{d}) & \text{Oriented} \quad (2) \end{cases}$$

The nature of variables and the type of problem under consideration determines the nature of the F, G functions. This model structure also shows how constraints F(z,u,d) can be propagated from higher to lower levels. The selection of $\underline{z}$ implies that the modelling exercise, expressed as an attempt to specify F, G includes the modelling of the interface of higher level operation to the level defined by $\underline{z}$. The model M ($\underline{u}$, $\underline{y}$, $\underline{d}$; $\underline{z}$) in (1), (2) will be referred to as a $\underline{z}$-*stage model*. The selection of the operational stage (i.e., logistics, scheduling, steady state optimisation, quality control, dynamic process control, state assessment, maintenance etc.) determines the nature of the internal vector $\underline{z}$ and thus also of the corresponding $\underline{z}$-stage model. The dimensionality and nature of $\underline{z}$ depends on the particular functionality under consideration. Describing the relationship between different stages internal vectors is an important problem and implies an understanding of interfaces between functions; this is closely related to the problem mentioned before Hierarchical Nesting, or Embedding of Function models.

The Control and Information Shell in the Hierarchical Structure

The fundamental shell of this hierarchical nesting architecture is described below.

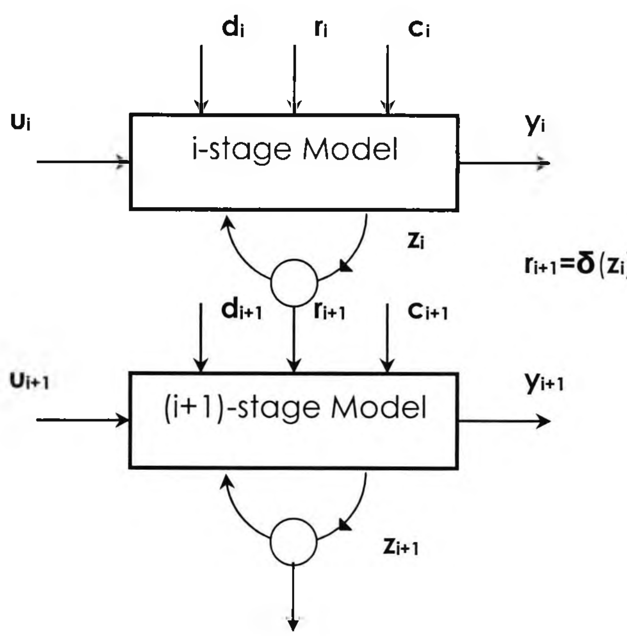


Figure 6.4: Nesting of models in the operational Hierarchy

Note that the vector reference image $\underline{r}_{i+1}$ of operational objectives of the (i+1)-stage is defined as a function of the ith-stage internal vector $\underline{z}_i$. A scheme such as the one described above is general and can be used to describe the meaning of the hierarchical nesting. Furthermore, such a scheme can be extended to describe relations between models associated with functions at the same level of the hierarchy, extend upwards to more general view of the problem and downwards to the area of the physical process.

The fact that each stage model in the hierarchy is of different nature than the others makes the overall system of hybrid nature. It is clear that the theory of hybrid systems [Ants], [Ner] is crucial in the study of the control problems defined on the overall process hierarchy. Most of the work in the hybrid area has been concerned with two types of models; the characteristic of the present paradigm is that we have a multilayer hybrid structure.

The Key Control and Diagnosis Problems

On this multilayer structure we have two fundamental problems:

- Global Controllability Problem
- Global Observability Problem

The first refers to the fundamental issue of whether a high level objective (possibly generated as the solution of a decision problem at a high level) can be realised within the existing constraints at each of the levels in the hierarchy and finally at lowest level, where we have the physical process (production stage). This is a problem of *global controllability*, or alternatively may be seen as a problem of *realisation of high level objectives* throughout the hierarchy. This open problem requires development of a multilevel hybrid theory and it can take different forms, according to the nature of the particular stage model. The Global Controllability problem described above is central in the development of *top-down* approaches in the study of hierarchical organisations, such as the process operations considered here.

The second problem refers to the property of being able to observe certain aspects of behaviour of the production layer of the hierarchy by appropriate measurements, or estimation subprocesses which are built in the overall scheme. This is a *global observability* property and it is related to the ability to define *model based diagnostics* that can predict, evaluate certain aspects of the overall behaviour of the manufacturing process. It is assumed that the observer has access to the information contained at all stages of the model apart from the production layer, where only external measurements provide the available information. The Global Observability problem is intimately linked to the *bottom-up* approach in the study of hierarchical organisations. The measurements, diagnostics defined on the physical process are used to construct the specific property functional models (quality, overall state assessment etc.) and thus global observability (ability to observe all types of behaviour of the physical process) is linked to the quality of the respective functional model.

Integration of Operations requires study of fundamental problems such as Functional Model, Global Derivation and Interfacing, Model Embedding of Function Models, Global Controllability and Global Observability of the Process Hierarchy. These problems have links between themselves and establishing such links is also a challenging problem that may be referred to as *Process Operations Design* (POD). These problems have been hardly addressed from the Systems viewpoint so far, with the only exception the recent work on hybrid systems, which covers only partially some of the issues raised in the above problems. Of course, Process operations are based always on a physical system, process. Establishing the links between Operational criteria (desirable goals) and Engineering Design Objectives – criteria, is a major challenge and it is referred to as *Operations–Design Interface* (ODI) problem. When operational objectives cannot be realised on the existing physical process, then the problem of Process Redesign arises and this is a problem that addresses together problems of Process Operations and Integration of Design simultaneously and can be considered within the current framework.

6.3 Data and general aspects of operational integration

The problem of Systems integration has a technical dimension, expressed by issues of Process Operations, Design and IT, as well as general aspects dealing with the Human support of the integrated framework and involving education and formation of interdisciplinary teams. Here we examine certain aspects of the IT framework, which have a system context and discuss briefly the educational requirements stemming from the needs to support the new integrating, multidisciplinary activities.

The development of methodology and techniques for integrating operations has also a software, information and data dimension such aspects support the local modelling, analysis and decision making and the problem of their integration is crucial for the design of integrated IT support for the Operations problem. The problem of software integration has dominated the overall area for many years and essentially is a problem of adopting common standards. Integrating data structures and information is, however, a more difficult problem since data structures for each of the production functions represent "primitive forms" of models, which support the functional modelling, and thus obey the same rules of connectivities and interrelationships coming from the production organisation. The interaction between data bases supporting the individual activities is thus a problem that has a systemic dimension and couples two key subproblems:

(SDP.1) Representation and Modelling of Lifecycle of Data Structures for individual Processes.
(SDP.2) Interconnection and Organisation of Data Structures of interacting Processes.

The first area deals with the study of data structures associated with a particular operational activity and aims to provide a system based approach that explains the process of transformations in the data and provides a suitable framework for database integration. The need for such work was motivated by the requirements of business process modelling where, the continuously changing and restructuring business can only be modelling where, the continuously changing and restructuring business can only be modelled by a dynamic system supporting life cycles of its components [ESPRIT]. The study of such problems reveals the existence of a new class of systems based on primitive objects and their relations, where transformations linked to their lifecycle are time and event driven. The distinguishing feature in this form of systems is that the notion of state space (attributes set) is not of fixed dimension, but may vary as time evolves and events occur and relations, connectivities, also follow a similar pattern. Describing the lifecycle of data linked to a specific functionality, provides the most primitive form of a model for this function; such models, provide the basis for the development of advanced behavioural models for the corresponding function. To handle the problems of this challenging area, a very general new class of systems has been recently introduce referred to as *Object Dynamic Systems* (ODS) [Leon]. The development of ODS was based on the time and event driven evolution concepts of classical systems theory, the structured lifecycle approach of Object Oriented methodology and the experience of conceptual modelling, data modelling, systems analysis and database design. This new family of systems belongs to the general area of evolving systems and brings a new dimension through the variability of dimension of state and respective relationships associated with the primitive element, the object. The development of the ODS framework is a major challenge in (SDP.1) area.

The family of Object Dynamic Systems belongs to the general cluster of Structure Evolving Systems. Their distinguishing feature is that their basic cell, the object, is characterised by a variation in the dimensionality of its state and by variability, evolution of the relations associated with it. Furthermore, the definition of the object and that of the associated relations are intimately linked. From this viewpoint, the modelling and dynamics of data structures is a paradigm that is closely linked to that of Business Processes. In fact, the life cycle of Business Operations involves a continuous structure modification (existing connectivities) and structure growth (development of new activities), as well as parameter changes. The example of Business Process Reengineering is a typical manifestation of this evolution of structure. The general experience from the technical Structure Evolution Systems, including that of ODS is expected to provide a new insight to the study of Modelling and Dynamics.

We have a variety of functions which are based on the given physical process and they are naturally interlinked, although not always in clear way. This implies that their respective databases are interacting as dynamic processes and this makes the problem of their integration described in (SDP.2) an important one. A systems framework based on ODS seems to be a natural way to address the problem of interconnecting databases, since the methodology of interconnected and organised systems may provide a useful avenue for study of such problems. Integration of data structures has also an alternative meaning, stemming from the need to link the different families of models in the overall hierarchy.

The linking and relationships between the different types of models (including data models) may be summarised in Figure 6.5 below.

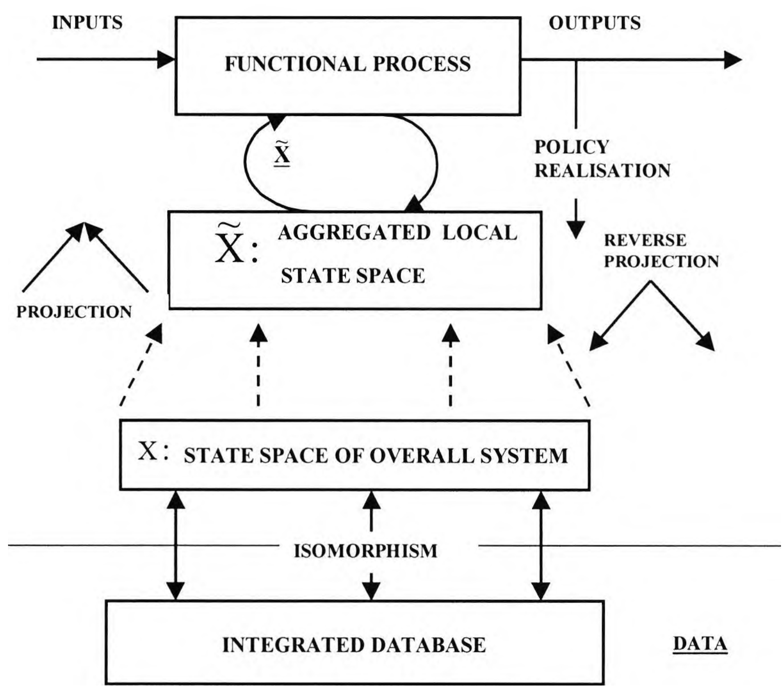


Figure 6.5: Abstract Functional Model and Dependencies

The modelling of individual functions is a process that has many more additional features than those described above. If the vector of internal variables is a state vector (independece of associated attibutes) $\underline{\tilde{x}}$, then its state space $\tilde{X}$ is linked to the overall system state space X in terms of projection (aggregation). The overall state space X of the system corresponds to all variables associated with the Overall System and expresses the event and time evolution of them. Defining $\tilde{X}$ and X involves modelling and definition of appropriate measurement schemes; such measurements are not only physical, but they may be linked to specific metrics associated with the functional process. The time evolution of overall process generates data. The integrated database of the system contains all measurable information about the time and event evolution of X and additional information (issues related to physical process etc.). Specifying the nature of such relations is not a simple problem; ideally, when all $\tilde{X}$ is made up from measurable variables, this relationship (Data→ $\tilde{X}$) is a projection. Creation of an integrated database that supports all processes and functions is a major challenge that cannot be addressed without understanding the more general aspects of integration of operations. Such knowledge is essential for the exact specification of links between individual databases. Issues of aggregation of data due to the projections involved in the operational hierarchy are also important, since they introduce additional dependencies between data structures at the different levels of the hierarchy.

Systems Modelling Measurement and Control Problems in the Integrated Hierarchy

In summary, the area of Integrating Operations involves a number of fundamental problems of the Systems, Modelling, Control and Measurement type which may be described as:

(O.P.1) Formulation of individual Process functionalities as dynamic decision making problems (particular problem aspects).

(O.P.2) Study of alternative forms of organisation of the Overall (extended) Process Operations and Business environment.

(O.P.3) Multimodelling aspects of the Integrated Extended Operations hierarchy and multilevel Hybrid Systems.

(O.P.4) Global Controllability of the Integrated Extended Hierarchy and realisation of policies, strategies.

(O.P.5) Global Observability of production process and Model based Diagnostics.

(O.P.6) Integrating design aspects of alternative process operations.

(O.P.7) Interfacing Operational issues and Engineering Design of the production process.

(O.P.1) involves the formulation of individual function studies in the standard control framework, whereas (O.P.2) deals with the alternative forms of organisation, rather than the traditional Hierarchy adopted here. The areas (O.P.3)-(O.P.7) have been already discussed. The term "extended operations" refers to that in the traditional manufacturing production operations we also include those linked to the business environment.

6.4 Overall system design and the problem of its integration

6.4.1 Description of the problem area and problem clusters

Integrated Design is an area that has been addressed in many branches of engineering. The greatest effort to develop some general methodology has been in the areas of Chemical Processes and Aerospace. Within the area of overall design of chemical processes, issues related to integrated design have been addressed in a number of areas [Perk], [Mor, 1]. Such problems may be defined on the conceptual level first, using generic classes of models and then addressed in more concrete forms. Here we consider the nature of these problems. The most central from them are:

(i) **Evaluation of Process Flowsheets** with operability, stability, controllability, as well as economic criteria (depending on the case).

(ii) **Steady State Process Optimisation**

(iii) **Selection of Control Structures**, sensors and actuators

(iv) **Evaluation of the overall performance** in terms of system reliability and economy.

(v) **Advanced Control System Design**.

For large dimension problems (systems with many inputs, outputs, internal variables) we have additional problems arising due to the large dimensions and the difficulties in computations, as well as coping with many design objectives simultaneously.

Further areas of interest for such cases are:

(vi) **Process decomposition**. Process decomposition is the reduction of a large problem into a sequence of smaller problems at the expense of having to deal with the co-ordination of the sequence of these subproblems.

(vii) **Decomposition into unit goal**. The decomposition of operations of each subsystems (unit) into Specific Unit Goals is considered, which in turn have to be co-ordinated. Interactions between process units introduce additional complications.

(viii) Sequencing of the design process. The decomposition of operations of each subsystems (unit) into Specific Unit Goals is considered, which in turn have to be co-ordinated. Interactions between process units introduce additional complications.

These problems are not trivial since the goals of each unit are not specified a priori and the relations between the goals and the respective nodes of the decomposition are also unspecified.

6.4.2 Cascade nature of the design process

The sequencing of the design is the result of the Process Design Decomposition, which in turn is the consequence of division of the design process and development of specializations. Thus, having effectively decomposed the plant into segments that may be treated independently, or having divided the design process into separate areas undertaken by different specializations, we have now to co-ordinate the individual goals into a sequence that involves the plant as a whole. A simplified scheme of the overall decomposition of design (in practice we have many feedback loops involved) is shown below:

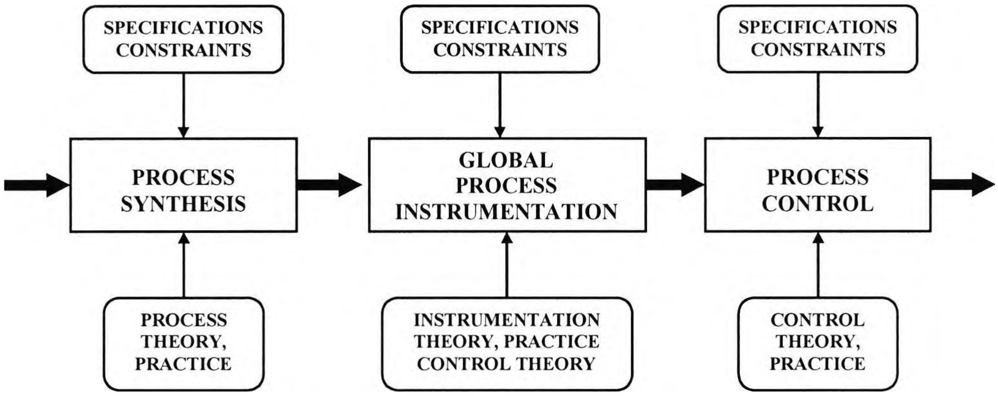


Figure 6.6: Simplified diagrams of main design stages

The cascade nature of this design procedure (with many feedback loops representing iterations) has a number of dominant characteristics. In fact, the main inputs at each design stage are the special skills, body of knowledge, theory, local objectives and specifications and the final result of the previous design stage expressed into the form of a model. Secondary inputs expressing transfer of information between different stages (a manifestation of the iterative nature of the design process) is mostly empirical, or expressing simple facts coming out of simulations. For most cases, there is no “a priori” knowledge of the implication of decisions taken on early design stages on the nature of possible results that may be achieved at the successive local stages. Defining “a priori” a tight set of specifications for every local design stage is also difficult, since what is best with local criteria, is not necessarily best when we look at the overall result. The cascade design procedure is dynamic, in the sense that what is feasible to achieve at a given design stage is influenced by the decisions taken at the previous design stages. The trial and error procedure may be essential for small corrections and changes, but major alterations are time consuming, uneconomic and very frequently not possible. A body of knowledge, theory, techniques that can guide the overall design process taking into account both local and global criteria may be referred to as “Global Integration Methodology” (GIM). The holistic nature of the task implies that system theory and modelling are central in the effort to build GIM. Given that the easy or difficult nature of the final control problems is usually the overall evaluator of the design that makes Control Theory and Design also a crucial ingredient in the effort to develop GIM.

Systems, Modelling, and Control Issues in Process Synthesis

Process synthesis is an act of determining the optimal interconnection of processing units, as well as the optimal type and design of the units within a process system. The structure of the system and the performance of the process units are not determined uniquely by the performance specifications. The task is then to select a particular system out of the large number of alternatives which meet the specific performance specifications. Some of the basic problems in Process synthesis are [Mor, 1]:

(i) The **Representation Problem**. It deals with the question of whether a representation can be developed, which is rich enough to allow all alternatives to be included without redundancy.

(ii) The **Evaluation Problem**. It deals with the question of whether the design alternatives can be evaluated effectively, so they may be compared.

(iii) The **Strategy Problem**. It deals with whether it is possible to locate quickly the better alternatives without totally enumerating all options.

Problems (i) and (iii) heavily depend on the specific applications areas, they are more towards the directions of Chemical Engineering (or the relevant engineering discipline). Systems and Control provide generic results which can be used to formulate alternative approaches based on generic concepts. Some of the emerging theoretic issues are considered next for each of the basic domains.

The Representation Problem

The key issue in the Representation area is the generation of process flowsheets, or required structure of interconnections and this is based on the specifics of the application area. There exist however degrees of freedom in any engineering design and it is this space we would like to explore with Systems and Control results.

Three important classes of problems linked to this area are:

(a) Variable Complexity Modelling (VCM)
(b) Feedback Representation of Process Synthesis (FRPS)
(c) Structure Evolving Systems (SES)

The first family of problems (VCM) is linked to the general procedure in design where we have a fixed interconnection structure but at the Early Stages we require *simple modelling* for subprocesses and Physical interconnections, whereas at the Late Stages of design *more detailed*, full dynamics models are required for subprocesses and physical interconnection structures. The study of such problems requires the development of a framework that permits the transition from simple graphs to full dynamic models and allows study of Systems and Control properties in a unifying way. Here, we essentially observe an evolution of the given structure of the system in the *design stage time* axis and this problem expresses the *Early-Late Design Variability of Model Complexity* [Karc, 5], [Karc, 6], where study of structure and property evolution becomes the key objective.

The generation of overall system models from the general graph (scalar, or vector) and the available for the different design stages models for the subprocesses in a concise and uniform way is the subject of the third area. The problems considered here describe a new family of Systems, where their structure-and thus properties-evolves as a function of the design time; we shall refer to such systems as *Design-Time Evolutionary Systems* (D-TES).

The second problem area (Feedback Representation Process Synthesis) deals with the problem of representing the overall interconnection of sub-processes as an equivalent feedback design problem for which traditional Control Theoretic Tools may then be used. For the case of linear systems this has been introduced in [Karc, 6], but for more general subprocess model families (nonlinear etc.) the problem is still open. Transforming synthesis problems to standard representations, such as the feedback, or multiport representation allows the use of existing methodologies; in fact, the equivalent feedback configuration allows the design of the interconnection graph in terms of feedback theory, where as the multiport representation permits the use of network synthesis techniques

The third area relates to the notion of *Graph Structure Evolving Systems* (SES). This also emerges in a different form in the context of transition from conceptual to qualitative and then quantitative models of the process synthesis area, as well as the area of redesign, modification of existing process flowsheets. The distinct feature here is that we start with an elementary system cell and then progressively develop the overall structure by adding new subprocesses and building appropriate interconnections. In this sense, the basic cell grows and eventually leads to the final design. This is a different form of evolution than those described above and it is reminiscent of similar processes in biology, crystalography, data structures and business modelling.

The Evaluation Problem

The variety of alternative structures, process flowsheets, generated have to be evaluated with a variety of criteria. There is a range of important issues which need consideration in the evaluation stage of emerging conceptual designs. These issues are:

(1) **Flexibility:** Is defined as the ability of the system to handle a new situation at steady-state and thus express the ability to operate at different steady states.
(2) **Switchability:** Considers ability of a plant to be moved from one steady state operating point to another. This also involves start up and shut down of the process.
(3) **Process Controllability:** Is the "best" dynamic performance (set point following and disturbance rejection) achievable for a system under closed loop control.
(4) **Safety:** Examines the hazards that may be involved with particular designs and using process dependent heuristics provides a classification.

It is worth noting that *Process Controllability* is a much more general notion than the traditional system controllability. Note also, that Flexibility depends mainly on the structure of the process, whereas switchability and controllability depend on the system structure, as well as the selected control structure. It is also worth noting that controllability requires flexibility. It is worth noting that in the study of the above clusters of problems the required complexity of the model is a significant issue. The **Prediction of Full Model System Properties** using information provided by simple models is a major challenge.

Systems and Control Aspects of Instrumentation

The instrumentation of a process that is the, selection of measurement variables (outputs), and actuation variables (inputs) has a "micro" (local), as well as a "macro" (global) aspect. The "micro" role of instrumentation has been well developed [Finkel] and deals with the problem of measurement or implementation of action upon given physical variables; instrumentation theory and practice deals almost exclusively with the latter problems.

An overview of this important area is given below:

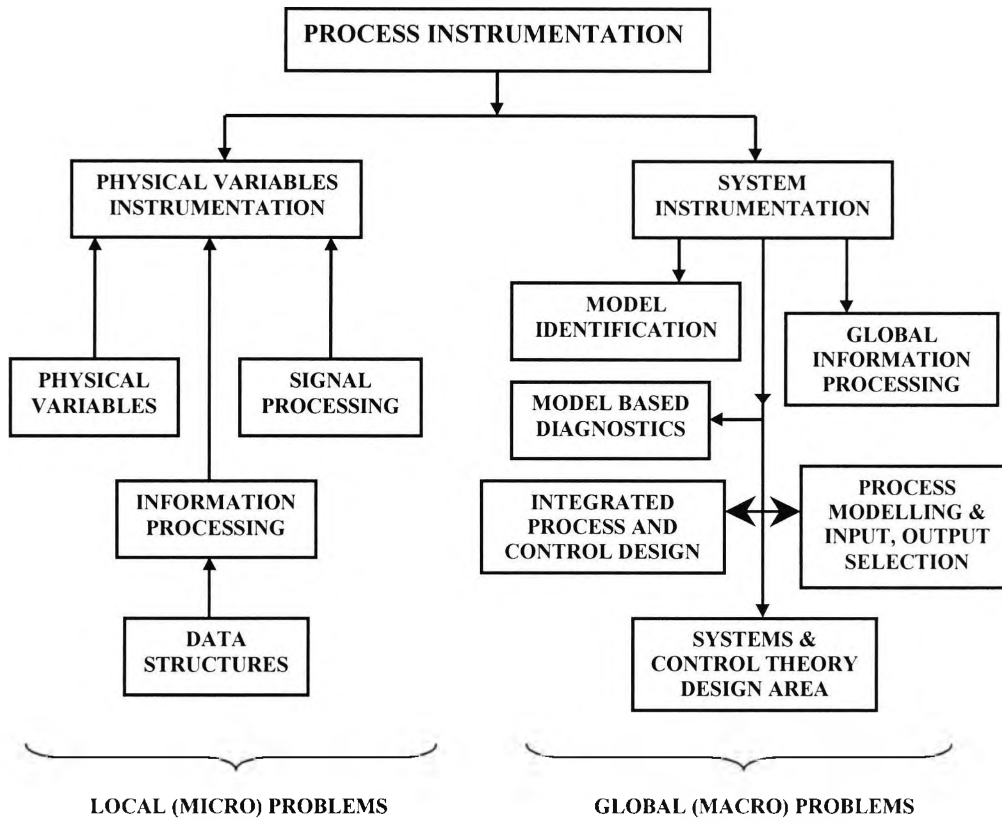


Figure 6.7: The area of Global Process Instrumentation

The "macro" aspects [Karc, 4] of instrumentation stem from that designing an instrumentation scheme for a given process (classification and selection of input and output variables) expresses the attempt of the "observer" (designer) to build bridges with the "internal mechanism" of the process in order to observe it and/or act upon it. What is considered as the final system, on which Control System Design is to be performed, is the object obtained by the interaction of the "internal mechanism" and the specification of the overall instrumentation scheme. Difficulties in control of the final system may be assessed in terms of certain structural characteristics of the final system model. These structural characteristics are formed through the various stages, where the design goes through; the stage of instrumentation is one of the stages of the overall design and has a decisive affect on the further shaping, evolution of the structure delivered at the end of process synthesis. Here we examine the fundamental system aspects of the Instrumentation process which refer to its model shaping role. This area will be referred to as *Systems*, or *Global Instrumentation.*

Global instrumentation is a problem of selection of inputs and outputs and its study revolves around the study of four fundamental problems which are:

- **Model Orientation Problems** (MOP),
- **Local- Global Structure Problems** (LGSP),
- **Model Projection Problems** (MPP) and
- **Model Expansion Problems** (MEP)

The above mentioned problems are issues of the general area of Control Theory and Design, but they have not been properly recognised and addressed there. These problems are essential in the effort to develop conceptual and design tools for assisting the "good" shaping of the system model as a result of the instrumentation process. The distinguishing feature of instrumentation as far as model shaping is that it acts on the shaping of the input-output structure, rather than the interconnection graph, shaped by process synthesis. Such problems may be addressed in a systematic way within the framework of linear systems etc. However, they are generic conceptual problems and their nature is briefly described below.

(i) Model Orientation Problems: The classification of system variables as inputs and outputs is referred to as model orientation. In many systems, the orientation is not known, or that depending on the use of the system the orientation changes. Question such as, when is a set of variables implied, or not anticipated by another, or when it is free, have to be answered, if model orientation criteria based on the nature of the process are to be delivered; the specific use of the system may provide additional model orientation criteria. It may happen, that the above two types of criteria do not provide a unique solution to model orientation; not that for each alternative orientation we have a different i/o model and thus criteria based on the resulting model characteristics have to be used for the final evaluation, selection. Problems of this type have been recently realised [Karc, 7] and their solution, as far as techniques is in its early stages. The implicit description may be used and the overall problem is then the partitioning of the $\underline{\xi}$ vector [Karc, 6, 7] such that the resulting vector has clearly identifiable inputs, outputs and independent internal variables, whereas the resulting model has a given structure. Note that in such problems both the dimensionality of input, output vectors, as well as the overall partitioning of the implicit vectors $\underline{\xi}$ is considered as design parameters. The overall objective may be expressed as a desirable system property (regular, normal system description), or a model with desirable structural characteristics (system invariant structure). The essence of the model evolution here is that from an implicit form we transcend to an oriented form and this implies a reorganisation of the internal model structure with corresponding effects on the resulting system properties.

(ii) Local-Global Structure Problems: A large family of problems dealing with the establishment of relationships between subsystem model and composite model properties is considered here and it is referred to under this general heading. Crucial essence of the completeness assumption is that all variables from the subsystem which feed through to the other subsystems in the interconnected structure are measured and that for every subsystem the number of exogenous independent variables is equal to the dimension of the vector that feeds through from the other subsystems. It has been shown [Karc, 6] that under the completeness assumption the family of system properties are the same for both the aggregate and composite system. Deviations from completeness, i.e., not all variables feeding through the other subsystems are measured and fewer control variables than those needed for completeness, imply that the interconnection structure, represented by the matrix F [Karc, 6, 7] starts to play a crucial role in shaping the structural properties of the resulting composite systems. The selection of input, output variables at the subsystems level may thus be treated as a design problem to either retain, or modify structural properties of the composite system by aiming for completeness (if aggregate properties are desirable), or appropriate deviations from completeness in a manner that alters appropriately the composite system properties. The process of model evolution in this context is that shaping of the local input, output structures affects the overall dimensionality of transmittances of the graph (graph dimensionality problem) and thus affects the resulting structure and properties of the composite system.

(iii) Model Projection Problems: The number of potential control variables p and potential measurements q, which ideally may be used, is quite large in many engineering designs. In an ideal design, unconstrained by resources and effort all possible inputs and outputs should be used; economic and technical reasons, however, force us frequently to select a subset of the potential inputs, outputs as effective, operational inputs, outputs. Engineering specifications and past experience with similar designs provide some guidance in how to select the effective ℓ-inputs and effective m-outputs, but they do not specify a solution uniquely. Developing criteria and techniques for selection of an effective input, output scheme, as projections of the extended input, output vectors respectively, is what we call Model Projection Problems (MPP). The basic questions which are handled within the general type of problems listed above are:

- Define the lowest bounds for the number of effective inputs, outputs, which are needed for certain control scheme, or family of alternative control schemes.
- Define the best location of effective inputs, outputs, as well as, the structure of actuator, sensor maps, which may guarantee structural controllability, and observability.
  - Evaluation of degree of dependence, independence of given input, output instrumentation schemes and its implications on process controllability, observability.
  - Evaluation of effect of a selected input, output scheme on the control quality, characteristics of the final system and selection of "best" schemes for easy, reliable control.

In this area model structure evolution emerges as the process of obtaining new forms by reducing a larger original structure. In this sense, projection tends to aggregate, reduce an original model to a smaller dimension with desirable properties. A special problem within this area that has been studied so far is the zero assignment problem [Kouv], [Karc, 8].

(iv) Model Expansion Problems: Defining input test signals and corresponding output measurements, is an integral part of the identification, modelling exercise. Defining input output schemes with the aim to identify, (or improve) a system model, or reconstruct an unmeasured internal variable, is what we mean by Model Expansion Problems. Questions related to the nature of test signals, or properties of the measured signals are also important here, on top of the more general questions related to the structure of the i/o scheme; the latter gives a distinct signal processing flavour to MEP. Some distinct problem areas are:

(a) Additional Measurements for Estimation of Variables: Frequently in process control, some important variables are not available for measurement. Secondary measurements have to be selected and used in conjunction with estimators to infer the value of unmeasurable variables.

(b) Input, Output schemes for System identification: The selection of input test signals and output measurements is an integral part of the setting up of model identification experiments. In fact, the identified model is always a function of the way the system if excited and observed, i.e., of the way the system is embedded in its experimental environmental. The study of effect of location of the group of excitation signals and corresponding group of extracted measurements on the identification problem is a key problem here.

Model expansion is one more example of model structure evolution, where additional inputs, outputs help a system model to grow to a more full representation of the existing system. The problems in this area express an alternative form of evolution of structure and properties of the model by manipulation of the input-output, external structure.

6.5 Conclusions

This chapter has provided an overview of those technological aspects of systems integration which have a Systems, Modelling and Control dimension and are derived on the basis of a conceptual analysis. It has been mainly preoccupied with the issues of Integrating

Operations and Design aspects for industrial processes and in doing so it has specified a range of new open issues of the Systems and Control type as well as new families of systems which are intimately linked to the new applications paradigm. From this viewpoint, it provides a very challenging agenda for research in the Systems, Modelling and Control area. The two central themes which emerge are the needs for control and measurement in a multimodelling context, which makes multilevel hybrid theory a key area, and the development of theory and methodology for the different types of Evolving Systems. Such systems emerge in many different areas and with variability in their statement and form and require a fundamentally different approach and methodology to those of the traditional non-evolving type of systems.

Chapter 7

Quantitative Analysis using Real Options Analysis

Conceptualisation of the system is the first stage of the modelling process that is based on the "general knowledge" about the system and relies on information from the past. It produces a structure, a framework that has to be substantiated and populated by "current" and "specific" knowledge. This is where the interaction of conceptualisation and the "measurement process" begins. It is generally accepted by researchers that an effort to combine the conceptualisation of the "Method" with the "Measurement" outcomes is always needed. This thesis strongly supports that the two are not only inseparable, but complementary as far as modelling is concerned. The conceptualisation provides the fundamentals of the structure of the model, whereas measurement comes to provide the means to identify details, fix parameters and validate assumptions. Conceptualisation influences measurements from certain global aspects, such as defining the basics of experiments that have to be set up, the location and type of measurements to be made, etc. Therefore, in this chapter an introduction will be given about the effort made to bridge the two fairly diverse approaches to a problem and thus illustrate with a case study, how the two approaches can work together to give solutions to a modelling problem.

7.1 Bridging Conceptualisation with Quantitative methods: Key Issues

From Wikipedia [Wik], the free encyclopedia, Quantitative research is the numerical representation and manipulation of observations for the purpose of describing and explaining the phenomena that those observations reflect. It is used in a wide variety of natural and social sciences, including physics, biology, psychology, sociology and geology. Quantitative research begins with the collection of statistics, based on real data, observations or questionnaires and definition of experiments that may lead to useful data that enhance our understanding of the model.

Quantitative methods are methods concerned with numbers, functions and anything that is quantifiable. Counting and measuring are common forms of quantitative methods and the result is a number based model or a series of a number based models, often presented in tables, graphs or other forms of statistics. Quantitative research methods were originally developed in the natural sciences to study natural phenomena. Examples of quantitative methods include survey methods, laboratory experiments, formal methods (e.g. econometrics), numerical models, physical modelling and modelling in general [Qual]. They are used in conjunction with Qualitative methods. Using Qualitative methods it is often possible to understand the meaning of the numbers produced by Quantitative methods; using quantitative methods, it is possible to give precise and testable expression to qualitative ideas.

Measurement is the assignment of numbers or other symbols by an objective, empirical process to attributes of objects or events of the real world, in such a way as to describe them [Finkel]. There is also the more general aspect of measurement, referred to as “global” or “system dimension” [Karc,2], which relates to the question of what to measure and where, and addresses the effect of measurement in general on the “shaping” of the system model. Identification is the determination of a mathematical model of a dynamical system or process, based on input and output measurement data of that process and in combination with specific experiments. The scheme of identification is divided into two activities [Eykh]:

- Modelling-deriving structural knowledge of the process under test by using, for example physical laws that hold for the process or system under consideration; and
- Estimation-deriving knowledge about parameters by using measurement data.

Identification is the combination of both these activities. The organized body of knowledge produced by measurement, forms part of Quantitative research.

In Chapter 4 of the thesis we came across for the first time a description of what we define here as quantitative methods, while we were identifying the main tasks of the observer-modeller. We defined as tasks of the observer:

- the definition of the problem area,
- the conceptualisation of that problem,
- *the measurements, observations, experimentations and data collection*
- *the interpretation, formal modelling.*

The last two tasks combined give a very good description of what are quantitative methods. Quantitative research is encountered as part of formal research, but is also sometimes used when conducting *exploratory research.* Through *measurements, observations, experimentations and data collection* the modeller collects all the available information for the problem area-the system- so as to have a complete idea of the main features that need to be modeled. Sometimes, when measurements and observations are simply not enough to collect the needed data, the modeller deals with the problem of defining appropriate ways of exciting the system to get responses, which are rich enough to provide useful knowledge on the system and its parts. This involves designing a sequence of experiment-observations which are independent from each other and each one of them highlights a different aspect of the system features, previously defined as *Design of Experiments Problem* (DEP). Clearly, such a design of experiments-observations requires some form of knowledge about the system and it is thus linked to the System Conceptualisation Process. Selecting measurements or actuation variables and setting up experiments, is an area of the general modelling process that is not governed by a systematic theory and holistic methodology, but it is usually driven by heuristics and past experience.

Then *interpretation of the results of experiment and data, and formal modelling* is an inseparable "component" of the modelling process. The observer will be the cornerstone associating data, observations and measurements with knowledge; this extraction process is inevitably strongly related to the observer, since the observer by his conceptualisation of the system, provides the conceptual model that guides the specific measurement activities. This task also incorporates the further steps from conceptualisation to formal modelling; from definition of key variables, attributes and identification of the relations between objects and the causality of those relations, to the formulation of a more substantiated, formal form of modelling. Here the introduction of the formal language, is made, which gives the specific character to the system and deals with the definition of appropriate machinery that can be either conceptual or formal and describes the given Modelling set. Within this area we have the problems of physical modelling, system identification, hybrid versions between the two, etc. This problem was defined previously as the *Model Construction Problem* (MCP).

Quantitative research uses methods adopted from the physical sciences to ensure objectivity, generalizability and reliability. The strengths of the quantitative paradigm are that its methods produce quantifiable, reliable data that are usually generalizable to some larger population. Furthermore, quantitative models provide the means for testing hypothesis, validate assumptions and suggest methods for improving existing models. This paradigm breaks down when the phenomenon under study is difficult to measure or quantify [Wein]. The greatest weakness of this approach is that it removes the event from its real world setting and ignores the effects of variables that have not been included in the model, sometimes because those effects are difficult to be expressed with variables or quantified in a concrete way.

The development of quantitative methods always assumes a system conceptualisation, which may be of the following two types:

(a) An unstructured or "black-box" conceptualisation.

(b) A structured or "internal" conceptualisation.

Both require specification of variables and design of specific experiments that generate "relevant" data that they may be used for the specification of parameters in the formal model. The following diagram summarises the relationship between the conceptualisation and the development of formal models.

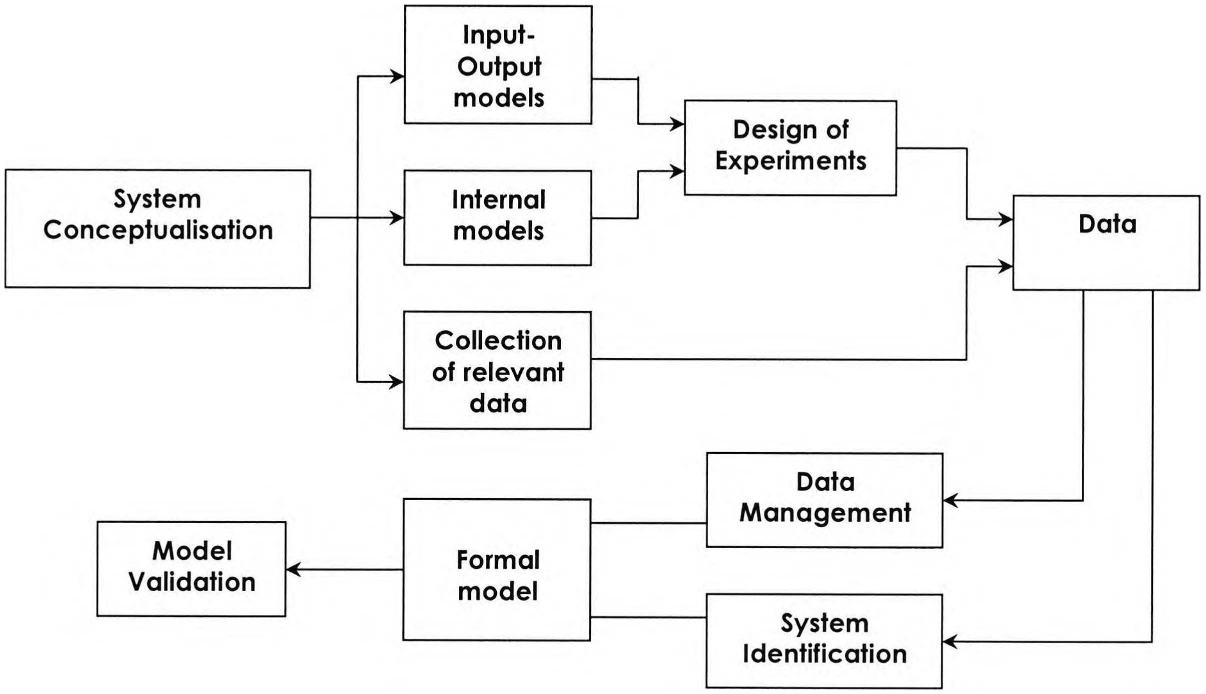


Figure 7.1: Conceptualisation and Formal Model development.

Major issues that arise in the development of formal models which define the building blocks to the formal modelling development are:

(a) Purpose of the model and model structuring.

(b) Design of experiments.

(c) Measurements.

(d) Data Management.

(e) Methods of System Identification.

(f) Model Validation.

Each of the above topics are issues that have taken considerable attention and have already been identified at the end of chapter 4; however, their study is beyond the objectives of the current work and they are not properly examined here. A brief summary of the fundamentals in each area is given below:

(a) Purpose of the model and model structuring. The model incorporates some of the important semantic information about the real world. A model should be developed for a specific purpose (or application) and its validity determined with respect to that purpose. If the purpose of a model is to answer a variety of questions, the validity of the model needs to be determined with respect to each question [Mac]. Several sets of experimental conditions are usually required to define the domain of a model's intended applicability. A model may be valid for one set of experimental conditions and invalid in another. A model is considered valid for a set of experimental conditions if its accuracy is within its acceptable range, which is the amount of accuracy required for the model's intended purpose [Sarg]. Furthermore, such a model should be able to predict behaviour excited under similar conditions to those covered by the original experiments.

(b) Design of Experiments (or DOE). An experiment is a set of actions, stimuli, performed in the context of obtaining observations and responses from the system we are interested in, to support (or falsify) a hypothesis concerning that system. DOE is typically applied in a constructive way; that is, one starts with a simple design and estimates a simple meta-model. For example, first use of a design to identify important (main) effects and to see if there are possible interactions. Only if the estimation results indicate other effects, a more complicated design is introduced. This approach is chosen to minimize the amount of work required [Groen]. Analysts spent most of their time on developing the System Dynamics model, and little time on the sensitivity analysis of their model. Nevertheless, it is important to answer questions such as: what are the effects of changing input values; are there interactions among inputs? This dynamic analysis uses simulation to check the results with those predicted from deep knowledge. Simulation involves the software or hardware realisation of the derived, formal model and its excitation by various inputs and initial conditions, provides knowledge for the predicted behaviour of the model. It is the comparison of this predicted behaviour with the real behaviour as this is derived by experimentation that provides an evaluation of the quality of the derived model. Experimentation, however, requires an appropriate *design and analysis*, if reliable results are desired. Experimental design in the "hard" sciences (sciences such as physics, chemistry, etc.) tends to focus on the elimination of extraneous effects, while experimental design in the "soft" sciences (social sciences) focuses more on the problems of external validity, often through the use of statistical methods. Occasionally events occur naturally from which scientific evidence can be drawn, which is the basis for natural experiments. In such cases the problem of the scientist is to evaluate the natural "design" [Wik]. Note that experiments with *real* systems have been frequently subjected to the design and analysis techniques developed in the field of mathematical statistics. In the 1930s Fisher focussed on agricultural experiments. Since the 1950s Box concentrated on chemical experimentation. Nowadays Taguchi's designs are very popular in industrial quality control. *Vital questions* are: what are the efficiency effects of changing input values; which input values give optimal output; which decision rule leads to the best result; are there interactions among inputs, and so on [Kleij].

(c) Measurements. At an early stage of model development, measurement is usually performed at only the basic level, that of classification. Measurement is usually performed by assigning numbers-though this is not a prerequisite for measurement. Assignment of numbers makes possible the application of the concepts and theories of mathematics [Rob]. At a further stage of the modelling process, measurement acts as foundation for developing tools for analyzing statements made in scale values. Measurements are usually based on powerful, well established theories. A major application of measurement theory is to problems of decision making. One of the basic problems of measurement theory is that of representation [Rob]. It is important to be able to state foundational axioms, because we must know under what circumstances certain kind of scales of measurement can be produced. Another basic problem is that of uniqueness [Rob]. As Hays [Hays, 1] points out, one can always perform mathematical operations on numbers. However, the key question is whether, after having performed such operations, one can still deduce true (or better, meaningful) statements about the objects being measured.

(d) **Data Management**. The data necessary for model building, model evaluation and testing, and conducting the model experiments to solve the problem must be adequate and correct. Data are needed for three purposes: for building the conceptual model, for validating the model, and for performing experiments with the validated model. To build a conceptual model we must have sufficient data on the problem that can be used in building the model, to develop the mathematical and logical relationships in the model that will allow it to adequately represent the problem identity for its intended purpose, and to test the model's underlying assumptions. In addition, behavioral data is needed on the problem to be used in the operational validity step of comparing the problem entity's behavior with the model's behavior. If these data are not available, high model confidence usually cannot be obtained. Unfortunately, there is not much that can be done to ensure that the data are correct. The best that can be done is to develop good procedures for collecting and maintaining it. When data is inconsistent or there is no pattern available, the theories used are probabilistic theories, which are built around more random data [Rob].

(e) **Methods of Systems identification**. The problem of system identification involves two main tasks. The first is the determination of the required structure of the model (i.e., interconnection topology of the sub processes, required order of dynamics, etc.). The second task deals with the use of experimental data (derived under appropriate experimental conditions) to provide a best fit for the parameters in the structured model that has been defined at the previous stage. Fitting a model within a given structure parameter estimation is in most cases a lesser problem. A basic rule in estimation is not to estimate what you already know. In other word, one should utilize prior knowledge and physical insight about the system when selecting the model structure. It is customary to distinguish between three levels of prior knowledge, which have been color-coded as follows [Sjob. et al]:

- White Box models. This is the case when a model is perfectly known it has been possible to construct it entirely from prior knowledge and physical insight.
- Grey Box models. This is the case when some physical insight is available, but several parameters remain to be determined from observed data. It is useful to consider two sub-cases:
  - Physical Modeling. A model structure can be built on physical ground, which has a certain number of parameters to be estimated from data. This could, for example, be a state-space model of given order and structure.
  - Semi-physical modeling. Physical insight is used to suggest certain nonlinear combinations of measured data signal. These new signals are then subjected to model structures of black box character.
- Black Box models. No physical insight is available or used, but the chosen model structure belongs to families that are known to have good flexibility and have been successful in the past.

A particular problem is to deal with the large number of potentially necessary parameters. This is handled by making the number of used parameters considerably less than the number of offered parameters by regularization, shrinking, or regressor selection [Sjob. et al].

(f) Model Validation. The developers and users of these models, the decision makers using information derived from the results of the models, and people affected by decisions based on such models are all rightly concerned with whether a model and its results are "correct" [Sarg]. This concern is addressed through model verification and validation. Model validation is usually defined to mean "substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model" [Schles. et al]. It ensures that the model meets its intended requirements in terms of the methods employed and the results obtained. The ultimate goal of model *validation* is to make the model useful in the sense that the model addresses the right problem, provides accurate information about the system being modeled, and to makes the model actually used. Unlike physical systems, for which there are well established procedures for model validation, no such guidelines exist for social modeling. In the case of models that contain elements of human decision making, validation becomes a matter of establishing credibility in the model. The task is to establish an argument that the model produces sound insights and sound data based on a wide range of tests and criteria that "stand in" for comparing model results to data from the real system.

Researchers who use primarily quantitative methods are accustomed to relying on certain principles of classical research design with demonstrated validity and reliability, systematic and "non-biased" analysis of data and interpretation of data based on assumptions of objectivity and generalizability [Npi]. An examination of the quantitative and qualitative paradigms will help to identify their strengths and weaknesses and how they complement each other. In many cases researchers fall into one of the two camps, but here we believe that an integral solution to the problem requires use of both approaches; the exploratory-explanatory nature of qualitative research with the interpretation and formalization provided by the quantitative research. The detailed consideration of the techniques and methodologies used in the above areas is beyond the scope of the thesis. We will consider these issues, however, in the context of a case study. The case study, is introduced here, as a means of providing an example of how the two approaches (qualitative and quantitative research) complement each other.

7.1.1 Introduction to the case study: the problem

The purpose of this chapter is to introduce quantitative methods and particularly focus on methods for Economic forecasting. This particular section will give an insight on why the focus of this chapter is methods of Economic modelling-forecasting and specifically Real Options analysis. This discussion is driven by the necessity to give a solution to the problem that will be formulated and further discussed as a case study at a later chapter. At this point it is essential to give a brief introduction of that case study, so as to understand the reasons why this chapter has evolved, the way it has-with Real Options Analysis being the main focus.

For the last ten or twenty years sustainable energy was thought of simply in terms of availability relative to the rate of use. The term originally applied to natural resource situations, where the long term was the focus. Today, it applies to many disciplines, including economic development, environment, food production, energy, and lifestyle. Basically, sustainability refers to doing something with the long term in mind, meeting the needs of the present without compromising the ability of future generations to meet their own needs. Today's decisions are made with a consideration of sustaining our activities into the long term future. In the context of the ethical framework of sustainable development, other aspects are equally important [Yam]. These include environmental effects and the question of wastes, even if they have no environmental effect. Safety is also an issue, as well as the broad and indefinite aspect of maximising the options available to future generations. There are many who see no realistic alternative to pushing Sustainable Development criteria into the front line of energy policy. In the light of concerns about global warming due to human enhancement of the greenhouse effect, there is clearly growing concern about how energy needs are addressed on a sustainable basis [Gov], [HMGov].

The problem described in this thesis is related to the energy sources of the UK today and the energy sources proposed for the future. In a forthcoming chapter, energy and the future of it will be discussed as a main concern for the world. The resources of oil and gas are finishing, alternative sources of energy such as renewable-wind, thermal, etc-cannot be considered as the main source of energy, because the percentage of the effective contribution is very low (in UK only 8%[1] of the total energy production comes from wind) and in addition the Kyoto agreement that initiated in 2005, puts boundaries in the CO_2 emissions and consequently to the sources of energy.

This thesis will not give a solution to the ever ending sources of energy but will give *an insight of the future prices of a specific source of energy*-that, being gas-and will also suggest the appropriate point in time to incorporate an appropriate percentage of another

1 This percentage is an outcome of current policies.

source of energy. Such an action will be inevitably essential so as to conform to the standards and the needs of that time.

While at this point the general path of how the problem was tackled can be given with a plethora of details, at the moment of taking decisions of how to deal with the situation, it was difficult to foresee the development plan. It has to be noted here that at this specific point, where the realization of the existing problem took place and the decision to take action, the first stage of modeling has began. The identification of the problem was in fact the realization that there is a problem with keeping up with the energy demand around the world and the resources available, as well as the restriction of the energy sources that emit large quantities of CO_2 and other environmental unfriendly gases. The variables, constraints and specifications are formed through learning about the problem, thus making every part of the research of vital importance.

The next stage was to identify a way forward, answer the question "what can be done?" This part is the most difficult and of the most significance. Taking the wrong decisions here is jeopardizing the rest of the project since this will be the point where the strategy is laid upon and the appropriate steps to move forward are examined. This strategy can affect the quality of the outcome and thus making the correct choice of method is considered a cornerstone. Due to the fact that the case study will look into future electricity prices (electricity produced by gas) and will identify and propose an alternative, more *economically* viable source of energy, an economic forecasting method is required. After researching the current methodologies and techniques used to implement the model for forecasting, a conclusion was drawn to use *Real Options analysis*. The nature of this problem, thus, prompted the economic analysis that will unfold in this and the following chapters, focusing on specific methodologies that have been selected as the way forward. It is essential here to note how the problem area and what the modeler needs to obtain from the modelling exercise are the deciding factors of the identification of the type of methodology to be used.

Real options analysis is a decision making tool, that is implemented in an already conceptualized model, even a roughly formed one. This model, however, even this rough version of a model, is a prerequisite for the application of Real options analysis. That is, there is no way Real options analysis can be used to build a model as it is not a modelling methodology; it provides a planning and analysis thinking framework that helps asses and analyse a way forward, after the conceptualization of the problem and the formulation of a first model.

For this thesis the realization of a general framework of modeling and the method-a step by step approach- towards the problem will be considered in the qualitative part of the thesis. In the second part of this thesis-the case study-an effort will be made to connect the qualitative part of the problem-the conceptualization of the specific case study-with the quantitative part of the problem, that is, using a "method of measurement" try to numerically conclude the problem. That way the qualitative analysis of the problem will be combined with the quantitative analysis, showing that the union of the two is essential and provides cohesion to the solution of the problem. In the next chapters the analysis of the problem described here, will unfold. The step by step approach will be discussed extensively, showing the flow from realizing the problem to providing a solution.

7.2 Quantitative analysis in Economics

Economics is concerned with modelling at the "micro", or "macro" level of economic life and activity. *Time series analysis* is a major tool for economic modelling and it is concerned with the dynamic consequences of events over time [Ham]. There are two main goals of time series analysis:

- identifying the nature of the phenomenon represented by the sequence of observations, and
- forecasting (predicting future values of the time series variable).

Both of these goals require that the pattern of observed time series data is identified and more or less formally described. Macroeconomists have recently been exploring what the time series data can tell us about the long run behavior of economic variables [Sims].

In economics, a model is a theoretical construct that represents processes by a set of variables and a set of logical and quantitative relationships between them. As in other fields, models are simplified constructions designed to represent complex processes. In general terms, economic models can be a simplification of and abstraction from observed data. It must be stressed here, that econometrics is different from a conceptual model expressing an economic theory. The two main purposes of econometrics are to give empirical content to economic theory and to subject economic theory to potentially falsifying tests. For example, a conceptual model of an economic theory may predict that a given demand curve should slope down. Econometric estimates can either verify or falsify that prediction, and shed light on the magnitude of the effect. Analysts therefore must make a reasoned choice of which variables and which relationships between these variables are relevant and which ways of analysing and presenting this information are useful [Wik].

The use of models in an economic context include [Wik]:

- Forecasting economic activity in a way in which conclusions are logically related to assumptions;
- Proposing economic policy to modify future economic activity;
- Presenting reasoned arguments to politically justify economic policy at the national level, to explain and influence company strategy at the level of the firm, or to provide intelligent advice for household economic decisions at the level of households.
- Planning and allocation, in the case of centrally planned economies, and on a smaller scale in logistics and management of businesses.

The details of model construction vary with type of model and its application, but a generic process can be identified. Generally any modelling process has two steps: generating a model, then checking the model for accuracy (sometimes called diagnostics). The diagnostic step is important because a model is only useful to the extent that it accurately mirrors the relationships that it proports to describe. Creating and diagnosing a model is frequently an iterative process in which the model is modified (and hopefully improved) with each iteration of diagnosis and respecification. Once a satisfactory model is found, it should be double checked by applying it to a different data set.

7.2.1 Economic Forecasting

Economic forecasting is the process of making predictions about the economy as a whole or in part. To be able to predict to some extend the outcome or outcomes of an investment is a precious tool, giving the analyst the chance to make and take opportunities. It is of vital importance at this point to make the distinction between economic models, or any other economic tools, that make predictions or projection in the future and others that take it a step further and with the appropriate handling can give you *options*, to undertake any business decisions. The following steps can assist a more formal forecasting process [Nam]:

Conceptualisation

- Identify the problem.
- Determine how you would use the forecast to deal with the problem.
- Select the particular items you would need to forecast.
- Determine the appropriate time horizon for a forecast that deals with this problem.

Development of formal model

- Research the techniques and theories used by others to forecast this variable in the past.
- Evaluate all possible options and consider pros and cons.
- Use a forecasting model that fits the problem given the specific constraints and limitations.

Use of model

- Make the forecast.
- Interpret the results.
- Make decisions and take action based on results.
- Implement a revolving recap of the forecast versus the actual figures.
- Modify the forecasting model or technique accordingly.

The time frame for a forecast can be short-range, medium-range or long-range. Short-range forecasts typically cover the immediate future and are used to deal with issues of daily or weekly operations of a business. Typically, a short-range forecast would cover a period of one or two months. A medium-range forecast usually covers the period from one to two months to a year and is generally related to something like a yearly production plan. A long-range forecast would be for more than one or two years and is used to plan for the production for new products or the expansion of production capacity, or in the consideration of long-term financing [Nam].

There are four components in the manager's toolkit for *valuing investment opportunities*, which are briefly described in the following sections:

- Payback Rules
- Accounting Rates of Return
- Net Present Values (NPV), and
- Real Options Analysis.

7.2.1.1 Payback rules-Discounted Cash flow

Payback rules ask how many periods management must wait before cumulated cash flows from the project exceed the cost of the investment project. If this number of periods is less than or equal to the firm's benchmark, the project gets the go-ahead. Subsequent cash flows, whether positive or negative, are not factored into the calculation.

In the field of finance, a discounted cash flow is the value of an investment (measured in terms of the cash you will put into and receive from it) adjusted for the time value of money. The future cash flows must be discounted in order to express their present values in order to properly determine the value of a company or project under consideration as a whole. [Wik]

The discounted cash flow for an investment is calculated by estimating the cash you will have to pay out and the cash you think you will receive back. The times that you expect to receive the payments must also be estimated. Each cash transaction must then be discounted by the opportunity cost of capital over the time between now and when you will pay or receive the cash. [Wik]

Both these measures enjoy the benefit of simplicity. Cash flows are easier to forecast in the near future than the distant future, so a payback rule can be implemented more accurately. And accounting rates of return are computed from data that is routinely compiled by management accountants, making comparison and monitoring relatively easy.

7.2.1.2 Accounting rates of return

The internal rate of return (IRR) is defined as the interest rate that gives a net present value (NPV) of zero. The NPV is calculated from an annualized cash flow by discounting all future amounts to the present. [Wik]

As an investment decision tool the calculated IRR is used to rate alternative investments. The investment alternative with the highest IRR is preferred. Note that placing the initial investment amount in the bank is always an alternative. Thus, any investments which do not match the bank's going deposit rate will not be realized [Wik]. It should also be noted that zeros of NPV as a function of IRR may lack existence or uniqueness if there is some alternation of positive and negative cash flow. The IRR exists and is unique if one or more years of net investment (negative cash flow) are followed by years of net revenues. [Wik]

7.2.1.3 Net Present Value

Net present value is a form of calculating discounted cash flow. It encompasses the process of calculating the discount of a series of amounts of cash at future dates, and summing them. It is a time consuming process, but not difficult at all.

Net Present Value can be calculated by the following formula, where t is the amount of time (usually in years) that cash has been invested in the project, N the total length of the project (in this case, five years), i the weighted average cost of capital and C the cash flow at that point in time.

$$NPV = \sum_{t=0}^{N} \frac{C_t}{(1+i)^t}$$

7.2.1.4 Real options

A real option is the right, but not the obligation, to undertake some business decision, typically the option to make a capital investment. For example the opportunity to invest in the expansion of a firm's factory is a real option. In contrast to financial options a real option is not tradeable - e.g. the factory owner cannot sell the right to extend his factory to another party, only he has the decision to make. The terminology "real option" is relatively new, whereas business operators have been making capital investment decisions for centuries. However the description of such opportunities as real options has occurred at the same time as thinking about such decisions in new, more analytically-based, ways. As such the terminology "real option" is closely tied to these new methods [Hull].

Real options Analysis is being used by different areas as a "mechanism" to see an investment into the future and predict available paths for it. The novelty with this tool is the fact that it gives the analyst the advantage firstly to see different pathways of the invstment and secondly gives the option to choose one of them according to the desired result. The immediate justification for the use of real options is that it provides a far better and much needed substitute for the current methods of project evaluation. Indeed, current practice uses discounted cash flow or, equivalently, net present value, to assign value to projects and, thus, to determine which should be funded and incorporated into design. Yet, these approaches are fundamentally flawed and inadequate whenever a project will exist in an uncertain environment. The difficulties, the fatal flaws of the discounted cash flow methods as practiced are both conceptual and mechanical [Neuf]. This method will be used for the economic modelling in this case study, and for that reason a further discussion will be held in the following sections.

7.3 Real Options Analysis, an introduction

7.3.1 Options and Options Pricing

An option is a security that gives its owner the right to trade a fixed number of shares of a specified common stock at a fixed price at any time on or before a given date. The act of making this transaction is referred to as *exercising the option*. The fixed price is termed the *strike price* and the given date, *the expiration date*. Consider the following example [Durb]:

> *Greta is a small investor who likes to trade stocks and options in her spare time. She believes the stock of the ZED coorporation, currently trading at $60, is undervalued and will increase over the next several months.*
>
> *Rather than buy the shares and hold them, Greta buys 6-month options on ZED with a strike price of $60. the options give her the right, but not the obligation, to buy ZED for $60 at any time over the next 6 months.*
>
> *In 6 months, ZED is trading for $62. Greta exercises her option and buys ZED for $60, realising a gross profit of $2 per share.*

Options have been traded for centuries, but they remained relatively obscure financial instruments until the introduction of a listed options exchange in 1973 [Cox]. Historically the pricing of options was entirely ad hoc. Traders with good intuition about how other traders would price options made money and those without it lost money. Then in 1973 Fischer Black and Myron Scholes published a paper proposing what became known as the Black-Scholes pricing model, and for which Scholes received the 1997 Nobel Prize (Black had died, and was therefore not eligible) [Wik]. The model gave a theoretical value for simple put and call options, given assumptions about the behavior of stock prices. The availability of a good estimate of an option's theoretical price contributed to the explosion of trading in options. Researchers have subsequently generalized Black-Scholes to the Black model, and have developed other methods of option valuation, including Monte Carlo methods and Binomial options models. These path-breaking articles have formed the basis for many subsequent academic studies. As these studies have shown, option pricing theory is relevant to almost every area of finance [Cox]. Since then, options trading has enjoyed an expansion unprecedented in American securities market and influence the techno-economic studies as the one to be carried out here.

7.3.2 Real options Analysis and the value of it

Real options analysis applies the financial principles of options to “real” or physical systems such as power plants, copper mines and so forth [de Neuf]. For the most part, the use of “real options” is almost identical with that of financial options. It is the right but not the obligation, to make a strategic investment decision (to expand, scale back, abandon, ealy, switch, to adjust the design of a system at a later date in a significant way that enables the system managers to either avoid downside consequences or exploit opportunities [Haber]) relative to an underlying asset. This right typically comes at a cost and exists for a specific and finite time period. Real options represent the flexibility in decision making that impact financial value typically available in some form, at some point in time, with many strategic business assets. This form of analysis gives weight to the opportunities, in addition to the traditional concern with losses and risks [Haber]. These options can exist individually, in sequence or in combination [Syn]. The problem is treated as a “black box” and the conceptual and analytic effort focuses on trying to work the available data into forms suitable for the tools of financial analysis; that is, variables with certain characteristics that are needed for the financial model each time.

Real Options Analysis is a systematic approach and integrated solution using financial theory, economic analysis, management science, decision sciences and econometric modelling in applying options theory in valuing real physical assets as opposed to financial assets in a dynamic and unceratin business environment where business decisions are flexible in the context of strategic capital investment decision making, valuing investment opportunities and project capital expenditures [Mun].

Consider the following example of the use of ROA to understand the role of it in the evaluation of alternative forms of technologies, design, etc. The food industry is characterised by relative maturity in terms of consumer markets. Major product innovations are driven by unique selling propositions, differentiation etc. Most technical innovation is focused on food processing, packaging and distribution. A manufacturer will be very careful in choosing new technologies and factory configurations and will have a strong need for flexibility during the phase of planning, evaluation and decision and even for the design and manufacturing phase of this equipment [Haber]. Real options analysis could help in the following way: if the manufacturer wanted to explore a different approach to food processing, that would be cheaper and faster, he/she had to be sure that by investing in new equipment, he/she would actually benefit. What ROA would do is weigh the option of keeping the same equipment over a period of time with the option of investing in new equipment over the same period. Aspects of the options such as money losses by keeping or changing the equipment, any risks of the change or not, or the correct timing for the change, could be seen depending on the analysis. Using this method the manufacturer could have an idea of if and when would be ideal for him to do the change if he/she saw that the new investment would be worthwhile. ROA tries to create future scenarios, given the right specifications, that will be weighed against each other and against the current scenario so as the person involved can take an informed decision of the way forward.

Real options are crucial in [Mun] :

- Identifying different corporate investment decision pathways or projects that management can navigate given the highly uncertain business conditions,
- Valuing each strategic decision pathway and what it represents in terms of financial viability and feasibility,
- Prioritizing this pathways or projects based on a series of qualitative and quantitative metrics,
- Optimising the value of the strategic investment decisions by evaluating different decision paths under certain conditions and using a different sequence of pathways to lead to the optimal strategy.

<u>Assumptions for Suitability</u>

Basic financial options models assume that the dynamic underlying asset value process is geometric, i.e. multiplicative. Therefore, the underlying asset (stock price or project value) distribution is lognormal [Haah], [Cop], [Cox]. The properties of lognormal distribution are that the value can not fall below zero, but it may increase to infinity without upper limit.

The distribution is positively skewed, having most of the values closer to the lower limit, and the natural logarithm of the distribution yields to normal distribution. The Black-Scholes Option Pricing Model calculates option value based on:

1. the projected value of the project (assumed to be distributed lognormal and non-negative),
2. the variability of projected value,
3. the cost of the project (assumed to be known with certainty),
4. the risk-free rate of return, and
5. the time until expiration.

Other Option pricing models have less strict assumptions. The multiplicative form of the Cox-Rubenstein Binomial model-which will be used in this thesis- allows non-constant variance, and the additive form of the model also allows project value to be normally distributed with the potential for negative values [Cop]. Schwartz and Zozaya-Gorostiza [Schwartz] have developed a model that explicitly models uncertainty in project costs, and allows an additional form of uncertainty, which is the possibility of catastrophic event during development.

Potential Use

Real options analysis should not be viewed simply as an academic method to calculate an investment. Individual and organisations should look at real options from three key perspectives [Syn]:

- As a strategic planning and analysis thinking framework
- As a risk assessment and management methodology
- And lastly as an investment analysis tool.

Much of the benefit of real options analysis is developing new or/and alternate approaches to actively manage assets utilising flexibility and in better recognising and taking advantage of existing investment flexibility. This "real options thinking" leads to increased opportunities to take advantage of upside market potential and mitigate downside risks.

Real options strategies put into action, simultaneously increase expected value as well as improve the risk profile of investment cash flows.

The three key functions of real options as described before, can be closely corelated with the step-by-step approach used for modelling in general; from a more abstract, foundation way of thinking, to finding an appropriate modelling methodology, to actually deriving results and making decisions according to those results. More specifically, the use of real options as a strategic and planing and analysis thinking framework, can be linked with the stage of the realisation of a problem and the first steps of analysisng a way forward, the conceptualisation of the problem. Its use as a risk assesment and management methodology, relates to the analysis based on models and diagnostics, while its use an an investment analysis tool, relates to the decision making stage of the modelling process.

The fundamental structure of Real options gives the modeller the advantage of creating possible scenarios of the way forward, and manipulates the best options. This ability to move from one pathway to another, providing a much needed basis of flexibility is something that has engaged many disciplines, one of whom is Systems engineering. The recent focus on "agility" in Systems Engineering is a manifestation of the increasing speed at which new products and systems are designed and introduced into the market place. More than speed, however, it is the existence of uncertainty in future needs and operating conditions and the resulting ambiguity in the requirements that drives new ways of developing systems. Agile Systems engineering, on the other hand, puts the emphasis on embedding agility in the systems themselves. This is usually done when the ability to predict the future demand or functional requirements of a system is severely compromised. An agile system is both flexible and has the ability to change from one state or operating condition to another rapidly, without large switching costs or increases in system complexity [Haber], something that can be correlated with the fundamental structure of Real options analysis.

The use of options in identifying possible decision pathways and prioritising them according to their investment value, its use as a strategic and planning tool, can be parallelised with a concept we have already discussed in a previous chapter; that of a *conceptual process flowsheet.* The design of any process in an early stage involves as a fundamental stage the problem of conceptual modelling, which transforms requirements and objectives to sets of preliminary designs referred to as Conceptual Process Flowsheets. The Conceptual Process Flowsheets enable the modeller to formulate different *scenarios*, and given sufficient detail, enable the modeller to *make a decision.* The introduction of each flowsheet is part of the conceptualization of the problem, as a way of describing the different possible strategies, but the actual formulation of a flowsheet is a modelling process of each one and will not be covered in this thesis. In this sense option analysis, may be seen as one of the decision making tools for early forms of models as process flowsheets reach other forms of models.

Options analysis provides a framework within which several alternatives are examined, on its conceptual level, and then the most attractive options are laid for comparison. The process by which different scenarios are compared and decisions are made according to the best alternative-best being the one that has been predefined from the requirements and generally, the criteria-is loosely similar to that of a conceptual process flowsheet. It must be noted here, again, the necessity of a formed model, a conceptualization of the problem, must exist as a basis where the analysis of Real Options will take place. The existence of such a model is considered given when using real options analysis and by no means do we consider Real options analysis a modelling methodology. The value of ROA comes as a result of the weight it has as an analysis and decision making tool for strategic decision making.

The similarity of Conceptual flowsheets with the structure of Real options analysis, is further supported by the nature of option pricing when it comes to pricing real life projects. The term "real life projects" here is emphasizing the fact that the underlying is not a stock but something more concrete, building a factory, or extracting oil from the earth-it is a real life situation, something "tangible". What is really interesting with the formulation of those situations is the fact that during the progression of the problem formulation, choices have to be made constantly, quite like in the progression of the flowsheets.

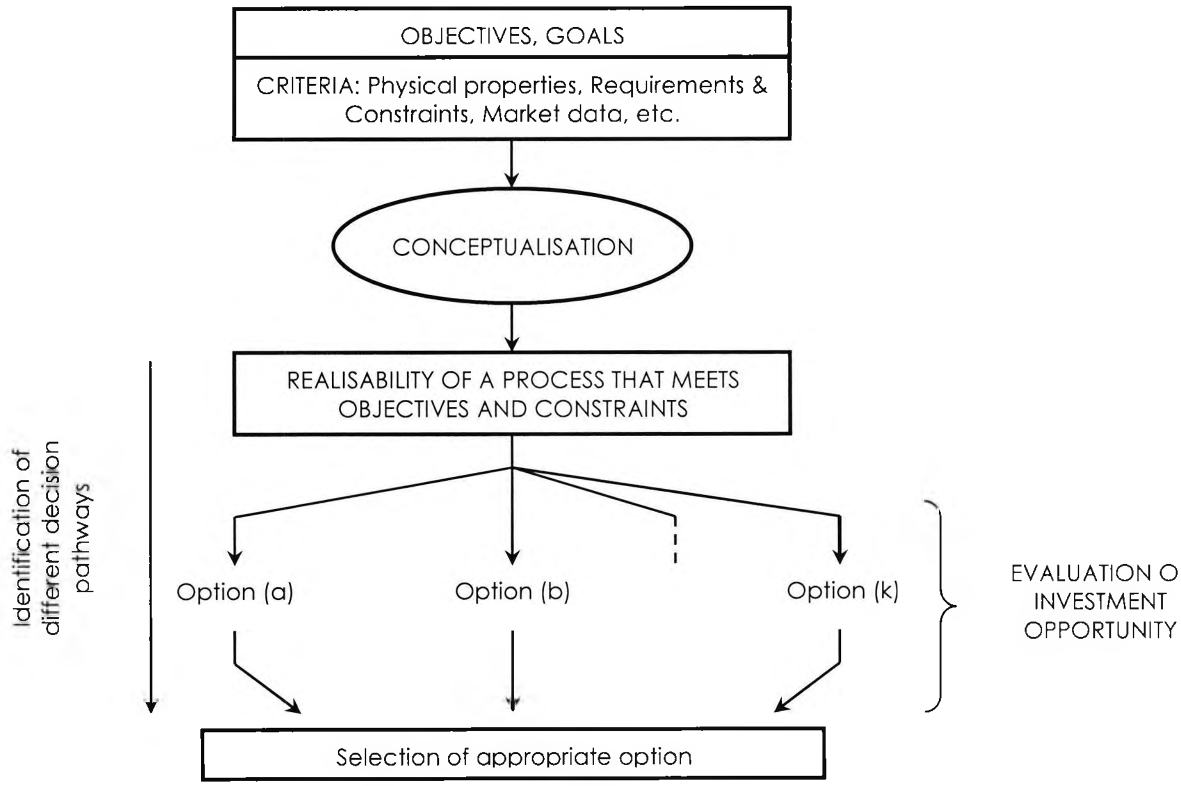


Figure 7.2: Real Options Analysis seen as a Conceptual Process flow sheet.

The figure above shows the progression of thinking when using Real options analysis and how it is linked with the evaluation of paths in Conceptual process flow sheets. At the stage of the evaluation of the different investment opportunities the modeller forms the different options and weighs them according the requirements he/she has set to the analysis, so as to get the optimum investment. Real options analysis provides a way of thinking that focuses on flexibility in the upstream process of conceiving, designing and implementing products and systems. This flexibility at the particular stage of the analysis gives Real Options the edge as a method of decision making.

7.3.3 Using Real Options in Strategic Decision Making

To maximize a firm's value its managers must match internal capabilities to external opportunities. Flexibility in timing of decisions about the firm's capabilities and opportunities give managers 'real options'.

Companies are not passive investors: managers have the flexibility to sell the asset, invest further, wait and see or abandon the project entirely. It is precisely the way in which real options deal with uncertainty and flexibility that generates their value. Real options are not just about “getting a number”; they also provide a useful framework for strategic decision making [Walt].

Real Options analysis is a tool for assessing investment decisions or developing strategic plans under uncertainty, and provides a quantitative method for monitoring, measuring, and adjusting decisions as economic conditions change. Traditional models such as NPV assume that managers have no flexibility to make changes; that the investment choice is an all or nothing decision. Real options analysis recognizes the value of active management in creating and utilising flexibility in the face of uncertainty. The traditional discounted cash flow methodologies (typically NPV-Net Present Value) used by the vast majority of coorporate practitioners are based on a static set of input values and assume passive management over the life of the investment. The problem with this approach, people argue, is that by extrapolating all the possibilities for the future into a single scenario, NPV doesn't account for the ability of executives to react to new circumstances; for instance, spend a little up front, see how things develop, then either cancel or go full speed ahead" [Smith]. The key to valuing these options, they argue, is to consider the uncertainty or "volatility" associated with the investment in the same way that Black, Scholes, and Merton did in their Nobel Prize winning work on valuing financial options. When valuing flexibility *the significant determinants of value are the degree of uncertainty in the anticipated returns of the underlying asset* (volatility), *the cost/benefit of implementing the option* (exercise or strike price) *and the amount of time available before the opportunity expires* [Hine].

Limitations

Real options focus on the evolution of a few complex factors over time that determine the value of investment and cash flows. These are factors about which decisions can be taken at any time over a period. Decision-tree analysis tends to consider great detail in the cash flow models and many uncertainties, but relatively little in the way of dynamic decision making. There are a large number of these factors with decisions made at discrete time periods

[Walt]. Critics of real options analysis point out that the value derived by conducting real options is not "real" unless the organisation using the analysis actually has the ongoing evaluation process and management decision ability to take the appropriate decisions at the appropriate time [Syn].

7.3.4 Real Options Analysis in application

Real option analysis is a relatively new methodology compared to traditional discounted cash flow methods. The oil and gas, mining and electric utility industries were the earliest adopters and remain significant users of real methodology. Later real option analysis expanded into biotechnology, pharmeceuticals and now into telecommunications, high-tech and across all industries.

The oil and gas industry is fraught with strategic options problems [McCard], [Trig]. This is because oil and gas exploration and production involves significant amounts of risk and uncertainty. For example, companies spend millions of dolars to refurbish their refineries and add new technology to create an *option to switch* their mix of outputs among heating oil, diesel and other petrochemicals as a final productusing real option as a means of making capital and investment decisions. This option allows the refinery to switch its final output to one that is more profitable based on prevailing market prices, to capture the demand and pricey cyclicality in the market[Mun].

In this thesis, the real options analysis will be used in the oil and gas industry, but from a point of view that was never really addressed before. This case study will give an insight of prices of electricity generated by gas up to the year 2050 and then an analysis of different options of alternative energy generators-those being coal,nuclear, etc- will take place. After the use of real options analysis, a prediction of the appropriate time to change to the suggested energy generator will be made.

A study of 39 companies conducted by Alex Triantis [Trian], found that most businesses using real option valuation use it in three key ways [Sam]:

- *As a mode of thinking.* Although real options are analytically vigorous, they are best understood as a way of thinking. From the modeller's perspective, it is a useful tool for stimulating thinking about a range of possible options and helping to make decisions on what to invest in, that means appreciating what types of options exist, how they can be created, how and why option values change, and how to capture their value [Maub]. In particular, ROA helps to keep investment options open, and enable riskier approaches to be explored, without making long-term commitments to them. It enables analysis of the range of options available both now and in the future, while identifying investment cut-off points
- *As an organisational process and an analytical tool.* The real options approach can help organisations improve their decision processes to take advantage of additional value in order to build or maintain competitive advantage, and to do so before the market punishes it or the competition extracts it [Syn]. Strategy guru Henry Mintzberg [Mint] makes the following point. He suggests that strategic planning, is really strategic *programming:* an articulation of strategies that already exist. He advocates strategic *thinking*- really *synthesis*- that incorporates intuition and creativity. Strategic planning is DCF (Discounted Cash flow)-based; strategic thinking is options-based. There are three steps in turning real options thinking into useable results. The first is to accurately identify a real option. The second is use of the options model itself. Finally, consideration must be given to the potential differences between option-model-derived value and real-world value. Defining the application is probably the most important part of a real options analysis. Amram and Kulatilaka, [Amram], break the defining task into four parts: the decision, the uncertainty, the decision rule, and the review:
    - **Defining the Application**
        - **The Decision**
            - What are the possible decisions?
            - When might they be made?
            - Who is making them?
        - **The Uncertainty**
            - What is the source?

    - What is the trend?
    - How has it evolved?
    - What other market factors are important?
- **The Decision Rule**
    - Create a mathematical expression
- **Look to the Financial Markets**
    - Is uncertainty private or market-priced?
    - Are there better alternate frames?
- **Review for Transparency and Simplicity**
    - Is the application definition clear?
    - Can managers understand the definition?
    - Is the definition clear to investors?

Appropriate definition of the application is a difficult balancing act. The modeler has to have some intuition about the problem with a model that maintains a degree of rigor. Pinpoint precision is neither an appropriate goal nor a likely outcome of the model: the thought process alone is valuable and leads to potentially important insights and opportunities [Maub].

Real options valuation shields strategic thinking and decision making in concrete financial analysis. The model also enables an organisation to recalculate the value of a project or an investment as it progresses and to understand what must happen before the project or the investment can move successfully into the next stage of development. But real options analysis is not appropriate for every investment because putting the model to work takes time, effort and expertise. "Real options is best suited for major investments with results far in the future and a lot of managerial choices in between," says John MacCormack, senior vice president and head of the energy practice at Stern Stewart & Co., a management consultancy in New York city [Sam]. Moreover, real options work only if a company is truly prepared to cancel a project if the numbers look bad after the initial investment. The flaw in the theory is not its complexity, as some have said, but the fact that it ignores the psychological and political realities of capital investments [Fink].

With NPV to capital budget decisions, an appropriate discount rate is applied to a project's anticipated cash flows. If the resulting present value of those cash flows exceeds the cost of capital and the NPV of alternative investments, the company proceeds with the project. With real options a decision tree is plotted showing a series of different scenarios that could develop at various points throughout the life of the project. Then probabilities and discount rates are attached to each scenario and the cash flow is discounted back to the present. If the outcome is positive, the company makes at least a partial investment.

Three conditions are prerequisite to using real options concepts to structure the evaluation and management of technology investments [Dix]:

- uncertainly regarding net payoffs,
- irreversibility in project costs, and
- managerial flexibility regarding how projects are structured.

While it seems nearly or entirely impossible to identify options within a given strategy or a project based on irreversible investments (since this is true for almost any investment), uncertainty and flexibility may signal the presence of options in a better way. Moreover, those factors may even reveal the importance or added value of the option, since their magnitude directly influences the option value and is quantifiable. Both factors are interdependent. Since flexibility constitutes the option to react to a state of resolved uncertainty, the later seems key to the presence of options. Hence, flexibility is the necessary condition to validate the use of options to mitigate an uncertainty. Thus the identification of real options inherent in a strategy could be driven by the identification of those uncertainties towards which a reaction is possible [Ben].

The trouble is, the flexibility created by a Real Option analysis must be supported by corporate discipline for the analysis to be vindicated. Traditional financial options, after all, are contracts with a specified expiration date, and their value tracks that of an underlying security. This explains why real options analysis can so readily be applied to projects in the energy business, where the use of futures contracts and exploration leases with definite terms is commonplace.

Real-options analysis could gain currency via its application in a number of functions that aren't industry-specific, such as supply-chain management. Build-to-order models, flexible assembly, contract manufacturing, and procurement contracts all offer numerous options that can be exploited. In high tech, as computer components become more commodity like, with futures, options, and spot markets developing for items like memory chips, supply-chain managers will need to become skilled financial engineers, predicts Alex Triantis, Real options could become one of their most valuable tools. [Teach]

The analogy between the pricing of a financial option on traded assets and the valuation of an R&D project as an option on future market introduction has strong managerial appeal. Any financial or real option can be seen as an initial investment offering the exclusive opportunity to keep open a specified follow-on investment trajectory at limited predetermined costs. R&D may therefore be considered as an initial investment opening up future market introduction of new products or technologies. The flexibility of that tool gives the perfect basis for projects related to energy.

Information technology is another fertile, cross-industry field for applying real options. IT now consumes the greater part of corporate capital budgets, and large applications are notoriously risky. But their deployment can be optimized, and the risk minimized, through real-options analysis, according to Mark Jeffery. In a recent paper, Jeffery and co-authors Sandeep Shah and Robert J. Sweeney demonstrated how real-option analysis can determine the optimal rollout of an enterprise data warehouse, via the phase-wise consolidation of data marts. [Teach]

7.3.5 Real Options compared with the DCF

Certain critical components of real options make them a powerful analytical tool. First, they recognize and value the flexibility that today's capital investments provide. Second, they recognize the nature of many investments and account explicitly for the reality that certain investments will never be made if -based on additional information developed over time-they are deemed unattractive. In these instances, it makes sense simply to abandon them, rather than sink additional money into a poor investment. By contrast, DCF

(Discounted Cash Flow) evaluates a series of investments as if they will be made, regardless of whether they still make sense at a later date.

Additionally, with real option analysis, uncertainty inherent in investment projects is usually accounted for by risk-adjusting probabilities. Cash flows can then be discounted at the risk-free rate. With regular DCF analysis, on the other hand, this uncertainty is accounted for by adjusting the discount rate (using e.g. the cost of capital) or the cash flows (using certainty equivalents). These methods normally do not properly account for changes in risk over a project's lifecycle and fail to appropriately adapt the risk adjustment. More importantly, the real options approach forces decision makers to be more explicit about the assumptions underlying their projections.

Another critical difference between DCF and real options is the effect of uncertainty (or risk) on value. Uncertainty is typically considered bad for the valuation of traditional cash flows. By contrast, uncertainty increases the value of real options.
Consider the following key points:

- As volatility (uncertainty) increases, so does the value of the real option.
- Initiatives with great uncertainty should be implemented in stages. Making a small investment up front can give management the ability to resolve uncertainty through data gathering and learning. The larger investment can be made in a future environment with less uncertainty.
- A series of initiatives should be looked at on a portfolio basis. The overall results of the investment portfolio are what ultimately matters, not the individual performance of each initiative.
- Real options recognize that abandonment is a viable alternative that must be contemplated from the outset. Furthermore, dropping a project does not necessarily mean that the team in charge of the particular initiative has failed.

One of the most important limitations of DCF is that it fails to account for the value of managerial flexibility that is inherent in many types of projects. Technology investments might often grant the possibility of pursuing an avenue in several months or a couple of years. But without the relatively small initial investment, an opportunity might be foreclosed forever. Although real options can be intuitively appealing, execution to arrive at a value is difficult. Determining the exact value of a real option is not necessarily critical. Instead, understanding the drivers of the valuation and the value relative to traditional methods is much more important.

7.4 Real options in detail

There are two types of options contracts. These are:

- CALL options
- PUT options.

CALL OPTIONS

CALL options give the holder (or owner) the *right to buy an asset* at a particular time and at a certain price (known as the Strike Price). If we are talking about a European option (European Option: An option that the user can exercise only at expiration and no other time before it), then the holder only has the right to exercise the option at maturity date and not at anytime until then this date is reached. If we are talking about an American Option (American Option: An option that the user can exercise at any time before and until the expiration date. This can be well before maturity, half-way through maturity or even right on the expiration date), then the holder has the right to exercise the option at any time up until and including the expiration date (i.e. at maturity).

PUT OPTIONS

PUT options give the holder (or owner) *the right to sell an asset* at a particular time (usually up until or at the expiration date) and at certain price (the Strike Price). If we are talking about a European PUT option, then the holder only has the right to exercise the option at expiration date and not at anytime until then. If we are talking about an American PUT option, then the holder has the right to exercise the option at any time up to and including the expiration date (i.e. at maturity). When buying the right to buy or sell an asset, then this means that the asset will be referred to as the *underlying* or even the *underlying asset*. Irrespective of which model or strategy is chosen for pricing any type of option there are six main factors that affect the actual price of an option. These are the following:

- The Current Underlying Price, So
- The Strike or Exercise Price, K
- The time to Expiration (i.e Maturity), T
- The volatility of the Price
- The Risk-Free Interest Rate, r
- The dividends (if any) expected during the life of an Option

7.4.1 Current Price and Strike Price (S_0, K)

The price of the underlying cannot be considered as a factor that can independently "exist" when dealing with options of any type (PUT or CALL, European or American). When pricing options, the underlying price and the strike price are interdependent. It is pointless to discuss about an underlying asset with a current price without knowing at which point exactly you have the right to exercise the right to buy or sell that asset. In simpler words, one must know the limit of the price of the stock. Depending on the position the investor takes on the option, the price of the stock must be less than or greater than the strike price in order for the option to be exercised.

If a CALL option is exercised at some future time then the payoff will be the amount by which the stock price S exceeds the strike price K (the difference of the two is the net profit S-K). Thus, a CALL option becomes more valuable as the stock price, S, increases. However, it becomes less valuable as the strike price, K, increases.

If a PUT option is exercised at some future time then the payoff is the amount by which the strike price, K, exceeds the stock price, S (i.e. K-S). PUT options become less valuable as the stock price S increases and more valuable as the strike price increases (i.e. the opposite from the CALL options).

7.4.2 Time to expiration-maturity (T)

As the time to expiration increases, this makes both American PUT and CALL options more valuable. The reason behind this is that the longer the life of the American Option, the investor has more opportunities to exercise his or her rights to buy or sell in comparison to the owner of an option that has a shorter time to mature. However, European PUT and CALL options do not necessarily become more valuable as the time to expiration increases. At this stage, it must be noted that when a stock pays dividends to the shareholders, then the price of this stock declines in the long-term. If however, this European option has a short time to maturity, then even if it still pays dividends, it will be more valuable than the other option that pays dividends too, but expires later. From this the rule of thumb that when a stock pays dividend, it causes the stock price to decline is verified.

7.4.3 Volatility (σ)

Volatility *is the measure of how uncertain the investor is about the future stock price movements* [Hull]. The stock's performance is highly dependent on volatility. After all, no investor or trader can be sure about the future stock price movements. There are so many (infinite) factors that need to be taken into account for one to predict the markets and that makes uncertainty (or volatility) a major ingredient for failure or success. In general, it can be seen that as volatility (or uncertainty) increases, the values of both PUT and CALL options also increase. In real options analysis, the standard deviation of the expected price change over a year is used to measure the uncertainty and volatility of the underlying asset. Volatility may be also regarded as the second moment of the value distribution. It is probably the most difficult input parameter to estimate in real options analysis [Mun], which is also the case with financial options. However, there exists historical data for financial markets that can be used for choosing and comparing different alternative stochastic models and their parameterizations to find the appropriate volatility measurement.

7.4.4 Risk-free interest rate (rf)

The risk-free interest rate affects the price of an option too but not in such a straightforward way. If the interest rates increase in the market by the Bank of England for instance, then the investors would expect to receive a better price for their underlying assets, which means that an increase in the interest rates would make the asset price increase. The current value of any future cash flow received by the investor who owns the option decreases. This actually increases the value of call options and decreases the value of put options. We are assuming that interest rates change while all other factors stay the same. More specifically, we are assuming that interest rates change while the stock price remains the same. In practice, when interest rates rise (or fall), stock prices tend to fall (or rise). The net effect of an interest rate increase and the accompanying stock price decrease can be to decrease the value of a CALL option and increase the value of a PUT option. Similarly, the net effect of an interest rate decrease and the accompanying stock price increase can be to increase the value of a CALL option and decrease the value of a PUT option [Hull].

7.4.5 Dividends

A dividend is a share of profits paid to people who own parts of an underlying asset, such as stocks. To get a better feeling of how dividends affect the value of the underlying and consequently the value of the option, it must be noted that a stock's return is its dividends plus capital gain, but this is not always the case. When an investor buys stocks from a company, then he or she becomes part owner of that company (shareholder). Many investors believe that the only way to receive income from their stocks is through dividends. This is not always the case though, because a company may indeed be in a position to pay dividends to shareholders, but may prefer to re-invest their net-profits in order to increase the stock price and thus in the long-term give its shareholders more value for money. Dividends represent the value that flows out of the stock price between the purchase of the option and its exercise (this is the modification to Black-Scholes that was added by Merton).

7.5 Valuation of Options

There are quite a few methods used today to value options. Methods such as the Black-Scholes model, the binomial lattices, simulation methods (such as Monte-Carlo simulation) and other numerical techniques. Models like the Black-Scholes and the binomial lattices are most widely used. They are easy, exact and quick to use.

7.5.1 The Black-Scholes Model

In the early 1970s, Fischer Black, Myron Scholes and Robert Merton made a major breakthrough in the pricing of stock options. This turned out to be the development of what is now commonly known as the Black and Scholes model. This model has been a tremendous contribution to the success of financial engineering in the 1980s and in the 1990s. However, it was not until 1997 that the importance of this model was recognized with the award of the Nobel Prize to Robert Merton and Myron Scholes. Unfortunately, Fischer Black had died earlier in 1995. For the pricing of an option using the Black and Scholes model equation five major factors need to be taken. These are the following:

- Current Asset Price, S(T)
- Strike Price of the Option, K
- Volatility, σ
- Dividend payment
- Maturity/Expiration Date (T)

The Current Asset Price (i.e. at time zero) S_0 is the major factor used in all the models for the pricing of options. Without knowing the current underlying asset price, it is impossible to value any option whatsoever. The formulas used for the pricing of options in the Black-Scholes model are different according to the type of option used. It must be noted that the following formulas given are used for European options on a non-dividend paying stock. In general the Black-Scholes model was designed with non-paying European stocks in mind. The price of non-dividend paying European call and put options is given by:

$$Call = S\Phi\left(\frac{\ln(\frac{S}{K}) + (rf + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) - Ke^{-rf(T)}\Phi\left(\frac{\ln(\frac{S}{K}) + (rf - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)$$

$$Put = Ke^{-rf(T)}\Phi\left(-\left[\frac{\ln(\frac{S}{K}) + (rf - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right]\right) - S\Phi\left(-\left[\frac{\ln(\frac{S}{K}) + (rf + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right]\right)$$

where S is the current underlying asset price, K is the strike price, T is the time to expiration, rf is the risk free rate, σ is the volatility and Φ is the cumulative standard-normal distribution.

7.5.1.1 Volatility Calculation from Historical Data

Probably one of the most difficult input parameters to estimate in real options analysis is the volatility of cash flows. The Logarithmic Cash flow Returns approach calculates the volatility using the individual future cash flow estimates and their corresponding logarithmic returns [Mun]. Below an example of this approach will be described.

Starting with a series of forecast future cash flows convert them into relative returns. Then take the natural logarithmic returns of these relative returns. The standard deviation of these logarithmic returns is the volatility of the cash flow series used in real options analysis.

Time period	Cash Flows	Cash flow Relative Returns	Natural Logarithm of Cash Flow Returns (X)
0	$125	-	-
1	$125	$125/$100=1.25	ln($125/$100)= 0.2231
2	$95	$95/$125=0.76	ln($95/$125)= -0.2744
3	$105	$105/$95=1.11	ln($105/$95)= 0.1001
4	$155	$155/$105=1.48	ln($155/$105)= 0.3895
5	$146	$146/$155=0.94	ln($146/$155)= -0.0598

Table 7.1: Future cash flow estimates and their corresponding logarithmic returns [Mun]

The volatility estimate is then calculated as:

$$volatility = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}$$

where n is the number of Xs, and $\bar{x}$ is the average X value. For the specific example the volatility is 25.58%

This method is very easy to implement, mathematically valid and is widely used in estimating volatility of underlying assets [Mun]. Models like the Black-Scholes where there exist equations that can be solved given a set of input assumptions are exact, quick and easy to implement but are difficult to explain because they tend to apply highly technical stochastic calculus mathematics. They are also very specific in nature with limited modeling flexibility [Mun].

7.5.2 The Binomial Lattice

Binomial lattices are easy to implement, easy to explain and highly flexible. The binomial options model provides a generalised numerical method for the valuation of options. It was first proposed by Cox, Ross and Rubinstein (1979). Essentially, the model uses a "discrete-time" model of the varying price over time of the underlying financial instrument. The option is then valued via application of the risk neutrality assumption over the life of the option, as the price of the underlying instrument evolves. Because it models the underlying over time, as opposed to at a particular point, this approach is able to handle a variety of conditions for which other models cannot easily be applied. (For example, the model is used to value American options which can be exercised at any point). The model is also relatively simple, mathematically, and can therefore be readily implemented in a software (or even spreadsheet) environment; its use is therefore widespread in finance [Wik].

7.5.2.1 Market Replicating Portfolios and Risk-neutral Probability

In the binomial world, several basic similarities are worth mentioning. No matter the types of real options problems one is trying to solve if the binomial approach is used, the solution can be obtained in one of the two ways. The first is the use of risk-neutral probabilities and the second is the use of market replicating portfolios. The use of market replicating portfolios is more difficult to understand and apply but the results obtained from replicating portfolios are identical to those obtained through risk-neutral probabilities. It does not matter which method is used although application and expositional ease should be emphasized [Mun].

Market replicating portfolio's predominant assumptions are that there are no arbitrage opportunities and that there exist a number of traded assets in the market that can be obtained to replicate the existing asset's payout profile. Suppose you own a portfolio of publicly traded stocks that pay a set percentage *dividend* per period. You can, in theory, assuming no trading restrictions, taxes, or transaction costs, purchase a second portfolio of several *non-dividend-paying* stocks and replicate the payout of the first portfolio of *dividend-paying* stocks. You can, for instance sell a particular number of shares per period to replicate the first portfolio's dividend payout amount at each time period. Hence if both payouts are identical although their compositions are different, the value of both portfolios should then be identical. However, in real options world where physical assets are being valued, financial purists would argue that this assumption is hard to accept, not to mention the mathematics behind replicating portfolios are also more difficult to apply [Mun].

Compare that to using an approach called risk-neutral probability. Instead of using a risky set of cash flows and discounting them at a risk-adjusted discount rate, one can risk-adjust the probabilities of specific cash flow occurring at specific times. Thus, using these risk-adjusted probabilities on the cash flows allows the analyst to discount these cash flows (whose risks have now been accounted for) at the risk-free rate. This is the essence of binomial lattices as applied in valuing options. In any options model, there is a minimum requirement of at least two binomial lattices; the first lattice is always the lattice of the underlying asset, while the second lattice is the option valuation lattice [Mun].

7.5.2.2 The Methodology of Binomial Lattices

The binomial pricing model uses a "discrete-time framework" to trace the evolution of the option's key underlying variable via a binomial lattice (tree), for a given number of time steps between valuation date and option expiration. Each node in the lattice, represents a *possible* price of the underlying, at a *particular* point in time. This price evolution forms the basis for the option valuation. The valuation process is iterative, starting at each final node, and then working backwards through the tree to the first node (valuation date), where the calculated result is the value of the option.

Option valuation using this method is, as described, a three step process [Wik]:

- Underlying asset price tree generation
- calculation of option value at each final node
- progressive calculation of option value at each earlier node

The underlying asset price tree

The tree of prices is produced by working forward from valuation date to expiration. At each step, it is assumed that the underlying instrument will move up or down by a specific factor - **u** or **d** - per step of the tree. (The *Binomial* model allows for only two states.) If S is the current price, then in the next period the price will either be *S up* or *S down*, where *S up* =S x u and *S down* =S x d. The up and down factors are calculated using the underlying volatility, **σ**, and years per time step, **t** (it is the original Cox, Ross, & Rubinstein (CRR) method):

$$u = e^{\sigma\sqrt{t}}$$

$$d = e^{-\sigma\sqrt{t}} = \frac{1}{u}$$

The figure below is a representation of a generalised binomial lattice.

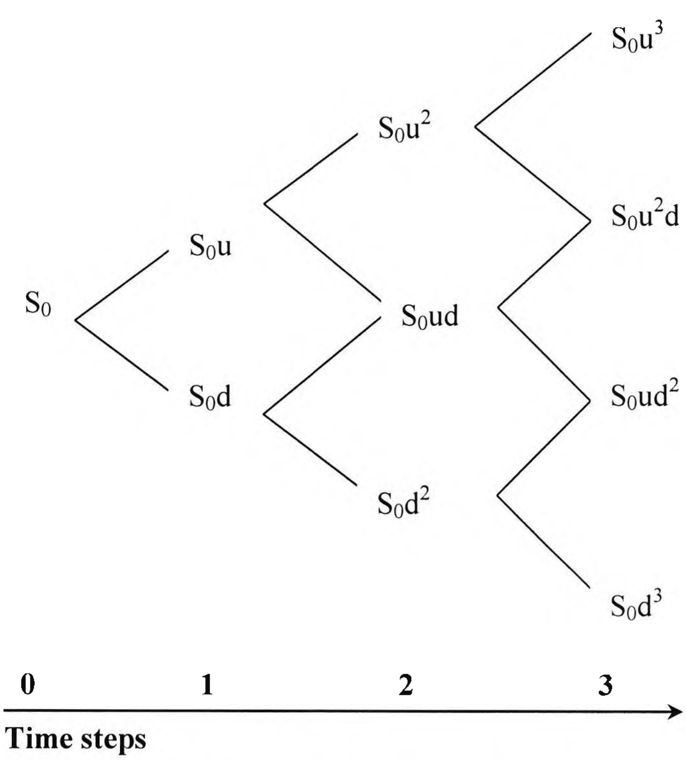


Figure 7.3: Binomial lattice of the underlying asset value (Recombining lattice).

(**Note:** Recombining and non-recombining binomial lattices yield the same results at the limit, so it is easier to use recombining lattices.)

Option value at each final node

At each final node of the tree-i.e. at expiration of the option-the option value is simply its intrinsic, or exercise, value.

For a call: value = S – Exercise price

For a put: value = Exercise price – S

Option value at earlier nodes

At each earlier node, the value of the option is calculated using the risk neutrality assumption. Expected value here is calculated using the option values from the later two nodes (*Option up* and *Option down*) weighted by their respective probabilities-"probability" **p** of an up move in the underlying, and "probability" **(1-p)** of a down move. The expected value is then discounted at **r**, the risk free rate corresponding to the life of the option. This result, the "Binomial Value", is thus the fair price of the derivative at a particular point in time (i.e. at each node), given the evolution in the price of the underlying to that point.

The Binomial Value is found for each node, starting at the penultimate time step, and working back to the first node of the tree, the valuation date, where the calculated result is the value of the option. The Binomial Value is calculated as follows:

Binomial Value = [p × Option up + (1-p)× Option down] × exp (- r × t)

Where:

$$p = \frac{e^{(rf-q)t} - d}{u - d}$$

where:

$$u = e^{\sigma\sqrt{t}}$$

$$d = e^{-\sigma\sqrt{t}} = \frac{1}{u}$$

q is the dividend yield of the underlying corresponding to the life of the option.

Next a generic example of the valuation lattice for the underlying asset (with black) and the valuation of the option in each node (with red) is given. The valuation lattice of the underlying asset (black) is sometimes refered to as forward model, while the valuation lattice of the option (red) is refered to as backward model, due to the direction of the calculations.

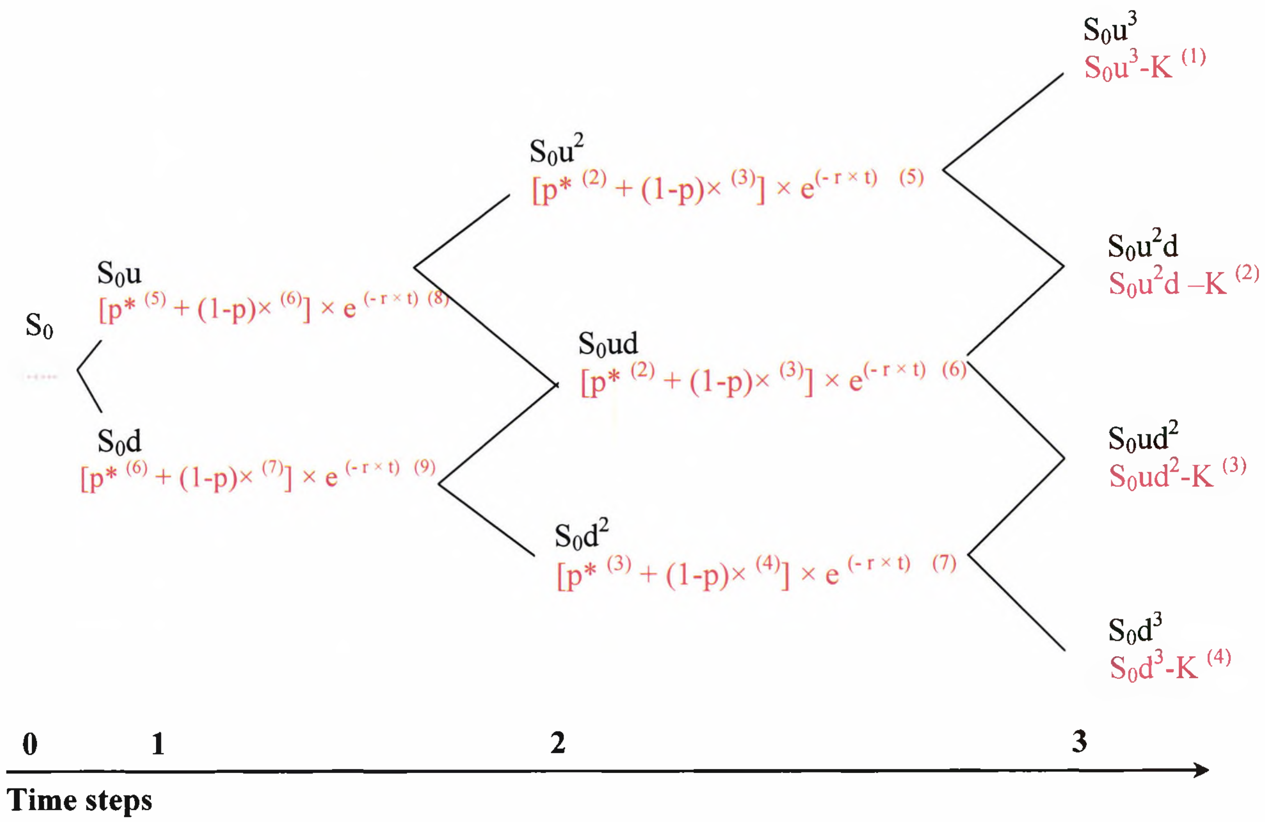


Figure 7.4: Option valuation lattice for a call option (Three time-steps Recombining lattice)

7.5.2.3 Relationship with Black-Scholes model

The use of binomial trees in the numerical valuation of options was first proposed by Cox, Ross and Rubinstein (1979). Although not as instantaneously recognised as the Black-Scholes (1973) options pricing model, binomial lattices are more easily generalisable and they are often able to handle a variety of conditions for which the former model cannot be applied. Binomial lattices give a record of the strategy to follow in any particular state as well as the value of the project with the option.

Both methods tend to provide the same results in the limit, but for ease of exposition the binomial lattice should be presented for management decisions. There are also other issues to contend with in terms of advantages and disadvantages of each technique. For instance closed form solutions such as the Black-Scholes model are mathematically elegant but very difficult to derive and are highly specific in nature. Binomial lattices, however are easy to build and require no more than simple algebra. They are also very flexible in that they can be handled easily to accommodate most types of real options problems.

Similar assumptions underpin both the binomial model and the Black-Scholes model, and the binomial model thus provides a discrete time approximation to the continuous process underlying the Black-Scholes model. In fact, for European options, the binomial model value converges on the Black-Scholes formula value as the number of time steps increases.

However, the attraction of the binomial lattice lies not in its ability to replicate the Black-Scholes model, but in its ability to handle more complex options. Amongst these are:

- Real assets options. These are options that are implicit in many capital investment decisions but which are generally ignored in financial evaluations using the net present value decision rule. The binomial lattice approach enables one to consider the value of options to abandon, expand and to defer start-up. Numerical techniques using the binomial lattice are being increasingly applied in finite reserve analysis such as the oil industry [Pad].
- A range of complex options that includes path dependent barrier options, look back options, options on options and American exchange options [Rub].

7.6 Final Thoughts

In this chapter we examined the setup for decision making, using primarily economic and financial tools. Decision making in conceptual models is quite challenging and requires special attention. Conceptualisation creates questions in the modelling process and in many cases dilemmas that need resolving before any decision is taken. In fact decision making is necessary to weigh possible scenarios that arise with the evolution of the modeling process. In our case (the energy problem) the decision making is primarily based on economic and financial criteria and thus a review of such tools has been undertaken in this chapter.

Due to the nature of the case study that will be discussed in detail in the following chapters, the focus was given to economic forecasting tools as a means to provide insight and elucidate certain grey areas where decisions are prominent. An overiew of some of the main, more efficient and widely used economic tools, as used by analysts for a plethora of diverse areas and businesses was examined. Real options analysis is the chosen method that will be used for the case study, and that choice is justified, giving a point by point description of the approach, how and when it is used, what are its strenghts and how it is compared to other leading tools. It has been emphasised that what we have been concerned with is the reviewing of a decision making tool rather than a new modelling methodology; that is a rough version of a model, with the basic variables needed to use Real Options analysis, is a prerequisite. The basic components of the methodology were described and finally the tool to be used was selected, that being the binomial lattice approach. It is the most convenient way to show graphically the evolution of the prices of the underlying asset, as well as, the valuation of the options prices, since the problem that will be discussed, and become the case study of this thesis, is basically a problem that needs a managerial solution, based purely on the economics of it. This case study will be the focus of further discussions in the following chapters; from the identification of the problem area, to the conceptualisation of the problem and the use of Real options analysis to that problem.