DECISION BEHAVIOUR, ANALYSIS AND SUPPORT

Decision making and how computers and analysis may support this

Simon French and Nadia Papamichail

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

simon.french@mbs.ac.uk
nadia.papamichail@mbs.ac.uk

Preamble

These notes are provide the basic reading for module CS635: *Decision Analysis and Decision Support Systems*. They are also supported by an extensive web-site which develops the material through more detailed notes and papers, case studies, links to other websites and software. The notes themselves are on the web-site and provide one of the entry points via hyperlinks (blue underlined text). The extensive reference list at the end covers both citations in these notes and in material on the web-site.

Please note that the web-site will be made available at the outset of the course.

For the present, these notes provide part of the prior reading for the course. The references at the end of the notes give an extensive bibliography of the area and provide the opportunity to read further on particular topics of interest, but it is not expected nor recommended that participants do before course week. Please also note that much of these notes focus on decision contexts, decision making and decision analysis. During the course more emphasis will be given to decision support systems, b, but these can only be satisfactorily understood and used if process of supporting decisions itself is first understood.

In addition as part of the prior reading, each participant should study the paper, specifically assigned to him or her. During the course week (w/b Monday, 17th February, 2003), you will be asked to make a short presentation (~10 min.) on an aspect of this paper. *Note*: in most case participants will *not* be asked to make a presentation the whole paper, just one part. The presentation will not be assessed. However, you should also write a 1500 – 2000 word summary of the key ideas in the *whole* paper. This may involve following up some of the references in the paper's bibliography as well as reflecting on related topics in these notes. This summary will form part of the course assessment (17% of the marks). The summary should be handed in by Tuesday February 25th, 2003. Thus will be a short time after the completion of course week for 'polishing' the summary in the light of the discussion during the presentations, seminars and laboratory work.

Table of Contents

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

1 Introduction

1.1 Decisions, decisions, decisions!

No two situations which call for a decision are ever identical. They differ because a decision changes the world in some small way and one can never go back to the original *status quo*. But there are many other ways in which decisions differ: see Figure 1.

- Problem context: e.g. what are the external characteristics of the problem? is it well structured? is uncertainty present? how many options need be considered?
- Cognitive factors of the decision maker (DM) or decision makers (DMs): how intelligent, imaginative, knowledgeable is she[1]? can the DM live with risk and uncertainty?
- Social Context: what are the characteristics of the social organisation in which the decision has to be made? who are the DMs? what are their responsibilities and accountabilities? who are the stakeholders?

We begin by considering a broad categorisation which will give shape to much of our discussion.

1.2 The Strategy Pyramid

Perhaps the most commonly discussed distinction between decisions is that between strategic, tactical and operational decisions: the so-called strategy pyramid: see Figure 2. Simon (1960) noted that strategic decisions correspond to ill-formed problems which are termed *unstructured* or *non-programmed*. DMs, e.g. a board of directors, seldom come to a strategic issue with a straight choice between, say, various acquisitions. Rather they first become aware that the company may need to grow. Through discussion, they formulate their objectives and the possible actions they might take. Only then do they have a strategic decision to make. In contrast, operational decisions are usually highly structured or programmed: e.g. should an inventory level be increased to support a production plan or in what order should the production of various items be scheduled? Another concept, which correlates well with the unstructured-structured dimension, is that of *time-span of discretion* (Jacques, 1989). Roughly speaking, this relates to the length of time before the consequences of a decision have their impact. The longer the time-span of discretion, the more unstructured the decision is likely to be.

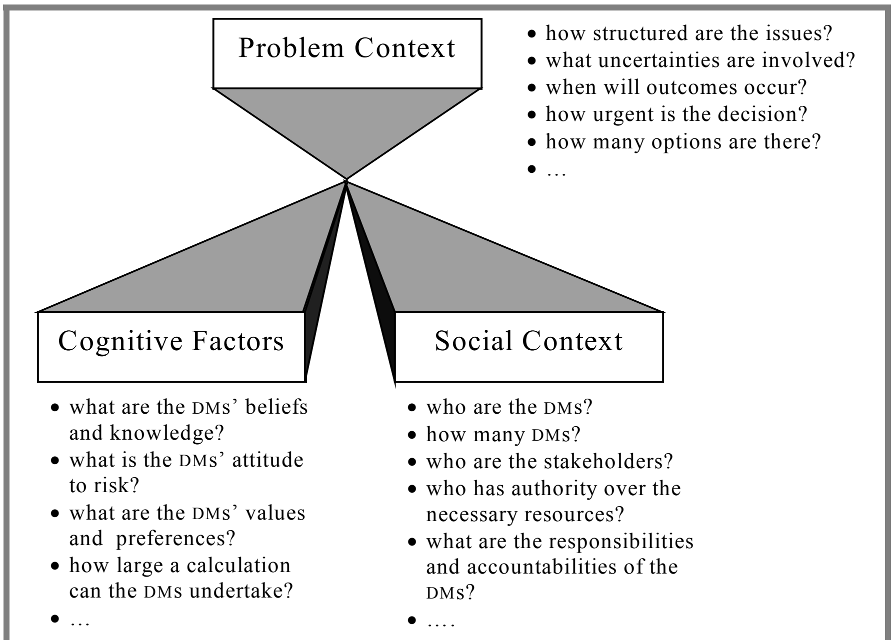


Figure 1: Factors that affect decision making (after Payne *et al*, 1993)

[1] We shall refer to the DM in the feminine and, later, a decision analyst in the masculine.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

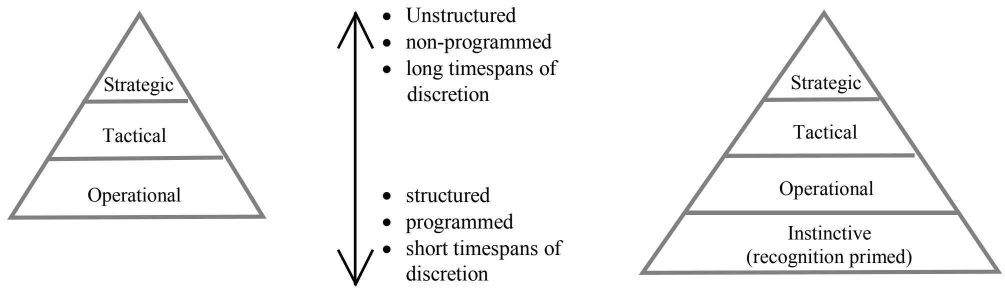


Figure 2: The strategy pyramid

The original 'three-level' strategy pyramid on the left of Figure 2 misses an important type of decision. In many cases, DMs seem to match the circumstances to something similar that has happened in the past and do roughly what they did then – or perhaps what they thought after the event they should have done. In such *recognition-primed* decision making, there is little or no comparison of options, just an instinctive choice of action. Thus we extend the strategy pyramid to include a fourth level. The term 'programmed' fits well with the idea of instinctive decision making based upon the (trained) recognition of similar circumstances. Within the discipline of artificial intelligence (AI) much effort has been expended on developing knowledge-based decision support tools (KB-DSS) which seek to 'automate' decision making. These tools operate at the lower levels of the strategy pyramid precisely because they need training. A research objective of AI is to develop KB-DSSs which need less training and operate at the highest levels of the strategy pyramid. However, for the present machines able to think strategically and creatively in unstructured, novel situations belong to the realm of science fiction.

Jacques (1989) argues that the tasks and decision making undertaken at different levels within an organisation may be characterised by the longest time-span of discretion required by their roles. Directors work on, say, a 2 to 5 year time frame. Although they may also deal with day-to-day operational issues, the defining imperatives of their role is that their 'vision' is focused on longer term strategy, which will bear fruit within 2 to 5 years. Line management focus their attention on shorter time frames, say 6 months to 2 years; while the production workers have much shorter time-spans of discretion, maybe 0 to 6 months. [*Note*: the time frames given here are examples only; they will and do vary between organisations.]

Jacques' theory is a mixture of the descriptive and normative, i.e. observations of how organisations *are* structured and reflections on how they *should* be. In many empirical studies he has shown that the concept of time-span of discretion provides a useful explanatory tool. However, he goes further and argues persuasively that organisations are best able to achieve their objectives when members of the organisation work at levels with time-spans of discretion within the limits of their ability to envisage the future. He terms the organisation *requisite* (Jacques, 1989).

In his observational studies, Jacques distinguished four domains of activity:

- the *corporate strategic* domain, which sets the guiding values and vision and develops strategy to take the organisation towards these;
- the *general* domain, which develops an implementation plan for strategy;
- the *operational* domain, which organises the detailed delivery of the strategy;
- the *hands-on work* domain, which delivers the work.

Note how these domains map on to the four levels (strategic, tactical, operational and instinctive) of the extended strategy pyramid (Figure 1). There is consistency here; any categorisation of decision making based on the how well structured a decision is correlates well with its time-span of discretion.

1.3 Players in a Decision

The simplest decisions involve just one person: the DM. She provides all the expert knowledge necessary, expresses her own judgements, performs her own analyses and makes her own decisions. However, in practice, this seldom happens. Rather a group of DMs, e.g. a management board or a government department, aware that they have a decision to make, will ask an analyst or consulting company to work with accountants, scientists, engineers and other subject experts to gain relevant information, then to formulate the problem, gather and analyse data and advise on the decision. Thus many will contribute to the process that leads to a decision; indeed, many may be party to the decision making.

The *decision makers* (DMs) are responsible for making the decision: they 'own the problem'. To be able to take and implement a decision, DMs need hold the appropriate responsibility, authority and accountability: *viz.*

- *Responsibility*. An individual or group is responsible for a decision if it is their task to see that the choice is made and implemented.
- *Authority*. An individual or group has the authority to take a decision if they have power over the resources needed to analyse and implement the choice.
- *Accountability*. An individual or group is accountable for a decision if they are the ones to take the credit or blame for the decision process and the choice that is made and implemented.

At various points in the decision process, responsibility may pass between different groups of DMs. When this happens, it is very important that the appropriate authority and accountability are also passed across. When responsibility, authority and accountability do not pass between groups in a coherent fashion, there is an obvious danger that the decision making process becomes dysfunctional.

The DMs are accountable to some, but not necessarily all the *stakeholders* in the problem. Stakeholders share – or perceive that they share –in the impacts arising from a decision. They have a claim, therefore, that their perceptions and values should be taken into account – and in many cases they are. The DMs are stakeholders, if only by virtue of their accountabilities; but stakeholders are not necessarily DMs. The obvious stakeholders in a business are its shareholders or partners, but there are many others: employees, customers, suppliers, local communities, etc. In public sector decision making, the government and its agencies generally have many stakeholders: the public, industry, consumers, political parties, etc. Accountability in the public sector is usually much broader than in the private sector, where the focus may be much more on

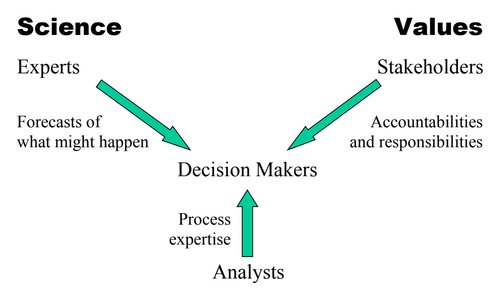


Figure 3: The players in a decision.

stakeholders closely associated with the 'bottom line'.

Experts provide economic, marketing, scientific and other professional advice used to assess the likelihood of the many eventualities. We shall often adopt the classical use of the term 'science' and use it to refer to a broad range of human knowledge. The knowledge that experts impart is used in the modelling and forecasting of outcomes of potential decisions. The DMs may have advisors who undoubtedly are experts in this sense, but they are unlikely to be the only experts involved. Other experts may advise some of the stakeholders, informing their perceptions and hence influencing the decision making.

Analysts develop and conduct the analyses, both quantitative and qualitative, which draw together empirical evidence and expert advice to assess the likelihood of the outcomes. They will also be concerned with a synthesis of the stakeholders' value judgements. These analyses are used to inform the DMs and guide them towards a balanced decision, reflecting the various expert and stakeholder inputs and the emphasis that the DMs wish to give these. Whereas experts support decision making by providing information – relevant economic data, assessment of physical risks or whatever – on the *content* of the decision; analysts provide *process* skills, helping to structure the analysis and interpret the conclusions. For this reason, analysts are sometimes referred to as process experts.

Figure 3 offers a simplified representation of the interrelationship between experts, stakeholders, DMs and analysts. This separation of roles is, of course, very idealised. Some parties to a decision may take on several roles. We have noted that DMs are necessarily stakeholders because of their accountabilities; but they may also be content experts and may conduct their own analyses. Similarly, experts may be stakeholders and *vice versa*. Analysts

	Data	→	Information	→	Knowledge
Description	observations of states and events in the world		data endowed with relevance to a context		general learning and understanding drawing on experience through reflection and synthesis
Characteristics	• easily captured • easily structured • easily represented • often quantifiable • raw resource		• needs agreement on meaning • built with analysis • reduces uncertainty within a context		• transferable between contexts • some knowledge explicit: e.g. Science • some tacit and personal: e.g. skills • hard to store and communicate
Human element	observation		judgement		experience

Table 1: From data to knowledge

may also be content experts and stakeholders, although there is a danger of bias entering the process if the analysts are too 'involved' in the decision itself. For this reason, it is common to arrange that at least some of the team of analysts are dissociated from the issues (see, e.g., Eden and Radford, 1990).

1.4 Inference, Prediction, Decision and Choice

Inference, also known as *induction*, is the process of learning from data: thus we have statistical or scientific inference. *Prediction*, or *forecasting*, is the process of building upon this learning to forecast (the likelihood of) future events and the consequences of possible actions. Inference and prediction should proceed decision. The DM should learn from all the available data and forecast what is likely to happen if she should take each of the possible actions before her, before committing to one of these actions. Inference, prediction and decision making are, therefore, intimately connected: see French and Rios Insua (2000) for a detailed theoretical exploration of these connections.

Some writers make a distinction between *decision* and *choice*, requiring that decisions are preceded by rational deliberation, while choices are unthinking acts of selection. Thus one might argue that the DM would 'decide' which car to hire, but 'choose' a mint imperial from that bag of sweets. We avoid this distinction because it is hard to maintain in the face of detailed examination. At the car hire firm suppose the DM is offered the choice between a black and a blue, but otherwise identical cars. She might choose blue automatically because she likes blue more than black. There is certainly reason for her choice: but is there deliberation? On the other hand, in looking at the bag of sweets she might see both mint imperials and mint toffees; and she might also much prefer mint toffees. However, she might also have a dental crown is liable to become detached if she chews toffees. Balancing up the threat of a lost crown, discomfort and an expensive visit to the dentist with the additional enjoyment of a mint toffee over a mint imperial, after some heart-searching she selects a mint imperial. What if she weighed things up on leaving the dentist last month and resolved never to eat another toffee. Then, when offered the bag of peppermints, she selects the mint imperial without any reflection. Did she choose or did she decide?

1.5 Data, Information and Knowledge

In everyday language we scarcely distinguish between data, information and knowledge. There is a feeling of increasing value to the user as we pass from data to information to knowledge, perhaps, but no clear distinction.

Following Laudon and Laudon (2001), Marakas (1999) and Turban and Aronson (2001), we define:

- *Data*. Facts about things, events, activities, transactions, etc. not organised for any specific context.
- *Information*. Data organised and, possibly, summarised to be meaningful within a specific context.
- *Knowledge*. Generic information, e.g. scientific understanding, that is relevant to several contexts, together with the skills

and values, which are used in conjunction with information that is more specific to a particular context to solve problems. Having knowledge implies having understanding, experience and expertise.

Whereas data and information can always be made *explicit* and *codified*, i.e. expressible in some permanent form, not all knowledge can. There is a distinction between a person's unexpressed *tacit* knowledge, e.g. a skill such as riding a bicycle, and *explicit* knowledge, such as encoded in an economic model. Many authors suggest that some tacit knowledge can only reside in a human mind, and we have much in sympathy with that view. However, if this is so, then AI will encounter much difficulty in developing true knowledge-based computer systems that operate autonomously from humans: again a view that we have much in sympathy with. We shall argue that KB-DSS can only aid DMs, not supplant them. Even an accepted scientific theory, which can be written down on paper and is hence part of our explicit knowledge, is difficult to codify such that it may be applied automatically. Knowing how to apply a piece of explicit knowledge is itself a tacit skill.

There are many qualifications which we should make. Perhaps the most important is that a piece of information for one person in one context may be quite irrelevant to another person in another context, and so simply an item of data to him or her. Indeed, in another context it may also be irrelevant to the first person and so simply become data again. Knowledge, however, is more long lasting. It includes *generic* information: e.g. the theories and models of science, economics and so on. Theories and models are structures which suggest how data should be organised, inferences drawn and predictions made in a range of contexts. Knowledge includes the generic skills which enable us to form and use information in specific contexts. Information becomes knowledge when the recipient recognises a new 'understanding': see, e.g., Earl (2000). Table 1 provides a summary.

There is a problem, however, with this linear perspective on the progression from data to knowledge. Consider the transformations, The DM begins with raw data, selects and assembles these into information. After doing this in a number of times, she may recognise certain parallels in the information she has used in several of the contexts and draw from this some general understanding, a recognition of a common pattern which she can learn and

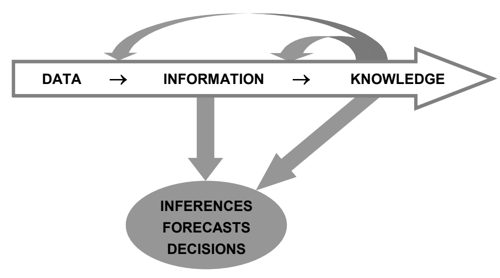


Figure 4: (Previously learnt) knowledge must be applied to transform data to information and information to knowledge; and also to make inferences, forecasts or decisions.

apply in the future. What criteria does she use to select data; how does she organise them to be appropriate to a specific context; how does she reduce her uncertainty in the light of this information; and by what criteria does she recognise common patterns and learn? Her cognitive processing of data into information into knowledge requires that she draws upon her experience, values and past learning to perform these tasks. Thus her creation of knowledge requires the application of her previous knowledge to help her filter and assemble the original data and information. Equally drawing upon information available to her in any context to make inferences, forecasts or decisions requires the application of knowledge too. See Figure 4. Boisot (1998) describes knowledge as providing the perceptual and conceptual filters which the DM uses to firstly to select and organise data into information and then to use that information to support an inference, forecast or decision.

Boisot (1998) also distinguishes between levels of knowledge.

- A *capability* is a general strategic skill which enables the user to deploy her knowledge in a general way. Thus she might be able, say, to design a car or develop a broad strategic direction to take an organisation forward.
- A *competence* is a tactical skill which enables the user to develop the details of an outline design or strategy. Thus she might be able to design the engine of a car to meet certain performance requirements or develop the details of a strategy.
- A *technology* is a more operational skill which enables the user to build the engine or implement aspects of the strategy.

Boisot's conception clearly maps onto the strategic, tactical and operational levels of the strategy pyramid; and thus one is drawn to identify a fourth level of skills that are actually deployed in doing things in the hands-on domain: the professional, analytic or craft skills one deploys on a daily basis. We have in mind the detailed *skills* and *abilities* that we use instinctively, albeit often only after a considerable period of training: e.g. a surgeon sewing up a wound or a mathematician undertaking the detailed calculations necessary to explore the implications of a model. Thus we think of skills and abilities as that knowledge we draw upon in responding to a situation in a recognition-primed manner.

- An *ability* is an instinctive skill such as deployed unthinkingly by a craftsman in executing some task.

1.6 Normative, Descriptive and Prescriptive Approaches

Moving from the contextual issues to the support of decisions, we need to make a distinction between *normative* and *descriptive models*. Normative models suggest how people *should* make decisions; descriptive models describe how they *do*. This distinction has been at the heart of many debates over the years, the central issue being that people seldom *do* make decisions according to the tenets of the normative theories which suggest how they *should*. This tension has lead to a third view on decision making[2], that of *prescriptive* decision analysis and support: see, e.g, Bell *et al* (1988), French and Smith (1997), French and Rios Insua (2000). Prescriptive analyses guide DMs towards a decision by providing models which capture aspects of the issues before them and of their beliefs and value judgements, while at the same time reflecting canons of rationality embodied in the normative theory. These models provide the DMs with perspectives on the issues which bring understanding and through this understanding they reach a decision. In communicating with the DMs and in eliciting their beliefs and value judgements, an analyst needs to understand how they draw inferences and decide intuitively because that is what they will do in answering his questions and understanding his reports. Thus, both normative and descriptive models contribute to prescriptive analyses: see Figure 5.

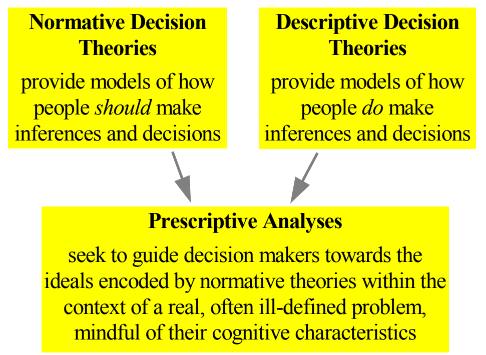


Figure 5: Prescriptive decision support

Descriptive theories are discussed in Chapter 2 and normative theories in Chapter 3. The latter is somewhat mathematical in nature and may be skip read initially. In Chapter 4 we illustrate some techniques that arise from these normative models, and before turning to a fuller discussion of prescriptive analysis in Chapter 5.

Man must learn to simplify, but not to the point of falsification.
Aldous Huxley
Brave New World

We shall often use the word *evolution*. Prescriptive analyses guide the evolution of the DMs' perceptions. During prescriptive analyses their perceptions change. Their perceptions evolve because of the analysis: it is the purpose of the analysis that they should. Thus it is vital to see the modelling process involved as creative, dynamic and cyclic. The DMs' beliefs and preferences are assessed and modelled; the models explored, leading to insights and a revision of their judgements, and thence revision of the models used. The process cycles until no new insights are found. Phillips (1984) describes this evolution, referring to the process as *requisite modelling*, the final model being requisite or sufficient for the inference or decision faced. Too often a static view is taken of decision analysis, in which the beliefs and preferences – indeed all the judgements of the DMs – are taken as fixed and immutable from the outset of the analysis. It is not a view which we shall adopt. We perceive decision analysis and support as providing an iterative framework in which judgements and understanding may evolve.

Although we shall discuss several approaches to modelling and analysing decision problems, the majority, if not all, derive in some sense from what is commonly called the Bayesian School. The key to understanding the Bayesian approach is that it explicitly focuses

[2] Beware: not all writers distinguish clearly between normative and prescriptive, even today.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

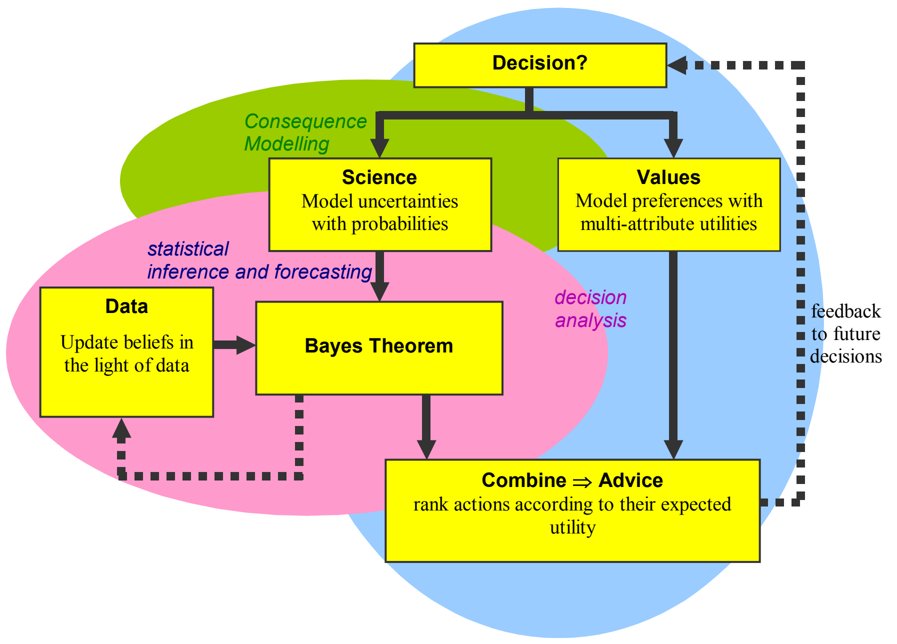


Figure 6: The Bayesian Paradigm

on a single DM and models her judgements. Bayesian methods focus on the individual. They have at their core a normative model of an idealised DM, who faces the same problem as the organisation, stakeholder groups and society as a whole does. She is idealised in that her beliefs and preferences are constrained to satisfy certain consistency conditions, the canons of rationality embodied in the paradigm, which define the Bayesian view of rational decision making. Subject to such consistency conditions, she can express the beliefs and value judgements of any individual.

How such an individualistic view of decision making leads to powerful methods to support groups, organisations and, indeed, society in its decision making we will discuss in later chapters, along with details of the Bayesian paradigm. Here we simply note its outline structure as illustrated in Figure 6.

The Bayesian approach models the DM's uncertainties about what might happen by probabilities and her preferences between the different potential consequences by multi-attribute utilities. Note how this separates the *Science*, predictions of what might happen as a result of possible actions, from the *Value Judgements* of how much each possible consequence matters. Compare this with Figure 3 in which Science and values are also separated. When move from a focus on an individual DM to group or organisational decision making, we shall see that experts become primarily involved in advising on the consequence modelling, while stakeholders have their concerns reflected in the multi-attribute modelling.

Once her uncertainties in the possible consequences are modelled by probabilities, the DM may conduct surveys or experiments to acquire further data in order to reduce her uncertainty. Bayes' Theorem provides the formalism for this – it is the use of this theorem that has led to the name *Bayesian Paradigm*. The process of applying Bayes' Theorem in this way distinguishes Schools of Bayesian Statistics and Forecasting (Barnett, 1999; French and Smith, 1997; Barnardo and Smith, 1994; French and Rios Insua, 2000; Migon and Gamerman, 1999; O'Hagan, 1994).

Her values are modelled by multi-attribute utilities. The term 'multi-attribute' reflects that this modelling allows that may factors or *attributes* may determine her preferences and that she may need to trade-off success in terms of one factor against that of another.

The subjective expected utility (SEU) model suggests that the DM should rank the actions according to their expected (multi-attribute) utility with respect to her probabilities after updating them with any data. We shall emphasise throughout these notes but particularly in Chapter 5 that this ranking advises and in no way prescribes her choice.

There are two further points that we should make about the Bayesian Paradigm as represented in Figure 6. Firstly, the methodology recognises that one decision usually leads to another and thus sets the context for future decision making. For this reason there is a feedback loop; and also theories of sequential decision making (French and Rios Insua, 2000; DeGroot, 1970). Secondly, in many applications one may be led to emphasise the analysis of uncertainty or the analysis of preferences. In particular, there is an extensive theory of *multi-attribute value* modelling in which the analysis of preferences in the absence of uncertainty is studied.

1.7 Decision Support Systems

There are many definitions of what is meant by a decision support system (DSS). We shall adopt the following: a decision support system is a computer-based system which supports the decision making process, helping DMs to form and explore the implications of their judgements and hence to make a decision based upon understanding. We shall emphasise support for the evolution of judgement and understanding, rather than more general support provided by, say, a summary of information in a database. In our view a 'true' DSS is as much about modelling and understanding the perspectives, views, preferences, values and uncertainties of the DMs as modelling external data.

Some categorise DSSs according to whether they are based on a model)., e.g. a linear programme or decision tree, or simply driven by data)., essentially by being built on a database. We shall not follow this route, since all DSSs are built on both, albeit with different emphases.

Primarily we shall categorise DSSs by the domain of managerial activity they support: *viz.* the corporate strategic domain, the general domain, the operational domain and the hands-on work domain – see definitions on page 6. Edwards *et al* (2000) make a similar classification in their discussion of the rôle of expert systems in business decision making. Secondly, we shall note the level of decision support provided, starting from minimal analytic support to full judgemental support.

Level 0:	Acquisition, checking and presentation of data, directly or with minimal analysis, to DMs.
Level 1:	Analysis and forecasting of the current and future environment.
Level 2:	Simulation and analysis of the consequences of potential strategies; determination of their feasibility and quantification of their benefits and disadvantages.
Level 3:	Evaluation and ranking of alternative strategies in the face of uncertainty by balancing their respective benefits and disadvantages.

Table 2: Levels of decision support

Table 2 defines four levels of decision support. The first, Level 0, refers simply to the presentation of data or, to be consistent with our distinction between data and information, the presentation of information. At this level the DSS simply extracts the relevant data from the database and presents it to the DMs with minimal analysis. In our terms, this is borderline for inclusion under the heading of DSS. Level 0 DSS's include executive and management information systems with their graphical and tabular summaries and geographical information systems relating spatial, temporal and factual data. Also included are statistical systems which provide exploratory and inferential analyses, but not statistical forecasting systems.

Level 1 systems take the available data and combine these with judgement, either expressed directly or through the use of one or more models, to forecast how the environment will evolve. Such systems predict the future, but stop short of predicting the consequences of the DMs' potential interventions. Thus here we include, *inter alia*, economic forecasting systems, market share predictions, environmental impact forecasts.

Level 0 and 1 systems do not recognise, *per se*, that the DMs face a decision. In terms of our definition of a DSS, they help the DMs' understanding grow only in relation to the external environment, either as it is (Level 0) or as it is likely to evolve (Level 1). Level 2 systems predict the consequences of the various alternative strategies facing the DMs. But, although they may predict the success of alternative actions against the attributes of concern: i.e. against a number of performance measures, Level 2 systems stop short of prescriptive decision support in our terms. They do not support the process of judgement that the DMs must undergo to make the decision. Finally, Level 3 systems do provide prescriptive support in that they do help the

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

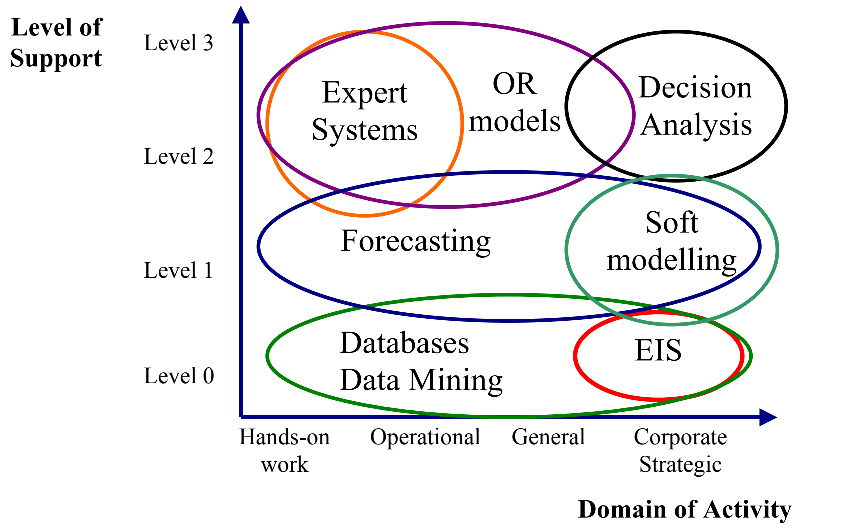


Figure 7: Categorisation of a variety of DSS according to domain and level of support.

DMs explore, evolve and act upon their judgements. They help DMs weigh together conflicting criteria and also balance potential benefits and costs with key uncertainties.

Figure 7 indicates rough categorisation of a variety of DSS tools according to both domain of managerial activity and level of support. Operational research (OR) modelling, e.g. linear programming, inventory models, and project planning tools (see Chapter 6), underpins many of the systems used in general operation and hands-on domains at levels 2 and 3, but OR techniques tend to assume too much structure to be used in the corporate domain. Expert systems, neural nets, and other AI techniques (see Chapter 7), again provide level 2 and 3 support, but are only really suited to the highly structured and repetitive situations found in the hands-on domain (Edwards *et al*, 2000). Databases and data mining can provide level 0 support over the whole range of activities, but are often referred to as executive information systems (EIS) in the case of the higher domains of activity(see Section 7.6). For supporting decision making in the highly unstructured contexts of the corporate strategic domain we need the broad range of tools from the discipline known as decision analysis (see Chapter 5). Moreover we may need soft modelling techniques to provide level 1 support in predicting how a situation may evolve and how different strategies change that evolution. In Chapter 8 we return to a general discussion of DSSs and pull much of the discussion in the earlier chapters together into the framework indicated in Figure 7.

1.8 Background Reading

This course draws together material from at many literatures. Thus many disparate texts serve to provide introductory background reading. We mention only a few. Firstly, we shall look at the behavioural decision science of how people do make decisions. General reading here is provided by Arkes and Hammond (1986), Bazerman (2002), Kahnemann *et al* (1982) and Wright and Ayton (1994). We shall then briefly discuss *decision theory*, which provides the normative models on which much of the support provided to DMs is based. Background reading, albeit of a technical nature, is provided by Bacharach and Hurley (1991), Bouyssou *et al* (2000), French (1986), French and Rios Insua (2000), and Lindley (1973) Drawing these ideas together to provide prescriptive decision support leads us to decision analysis: see Clemen (1996), French (1988), French and Smith (1997), Goodwin and Wright (1999) and Keeney (1992). The book by von Winterfeldt and Edwards (1986) covers decision behaviour, normative theory, and decision analysis. Operational research models are covered by many texts. We note Daellenbach (1994), Denardo (2002), Ragsdale (2001) and White (1985). Computing perspectives on DSSs may be found in Klein and Methlie (1990), Mallach (2000), Marakas (1999), Sauter (1997), Silver (1991), Srinivasan *et al* (2000), and Turban and Aronson (2001). Finally, we note that Watson and Buede (1987) and Kleindorfer *et al* (1993 provide a multi-disciplinary overview.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

2 Decision Making Behaviour

2.1 Introduction

Behavioural decision science is concerned with how people *do* make decisions: i.e. descriptive studies of human behaviour. One classic strand of work began in the 1950's, stimulated by the Allais Paradox (see Section 2.6) and by the seminal paper of Edwards (1954); however, arguably the most influential work is that of Kahneman and Tversky and their co-workers: see, e.g., Kahneman *et al* (1982). They conducted many laboratory studies which compared actual decision behaviour with that predicted by several normative theories, notably the subjective expected utility model (SEU) which will be central to much of our development. Nearly as influential have been the studies led by Janis and Mann (1977) and Simon (1960). These were both founded in analysis of case studies of actual decision making. Moreover, much has been learnt about group behaviour: i.e. how do groups make decisions in practice?

2.2 Psychological biases

Kahneman and Tversky (1974) has been cited in almost all subsequent texts on decision analysis and innumerable research papers since. Essentially, they took the SEU model (see Section 3.8, but you will not need details here) as a representation of rational choice and shown that, left to their own devices, DMs do not instinctively follow the tenets of the model. Moreover, taking probability as a rational representation of the relative likelihood of unknowns and future events, they have shown that people do not organise their uncertainty judgements in such a way. In particular, they revise their beliefs in the light of data quite differently to the prescription of Bayes' Theorem. Their conclusion is that unaided human judgement is susceptible to many biases. We and they are well aware that, in using the word 'bias', there is an implicit assumption that some model of rationality is correct or ideal in some way, but we ignore that issue for the present.

Misconceptions of 'randomness'. In a sequence of ten tosses of a fair coin, HHHHHTTTTT is as equally likely as HHTTTHTHHT to occur, yet many subjects behave as if the latter is the more likely. In making judgements of likelihood, individuals often expect events to 'look' random. For instance, in one experiment (Kahneman *et al*, 1982, p34) 92 subjects were asked the following question. "All families of six children in a city were surveyed. In 72 families, the *exact order* of births of boys and girls was GBGBBG. What is your estimate of the number of families surveyed in which the *exact order* of births was BGBBBB?" Of the 92 subjects, 75 (82%) judged the number to be less than 72. Yet the chances of either sequence of births are equal (assuming that the number of male and female births are about equal across the population). Thus on the information given, the estimate should be 72. Subjects seem to judge the second sequence less likely because it looks less representative.

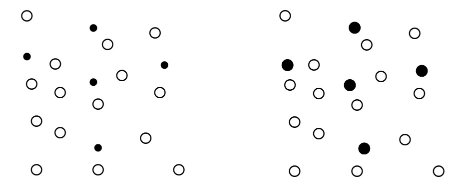
If a coin falls heads repeatedly one hundred times; then the statistically ignorant would claim that the law of 'averages' must almost compel it to fall tails next time. Any statistician would point out the independence of each trial, and the uncertainty of the next outcome. But any fool can see that the coin must be double headed.

Ludwik Drazek

Representativeness, ignoring base rates and overweighing recent evidence. Many problems require that a DM judges the likelihood that an individual belongs to a particular group. Kahneman and Tversky (1974) report several results on such judgements. First, they showed their subjects a number of descriptions of the personalities of individuals. The subjects were told that these descriptions had been drawn at random from the descriptions of 100 individuals, 70 of whom were engineers and 30 of whom were lawyers. The subjects were asked to assess the probability that a given description referred to an engineer. In a similar experiment, the same questions were asked, but the subjects were told that the population consisted of 30 engineers and 70 lawyers. The results in the two experiments were essentially identical. The subjects seemed to judge the probability that a description matched an engineer purely on the basis of their stereotypes of lawyers and engineers, ignoring that in the first population there were more than twice as many engineers as lawyers, whereas these *base rate* odds were reversed in the second experiment.

In a related experiment, subjects were also told that a group of 100 men contained 70 engineers and 30 lawyers. They were asked to judge the probability that one of them, Dick, was an engineer. Comfortingly, most of them gave a probability of 70%. However, they were then told: "Dick is a thirty-year-old man.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB



(a) (b)

Figure 8: An illustration of the availability bias.

Suppose that the subject has to estimate the probability of events • in a population of •'s and ○'s. If there is nothing to distinguish the impact of the events in the subject's memory, as in (a), then the ease with she recalls •'s will reflect their probability. But if the •'s have more impact in her memory, as in (b), they will be easier to recall and hence their probability will be overestimated.

He is married with no children. A man of high ability and motivation, he promises to be quite successful in his field. He is well liked by his colleagues." This description was designed to convey no information relative to the question of whether Dick was an engineer or lawyer. So the subjects should have held to their original assessment of the likelihood that Dick was an engineer. But most changed their assessment to 50%, seemingly agreeing that their was no information in the description, but forgetting their knowledge of the base rates.

Availability. In judging probabilities, subjects often recall series of similar circumstances and ask themselves in what frequency the event of interest occurred. The probability that they ascribe to the event will be highly correlated with the ease with which they recall the events. But subjects may generate very atypical sets of 'similar circumstances' and hence produce biased estimates. For instance, if asked for the probability that a person will fall ill with skin cancer in the next year, an individual may think of what proportion of her friends have such a disease. Since such diseases can be memorably horrific, she may overestimate the incidence by forgetting how many friends she has who do not have the disease. The idea is illustrated in Figure 8.

There is a similar bias relating to the subject's ability to imagine potential futures. A DM may need to assess the uncertainty of an event for which has not occurred before; e.g., in risk analyses, one often has to estimate risk related to foreseeable, but completely unexperienced events, perhaps an air-crash onto a nuclear power station. There is some evidence that the probability subjects ascribe to such events relates to the ease with which they can imagine them – which is subtly different from the ease with which they can happen.

Anchoring. People tend to 'anchor' on the first number they hear or estimate in a given context. The classic example of this is an experiment in which subjects are shown very quickly (too quickly to complete the calculation) either the multiplication problem:

$$1\times2\times3\times4\times5\times6\times7\times8 = ?$$

or

$$8\times7\times6\times5\times4\times3\times2\times1 = ?$$

Estimates of the answers in the first case are consistently smaller than in the second. Subjects anchor on the initial partial products. Similar effects occur if subjects are asked whether they agree with or how they would alter an initial estimate or numerical assessment. Thus an analyst should never ask: "Is the probability of rain tomorrow greater or less than 80%?" because the answer would be nearer 80% than a more neutral question would have elicited. Anchoring effects happen not just in probability assessment, but also in judgementally assessing any quantity.

Calibration. Taken together these effects mean that judgemental assessments of probability are likely to be biased. But how are they biased? Can we assess that bias and what can we do about it? Suppose that a person assesses the probability of a number of events. Gather together all the events for which his estimate was, say, 40%. Now look at the occurrence of these events. If the person is well attuned with reality, one would expect that about 40% of the events would actually happen. Similarly, of those events to which she ascribed a probability of occurrence of 70%, one would expect 70% to occur. Such is not usually the case. Figure 9 shows a graph of a *calibration curve*. This is a plot of the frequency with which an event occurs against the probability ascribed to it by the subject (Lichtenstein *et al*, 1982). Perfect calibration gives the 45° line. Most people depart from this as illustrated in the hypothetical, but typical curve. They tend to overestimate high probabilities: i.e. there is rather less than 100% occurrence of the events which they were certain would occur. Equally, they tend to underestimate low probabilities. A person's calibration curve is context specific: see, e.g.,

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

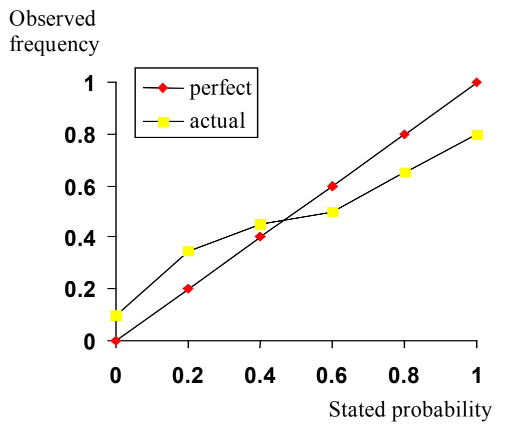


Figure 9: An hypothetical calibration curve.

Lichtenstein *et al* (1982). A financial expert may have quite different calibration in forecasting movements on the London Stock Market than currency rate fluctuations because of differing expertise in the two contexts.

On the subject of assessing zero probabilities, Dennis Lindley has coined the term *Cromwell's rule* for the advice: *ascribe an event zero probability in an analysis at your peril*. It is very dangerous to analyse a decision or inference on the basis that something is impossible, because no number of data to the contrary will change that opinion.

Framing. The manner in which a question is framed can affect the answer dramatically. For instance, Tversky and Kahneman (1981) report an experiment which has been repeated by hundreds of lecturers with their students since, subjects were given the situation shown in Figure 10 and asked to choose between programmes A and B, without being shown options C and D. About three quarters of all subjects prefer A. Other subjects have been presented with the same scenario and asked to choose between programmes C and D, without being shown options A and D. About three quarters of these prefer D. But, of course, the choice in each case is the same: programmes A and C are identical in their outcomes as are programmes B and D.

Thus whether a question is framed positively or negatively matters. Roughly speaking, framing an issue in positive terms (here, in terms of lives saved) leads to avoidance of risk and a preference for certainty, whereas the behaviour is reversed when the issue is phrased negatively (here, in terms of deaths). It should be noted that the same effect has also been observed with less extreme impacts, such as monetary gains and losses.

Imagine that you are a public health official and that an influenza epidemic is expected. Without any action it is expected to lead to 600 deaths. However, there are two vaccination programmes that you may implement.

Programme A would use a well established vaccine which would save 200 lives.

Programme B would use a new vaccine which might be effective. There is a 1/3rd chance that it would save all 600 lives and 2/3rds chance of it saving none of them.

Programme C would use a well established vaccine which would lead to 400 of the population dying.

Programme D would use a new vaccine which might be effective. There is a 1/3rd chance of no deaths and 2/3rds chance of 600 deaths.

Figure 10: The influenza example

Intransitivities of Preference. Transitivity is simply a property of preference which demands that if a DM prefers A to B and B to C then she should prefer A to C whatever A, B and C are. Most normative theories have an implicit assumption that preference should be transitive: see the money pump argument (Figure 11, page22) for a justification. Yet there are many examples of intransitivity arising in decision behaviour, some of which seem to be justified on quite rational grounds.

For example, suppose a company has to decide which of a number of new products to manufacture and launch. For each product two forecasts are available: the expected net present value (NPV) of cash-flows and the expected market share. Both forecasts are subject to error. NPV is more important as a decision criterion to the company than market share. Thus it might decide to choose between pairs of products according to higher NPV unless the difference is less than $50,000. If the NPVs differ by less than $50,000, the choice will be made by going for the higher expected market share. It is very easy to construct examples using this rule which exhibit intransitivities. The rule is known as the *lexicographic semi-order* decision model.

Tversky (1969) performed some experiments which indicate that the lexicographic semi-order does model describes one way in which subjects form their preferences when there are conflicting factors that they wish to balance. Other heuristic rules can also lead to intransitive behaviour and many psychological studies have provided empirical support for such descriptive decision models.

2.3 Janis and Mann's Theory

The influential book of Janis and Mann (1977) had the aim of providing "a comprehensive descriptive theory of how people actually cope with decisional conflicts." Note the verb 'cope'. Their aim was not primarily to describe the cognitive process of deliberation leading to a choice; rather they were interested in the DMs' behaviour in handling stress and conflict. In a sense, this perspective runs counter to our direction: we are concerned with the process of deliberation. Nonetheless, their studies identified five primary patterns of behaviour, one of which – *vigilance* – corresponds to the staged analytic pattern that prescriptive decision analytic techniques seek to foster and support.

Janis and Mann suggest that a good decision making process goes through eight stages. Failure to pass through one or more would constitute a flaw in the process. A person who tackles a decision and passes through each of these stages is said to be *vigilant*.

1. *Thorough canvassing of alternatives.* The DM investigates alternatives to her present course of action – the *status quo*.
2. *Thorough canvassing of objectives*. The DM pauses and reflects upon what she is really trying to achieve: her objectives and goals: *c.f.* value focused thinking (Keeney, 1992).
3. *Careful evaluation of the consequences of current policy.* What is she likely to achieve under her present course of action? What might she achieve if things do not run according to plan, if her forecasts of the external world do not come to pass?
4. *Careful evaluation of the consequences of new policies.* What may happen if she changes her course of action? Are alternative policies more likely to bring her better consequences in the future?
5. *Thorough search for information*. The DM seeks out information to support her decision making. Such information may help shape her beliefs about what her actions might lead to, about how other players may react and generally about how the future may evolve.
6. *Unbiased assimilation of new information.* The DM takes account of new information and (expert) advice that she receives in a balanced way, according each due weight.
7. *Careful re-evaluation of consequences*. The DM re-examines the positive and negative consequences of all alternatives before her in the light of any new information.
8. *Thorough planning for implementation and contingencies*. She commits to her decision and plans for its implementation, paying careful attention to the ways in which she will deal with the hazards that she has considered in her deliberations.

From their studies, Janis and Mann identified three patterns of good decision making and two aberrant ones. Vigilance, we have noted, passes through all eight stages and represents an ideal. Sometimes, however, good decision making need not be so wide-ranging; the course of action may be obvious. Thus, they also applaud *unconflicted adherence*, in which none of the incoming information causes the DM to doubt her present course of action. There are no serious risks and she calmly maintains her present policy. Equally, the incoming information may lead her to realise her present course is not adequate but there is an obvious alternative which counters the new risks and to which she can calmly move in a process of *unconflicted change*.

As opposed to this rational organised approach to decision making, Janis and Mann found two poor decision making behaviours. Firstly the DM may avoid receiving some disturbing information or distort its implications by wishful thinking. She is then said to exhibit *defensive avoidance*. Alternatively, the DM may become involved in a frantic and ill organised search for a way out of her difficulties. In common parlance, she may panic or come close to it. This Janis and Mann term *hypervigilance*.

There are many other patterns of behaviour which might be adopted but generally Janis and Mann did not observe them. Nor have subsequent studies (Maguire *et al*, 1997). However whether or not other patterns exist matters little to us. Our task in prescriptive decision analysis and support is to help DMs recognise straightforward circumstances in which unconflicted adherence or unconflicted change are appropriate and equally to recognise more complex circumstances when vigilant decision making is called for and help them move calmly through each of its stages.

2.4 Bounded Rationality

We have been implicitly assuming that DMs seek to make an *optimal* decision in some sense. H.A. Simon in his seminal work on *bounded rationality* argues that this need not be so (Simon, 1960). There are two broad groups of criticisms directed against seeking optimality. Firstly, doing so makes extraordinary demands upon the time and information processing capability of DMs. It assumes that both time and information are unlimited resources. Secondly it assumes that preferences are well defined and unproblematic. Even for individuals choosing for themselves, this is a dubious assumption. In many circumstances we encounter choices which may lead to consequences which we have not imagined before. Then we need to think through what our preferences are. Not all our tastes and values lie completely defined in our subconscious: some have to be constructed or developed when we find a need to articulate them in particular choices. In organisations, the definition and construction of values is even more problematic. Empirical studies suggest that conflict and, hence, ill-definition of communal values are endemic (House and Singh, 1981). Once defined, values are not necessarily stable. Some evolve over time, adding to the complexity of decision making. Optimality against such evolving preferences is a nebulous concept.

DMs are well aware of these limitations and seek to live within them. They know that they have not the time nor the discriminatory and intellectual powers to define, much less find a perfect optimum to a problem. So they *satisfice*: i.e. search for a course of action that is satisfactory in meeting reasonable aspirations. Satisficing behaviour is commonly observed in organisations. Indeed, Simon's theory of bounded rationality is closely entwined with his views on organisational decision (Simon, 1960).

2.5 Group Behaviour

Most decisions are the result of interactions of groups of individuals. Thus understanding and recognising the group context of many decisions is important. Groups may place many pressures on individual members, e.g.

- *Pressures to conform*. Individual members may not voice disagreements with a growing consensus for fear of 'rocking the boat' in some way or fear of the power of a superior or an influential grouping.
- *Individual status*. Some members with ambitions of leadership may moderate or misrepresent their views to fit better with the politics of the group.
- *Factions*. Similarly, a subgroup may try to control the issue, excluding the rest of the group, either overtly or more subtlety, perhaps by using a jargon which they share but the rest do not (fully) understand.
- '*Them and us*'. The group members see themselves as different – usually, as having a higher status in some sense – to those outside the group.
- *Societal prejudices*. Since a group is a microcosm of a larger society, any prejudices which exist outside the group may be carried into it and lead to behaviour changes in those who discriminate or are discriminated against.

The effects of these pressures are various. The most well know is perhaps that of *groupthink*. This phenomenon, documented by Janis (1972), tends to afflict decision making groups which are:

- highly cohesive;
- insulated from many external influences;
- lacking in procedures for evaluating and reviewing alternatives;
- under the influence of a strong, directive leader;
- under some stress, maybe because of the urgency or importance of the decision.

The result may be a collective pattern of defensive avoidance with a reluctance to acquire further relevant data, biased information processing of that data which is to hand and incomplete canvassing and evaluation of alternatives. The symptoms of groupthink are:

- a false belief in the invulnerability of the group;
- a common belief in the innate morality of their decision;
- direct internal pressure to conform;
- an unquestioned and unanimous rationalisation of their choice.

Groupthink has been offered as an explanation of many disastrous or near disastrous events: the Cuban missile crisis being a commonly quoted example.

Choice 1: Which of the following options would you choose?			
Option A:	£1 000 000	for certain	
Option B:	£5 000 000	with probability	0.10
	£1 000 000	with probability	0.89
	£0	with probability	0.01
Choice 2: Which of the following options would you choose?			
Option C:	£1 000 000	with probability	0.11
	£0	with probability	0.89
Option D:	£5 000 000	with probability	0.10
	£0	with probability	0.90

Table 3: The Allais Paradox

Other 'irrationalities' in decision making in groups are legion: see, e.g., House and Singh (1987). For instance, individuals may fight to join a group (thereby enhancing their status in some way?), yet then not take part in the decision making as fully as they might. Groups may seek more information than is strictly needed for the decision – and then ignore this information when it is available. Justifications for a course of action may be developed retrospectively to the choice rather than *vice versa*.

Perhaps the most emotively named descriptive theory is the *garbage can model*. If one subscribes to a theory of rational decision making, one might hope that at least the sequences of stages in the decision making process are ordered sensibly, even if the stages themselves are limited by the bounded nature of human cognition. Sadly, empirical observation does not find such logical ordering pervasive. Rather, decision making in organisations seems to be based more upon the random "confluence of independent streams of problems, solutions, DMs and choice opportunities" (House and Singh, 1987). In short, which decisions get taken in organisations is more the result of the timing and juxtaposition of choices, problems and DMs than of rational design. This may seem an observation quite contrary to the apparent success of the majority of organisations. Surely, they cannot be that chaotic! However, simulation and empirical studies have shown that, despite the randomness, important problems are more likely to be resolved than unimportant ones. Organisations do 'muddle through' (Lindblom, 1959).

2.6 The Allais Paradox

In 1952 Maurice Allais presented the example given in Table 3. Pause for a minute and think about which option you would choose in each of the two choices. A majority of individuals choose Option A in the first choice and Option D in the second. They argue that Option A makes them 'rich' beyond their wildest dreams so why should they risk the small chance (1%) in Option B of receiving nothing. In the second choice, however, there is roughly the same high probability of their receiving nothing whichever option they choose. So they select Option D which has the possibility of the larger prize.

Sensible and rational though these arguments sound, there are strong *prime facie* arguments why choosing A in the first choice is inconsistent with choosing D in the second. For instance, Savage (1954) offers the following argument. Consider the 'implementation' of Allais' Paradox illustrated in Table 4. Imagine 100 lottery tickets placed in a hat. One will be drawn at random and the prize in each option allocated as illustrated. Now in the first choice between Options A and B, there is no difference in the outcome on tickets 12-100. The distinction between the options arises only on the outcomes when tickets 1-11 are drawn. Similarly in the second choice between Options C and D, there is no difference in the outcome on tickets 12-100. The distinction between the options arises only on the outcomes when tickets 1-11 are drawn. Moreover, the pattern of outcomes on tickets 1-11 for Options A and B is the same as that for Options C and D. So consistency would seem to suggest that if one chooses A in preference to B, then one should choose C in preference to D. Equally if one chooses B in preference to A one should choose D in preference to C.

This argument based upon Table 4 reflects one of the basic normative assumptions of SEU theory, namely the *sure thing principle*. This

	Lottery ticket number		
	1	**2-11**	**12-100**
Option A:	£1 000 000	£1 000 000	£1 000 000
Option B:	£0	£5 000 000	£1 000 000
Option C:	£1 000 000	£1 000 000	£0
Option D:	£0	£5 000 000	£0

Table 4: Allais' Paradox explicated in terms of a lottery

suggests that in ranking options a rational DM should ignore outcomes in which the consequences are identical and focus only on those in which they differ. In the abstract this principle seems persuasive. Yet, over the years many behavioural experiments have been conducted and demonstrated that many people hold a preference for A over B and for D over C, even when participants have been exposed to arguments which articulate the sure thing principle as in Table 4 (Slovic and Tversky, 1974).

Further discussions of this 'Paradox' may be found in Allais and Hagen (1979), French (1986), French and Rios Insua (2000), and French and Xie (1994).

3 Decision Theory

3.1 Introduction

Decision theory may be defined as the study of (mathematical) models of the judgements involved in and leading to deliberate choice, usually – and certainly for us – rational, deliberate choice. We shall be interested in formulating the canons of rationality referred to in Section 1.6 in mathematical terms and identifying models of preference and beliefs which are consistent with them. We shall not venture too far into the technical depths of decision theory, avoiding stating the principles axiomatically and giving full proofs of our assertions. Rather we shall try to give an overview of the ideas and refer to the literature for the details.

We begin by categorising three types of decision problem. This categorisation is more important perhaps for its place in the history of decision theory than for modern applications; but it does introduce terminology which will be useful. We then turn to modelling preference in the absence of uncertainty. Next comes the modelling of uncertainty and, finally, we draw all the elements together into the SEU model.

3.2 Decisions under Certainty, Risk and Strict Uncertainty

Generally in a decision there are some things under the DM's control and some beyond it. The latter are called *exogenous factors* by economists and *states of the nature* by statisticians – we shall call them *states*. How much the DM knows about these states may vary between contexts and this leads to a further classification of decisions.

Decisions under certainty. In these the DM either knows or learns the 'true' state before she has to make her choice. Thus there is no uncertainty in her decision: she simply has to choose the option that brings her the best outcome. Of course, identifying which outcome she feels is best may not be trivial,

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

requiring her to balance conflicting objectives: e.g. safety cannot usually be maximised as the same time as profit.

Decisions with risk. Although the DM does not know the true state for certain, she does have some knowledge which makes some of the possible states seem to her to be more likely than others.

Decisions under strict uncertainty. Here the DM feels that she can say *nothing at all* about the true state. She is only prepared to identify what states may be possible.

One representation of decision problems use is a *decision table* or *consequence table*: see Table 5. We shall encounter alternative representations, e.g. decision trees and influence diagrams, which are useful in applications, but for discussions of decision theory the decision table is sufficient.

There are some things under the DM's control and some beyond it. The former define the *action space*, $A = \{a_1, a_2, \ldots, a_m\}$: i.e. the set of options from which the DM may choose. The latter, the *states*, form the *state space*, $\Theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$. The DM will receive a *consequence*, c_{ij}, lying in some *consequence space*, C, determined both by the chosen action a_i and the state θ_j that pertains: *viz*.

$$\begin{array}{ccccc} \text{Action} & \oplus & \text{state} & \rightarrow & \text{consequence} \\ a_i & \oplus & \theta_j & \rightarrow & c_{ij} \end{array}$$

If she knows the state that actually holds, which we shall call the *true state* without venturing into philosophical questions of what is 'truth', then she can predict the consequence of her choice with certainty: i.e. she is facing a decision under certainty and the decision table essentially reduces to a single column. If she knows nothing about which state holds, then the problem is one under strict uncertainty. If she does have some beliefs about the true state but is uncertain about which it is, she faces a problem with risk.

We shall concentrate the support of decisions under certainty and with risk, but be almost silent on decisions under strict uncertainty. Many attempts have been made to characterise strict uncertainty, but ultimately the task has proved impossible. At its heart strict uncertainty is a self-contradictory concept: to define the states about which she 'knows nothing', the DM must know something (see, e.g., French, 1986, Chapter 2).

		State			
		θ_1	θ_2	…	θ_n
	a_1	c_{11}	c_{12}	…	c_{1n}
	a_2	c_{21}	c_{22}	…	c_{2n}
Actions	.	.	.	…	.
	.	.	.	…	.
	.	.	.	…	.
	a_m	c_{m1}	c_{m2}	…	c_{mn}

Table 5: A decision table

3.3 Preference Orders and Value Functions

We begin with a mathematical model of preference rankings or orderings that underpins all the normative theory that we shall discuss. We focus on decisions under certainty. The same ideas will be used later in discussions of decision making under risk. Each available action leads to an unambiguous consequence, and the DM has full knowledge of everything that she considers relevant to her problem. Our purpose is, thus, to discuss and model a rational DM's preferences between the possible consequences. These preferences completely determine her choice of action, for we assume that a rational person will always choose an action that leads to a most preferred[3] consequence; and thus for the present we talk about her preferences between consequences and actions interchangeably. In modelling preference we must be careful to avoid dictating the actual preferences that a rational DM should hold. For instance, it would be wrong to demand that all rational people prefer tea to coffee. However, we shall demand that a rational person's preferences should be mutually compatible. For instance, if she prefers ice cream to raspberry jelly and, in turn, prefers raspberry jelly to apple pie. Then surely she must prefer ice cream to apple pie. It is with this and similar requirements that we begin our study.

We shall write $a \succeq b$ to x mean the DM *weakly prefers a* to *b*. An alternative, perhaps more expressive interpretation is that she holds *a* to be at *least as good as b*. Operationally we take this to mean that, if offered the choice of *a* and *b*, she would not be disappointed if she were

[3] A trivial but often forgotten point: there is no reason for there to be a unique most preferred action.

An agency has three secretaries a, b, c on its books and the DM has interviewed them all and (despite our better judgement) prefers a to b, b to c and c to a. Suppose that between the interviews and the appointment, c becomes unavailable. Her choice now being between a and b, the DM will pay the agency and employ a. Next the agency 'discovers' that c was not unavailable after all, although b, having not been selected, has gone off after another job. The agency has c; the DM is employing a; and she prefers c to a. The agency will not find it difficult to persuade the DM to swap a for c and, moreover, to pay the agency a suitably small charge, say a penny, for the privilege. At this point the agency 'discovers' that b did not get the other job after all, but that a is no longer available having suddenly succumbed to a terrible cold. Since the DM prefers b to c, she will need little persuasion to part with a further penny and swap c for b. Needless to say, there is a miraculous recovery on a's part, but not before c has caught the cold while visiting a's sick bed. Inevitably, the DM pays a further penny and swaps b for a. We leave the story as the cycle begins afresh.

Figure 11: The money pump argument

forced subsequently to take a. Let A be the set of objects over which the DM's preferences are expressed. We will make two specific demands on the consistency that we expect of the rational use of $\succeq$.

Firstly, we demand that $\succeq$ is *comparable*: *viz.* for all objects a, b in A, the DM holds either $a \succeq b$ or $b \succeq a$. Comparability may be restated as: there is no pair of objects a, b in A such that the DM holds neither a to be a least as good as b nor b to be at least as good as a. In yet other words, if we do not assume comparability there may be a pair of objects such that, if offered the choice between them, the DM would feel disappointment if she were subsequently forced to accept either one. In such a case it would appear that the act of choosing is more important to the DM than the receipt of the object of this choice. It may be true descriptively that people ascribe more value to the act of deciding than to the consequences of their decision, but it does not seem rational that they should do so. Or, rather, if we charge someone with making a decision for us it is immaterial to us whether or not she enjoys her task; our concern is with the result of her decision making.

Secondly, we demand that her preferences are *transitive*: namely, for all objects a, b, c in A, if $a \succeq b$ and $b \succeq c$, then $a \succeq c$. The assumption of transitivity seems more than reasonable, but it cannot be motivated other than by an appeal to self evident good sense. In the case of strict preference (see below), a simple money pump (Figure 11) argument suggests that transitivity should hold; but without strict preference there is no imperative for the DM to pay to swap one secretary for another.

There are two further preference orders related to weak preference: *indifference* and *strict preference*.. We shall write $a \succ b$ to mean that the DM *strictly prefers* a to b; in other words, if she were offered a straight choice between a and b, she would be disappointed if she were forced subsequently to take b. We shall use the notation $a \sim b$ to mean that the DM is *indifferent* between a and b; in other words, she is equally happy to receive either a or b.

We shall demand that a rational DM uses the notions of weak preference, strict preference and indifference in a consistent fashion. Specifically, she considers $a \succ b$ if and only if she holds $a \succeq b$ and does not hold $b \succeq a$; i.e. she considers a to be at least as good as b but not *vice versa*. Also, she considers $a \sim b$ if and only if she holds $a \succeq b$ and $b \succeq a$; i.e. she considers each to be at least as good as the other. It may be shown that because of the demand of this consistency, both indifference and strict preference are transitive, that indifference is symmetric, i.e. the DM holds $a \sim b$ if and only if she holds $b \sim a$, and that strict preference is asymmetric, i.e. if the DM holds $a \succ b$ then she does not hold $b \succ a$ (see, e.g., French, 1986). All these results seem sensible, with the money pump argument (Figure 11) giving normative weight to the conclusion that $\succ$ is transitive.

Indifference classes are sets of objects between which the DM is indifferent. Indifference classes (or curves) are constructs well known to economists. The definition of indifference made here ensures that indifference classes have the properties assumed of them by economists, namely that they are equivalence classes: see, e.g., French (1986).

The similarity between weak preference $\succeq$ and the numerical ordering $\geq$ cannot have passed unnoticed; and there is much to be gained from exploiting this similarity. It allows us to model preferences numerically. We say that $v(.)$ is an (*ordinal*) *value function* representing the DM's preferences if $v(.)$ is a real-valued function on

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

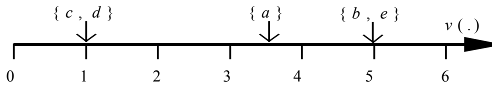


Figure 12: Ordinal value function representation of the preferences.

A such that $v(a) \geq v(b)$ if and only if $a \succeq b$; and we say that $v(.)$ *agrees with or represents* $\succeq$ *over A*. Note that the adjective ordinal emphasises that such value functions only represent orderings: they do not, e.g., represent strength of preference, for which we need more specific forms of value function. Unless we need to be precise, we shall often drop 'ordinal' and simply refer to value functions.

The advantages of a value function representation of preferences are twofold. Firstly, it is very compact. To represent a set of preferences over n objects we need only n real numbers; our great familiarity with the real line means that we instinctively know the ordering of any two numbers. Secondly, our discussion can become conceptually easier. Most of us find it simpler to identify a most preferred object by maximising a value function than by searching the set of objects, A, to find a maximal element, $a_{\max}$, such that $a_{\max} \succeq a$, for all a in A, even though the two tasks are essentially the same.

Notwithstanding the advantages that a value function brings, we must be careful, because we are only using the ordering of the real line in this representation; addition, subtraction, multiplication and division, for instance, have no part to play. It is meaningless to ascribe any interpretation to the mean value of $v(.)$ over a group of objects; yet it is very tempting to try to do so.

Example 1

Consider a set of preferences over five objects, $b \sim e \succ a \succ c \sim d$. These preferences may be represented by an ordinal value function as in Figure 12. Thus:

$$v(b) = v(e) = 5 > v(a) = 3.5 > v(c) = v(d) = 1.$$

Note that instead of choosing 1, 3.5 and 5 we could have chosen any increasing sequence of numbers such as -1, 0 and 29. Comparing these two representations, *viz.* $v(.)$ as above and

$$w(b) = w(e) = 29 > w(a) = 0 > w(c) = w(d) = -1$$

we may confirm our earlier remarks about the danger of reading too much into the numerical representation. The mean of $v(.)$ over the five objects is quickly calculated as 3.1, which is less than $v(a) = 3.5$, whereas the mean of $w(.)$ is found to be 11.2, which is greater than $w(a) = 0$. So we cannot meaningfully say that a is worth more or less than

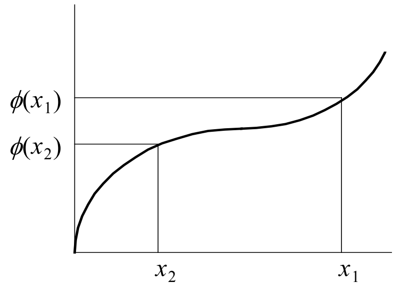


Figure 13: A strictly increasing function

the average. Similarly, $(v(a) - v(c)) > (v(b) - v(a))$ but $(w(a) - w(c)) < (w(b) - w(c))$, so we cannot ascribe a consistent meaning to value differences. It is meaningless to say that the increase in value of a over c is greater or less than that of b over a. At least, these statements are meaningless under the assumptions that made so far.

This and other similar examples quickly show that an ordinal value function represents only the ranking of objects in terms of preference: value differences, mean values, etc. are meaningless. Although ordinal value functions *per se* will not be central to our methods, the issue of meaningfulness will be. There have been many suggestions for normative decision models which have falsely drawn on some numerical properties in the quantitative representation that do not correspond to the properties assumed of the underlying preferences (or beliefs).

Meaningfulness is closely related to the *uniqueness* of the value function representation. In Example 1 two functions represented the same preferences. It can be shown that two ordinal value functions represent the same weak preference order if and only if they are related by a strictly increasing function: *viz.* $v(.)$ and $w(.)$ both represent the same weak preference $\succeq$ if and only if $v(a) = \phi(w(a))$ for all objects a, where the real function ϕ is such that for all x_1, x_2: $\phi(x_1) > \phi(x_2) \Leftrightarrow x_1 > x_2$, see Figure 13 (French, 1986; Krantz *et al,* 1971; Roberts, 1979). We say that ordinal value functions are *unique up to strictly increasing transformations*. Indeed, the adjective 'ordinal' means precisely that this 'uniqueness' holds.

3.4 Measurable Value Functions

In Example 1 we also saw that differences in the numerical values assigned to objects were meaningless. The fact that in one ordinal value representation $(v(a) - v(c)) > (v(b) - v(a))$ does not mean that a is preferred to c more than b is preferred to a, because in another

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

representation agreeing with the same preference the numerical differences may be ordered in the opposite sense. However, we have the imperatives of everyday language to consider notions of strength of preference. We could appeal to an innate feeling of strength in an individual's preferences, but we take a more action oriented approach. We say that a DM prefers a to b more than she prefers c to d if and only if she would prefer to give up b in exchange for a than to give up d in exchange for c. Let $(a \leftarrow b)$ represent the *exchange* of b to receive a. The use of this rather cumbersome notation emphasises that it is the complete act of exchanging b for a that is being denoted. There are now two types of entity about which the DM may express preferences: the objects themselves and exchanges between objects. Thus there are two weak preference relations.

$\succeq$ – weak preference between objects;

$\succeq_e$ – weak preference between exchanges.

As before, $a \succeq b$ means that the DM holds a to be at least as good as b; and, now $(a \leftarrow b) \succeq_e (c \leftarrow d)$ means that she holds the exchange of b to receive a to be at least as good as the exchange of d to receive c.

We define a *value difference function* or a *measurable value function* $v(.)$ to be such that:

$$a \succeq b \iff v(a) \geq v(b)$$

and

$$(a \leftarrow b) \succeq_e (c \leftarrow d) \iff v(a) - v(b) \geq v(c) - v(d).$$

Clearly this demands that a measurable value function is an ordinal value function; i.e. we are specialising the concept of an ordinal value function. What extra assumptions or are necessary and sufficient to justify this specialisation? Firstly we still need to demand that $\succeq$ is comparable, transitive and that $\succeq$ is related to $\succ$ and $\sim$ consistently in the sense discussed above. We need to demand precisely the same of $\succeq_e$ in relation to exchanges. We also need to demand that further consistency properties. For any object c, $(c \leftarrow c)$ is a null exchange. The DM exchanges c for itself and, hence, is no better or worse off. So we demand that for all objects a, b and c, $a \succeq b \iff (a \leftarrow b) \succeq_e (c \leftarrow c)$. If the DM feels that the gain in giving up b to receive a is no less than the gain in giving up d to receive c, she must also feel that the loss to her in giving up a to receive b is no less than the loss in giving up c to receive d. Hence we demand for all objects a, b, c and d $(a \leftarrow b) \succeq_e (c \leftarrow d) \iff (d \leftarrow c) \succeq_e (b \leftarrow a)$. Next consider exchanges via an intermediate object: i.e. first $(b \leftarrow a)$ then $(c \leftarrow b)$. We demand that if the DM holds the exchange $(a \leftarrow b)$ to be at least as good as $(d \leftarrow e)$ and the exchange of $(b \leftarrow c)$ to be at least as good as $(e \leftarrow f)$, then she must hold the exchange $(a \leftarrow c)$ to be at least as good as $(d \leftarrow f)$: i.e. for all objects a, b, c, d, e and f:

$$\left.\begin{array}{l}(a \leftarrow b) \succeq_e (d \leftarrow e) \\ (b \leftarrow c) \succeq_e (e \leftarrow f)\end{array}\right\} \Rightarrow \quad (a \leftarrow c) \succeq_e (d \leftarrow f).$$

A little algebra shows that these assumptions are necessary if a value difference function is to represent simultaneously weak preferences between objects *and* exchanges. But the assumptions are not sufficient.

Consider a sketch of how a measurable value function might be constructed to represent a DM's preferences for monetary gains. Assume that she is fully aware of her current assets. The objects a are positive sums of money that she will be given to increase her assets. The exchange $(a \leftarrow b)$ means that she is first told that she will receive b, but subsequently before she is given b, she is told that she will receive a instead. We shall assume that the DM always strictly prefers larger gains to smaller ones.

The first step in the construction of $v(.)$ is to define the unit of measurement. Pick an arbitrary positive gain a_1 and set $v(0) = 0$ and $v(a_1) = 1$. For consistency of notation, we define $a_0 = 0$, so that $v(a_0) = 0$, $v(a_1) = 1$. Next we ask the DM to state a sum of money a_2 such that she holds $(a_2 \leftarrow a_1) \sim_e (a_1 \leftarrow a_0)$, i.e. such that she would be equally happy to exchange a_1 for a_2 as to exchange a_0 for a_1. Since $v(.)$ is to be a measurable value function,

$$v(a_2) - v(a_1) = v(a_1) - v(a_0),$$

Hence $v(a_2) = 2$. Having identified a_2, we now ask the DM for a_3 such that $(a_3 \leftarrow a_2) \sim_e (a_2 \leftarrow a_1)$ and then for a_4 such that $(a_4 \leftarrow a_3) \sim_e (a_3 \leftarrow a_2)$ and so on. In this manner we identify a sequence of gains $a_0, a_1, a_2 \ldots$ such that:

$$(a_1 \leftarrow a_0) \sim_e (a_2 \leftarrow a_1) \sim_e \ldots \sim_e (a_n \leftarrow a_{(n-1)}) \sim_e \ldots.$$

Moreover, since $v(.)$ is a measurable value function, the assignments imply that

$$v(a_0) = 0;\ v(a_1) = 1;\ v(a_2) = 2;\ \ldots;\ v(a_n) = n;\ \ldots.$$

Suppose that we wish to find $v(a)$ for some point a in A. If $a = a_n$ for some n, then $v(a) = n$, and we are done. If $a \neq a_n$ for any n, we proceed as follows. We assume that for some n, $a_{(n+1)} \succeq a \succeq a_n$. This important assumption, to which we will return shortly, implies that

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

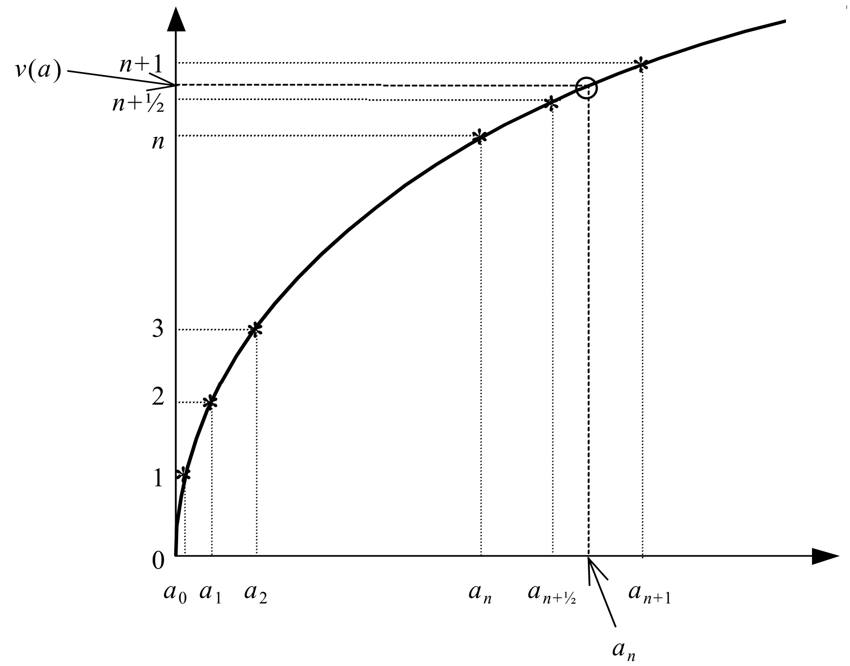


Figure 14: The procedure for constructing $v(.)$. The points marked $*$ on the curve are determined directly by the DM's indifferences. Values such as $v(\mathrm{a})$, marked ○, may be approximated to any desired accuracy.

$n + 1 \geq v(a) \geq n$. Next we ask the DM for $a_{(\mathrm{n}+\frac{1}{2})}$ such that $(a_{(n+1)} \leftarrow a_{(\mathrm{n}+\frac{1}{2})}) \sim_{\mathrm{e}} (a_{(\mathrm{n}+\frac{1}{2})} \leftarrow a_{\mathrm{n}})$, i.e. such that she would be equally happy to exchange $a_{(n+\frac{1}{2})}$ for $a_{(n+1)}$ as to exchange a_n for $a_{(n+\frac{1}{2})}$. In this sense, $a_{(n+\frac{1}{2})}$ is the 'preference midpoint' of the interval between a_n and $a_{\mathrm{n}+1}$. Since $v(.)$ is a measurable value function, $v(a_{(n+\frac{1}{2})}) = n + \frac{1}{2}$. Now either $a_{(n+1)} \succeq a \succeq a_{(n+\frac{1}{2})}$ or $a_{(n+\frac{1}{2})} \succeq a \succeq a_n$. So $(n + 1) \geq v(a) \geq (n + \frac{1}{2})$; or $(n + \frac{1}{2}) \geq v(a) \geq n$. Suppose that the first case holds. Then we ask the DM for $a_{(n+\frac{3}{4})}$ such that $(a_{(n+1)} \leftarrow a_{(n+\frac{3}{4})}) \sim_{\mathrm{e}} (a_{(n+\frac{3}{4})} \leftarrow a_{(n+\frac{1}{2})})$. Continuing 'halving' appropriate intervals, we may determine the numerical value of $v(a)$ to within an accuracy of $(\frac{1}{2})^k$ for any desired value of k. Figure 14 gives a diagrammatic representation of the procedure.

The procedure is very intuitive. The exchange $(a_1 \leftarrow a_0)$ defines a 'unit of preference' and, using this, the set of monetary gains is marked off into 'unit intervals of preference'. Just as the notches on a ruler divide it into equal units of length, perhaps inches, so the sequence a_0, a_1, a_2, a_3,… divides the set A into equal units of preference; and just as the inches are subdivided into half-inches, quarter-inches, eighths, etc., so the intervals of preference $(a_n \leftarrow a_{(n-1)})$ are subdivided into halves, quarters, eighths, etc.

The procedure is so intuitive that it is difficult to see that it takes much for granted. First, throughout we have assumed that the DM can answer questions of the following two types:

- For gains b, c, d, what gain a satisfies $(a \leftarrow b) \sim_{\mathrm{e}} (c \leftarrow d)$?
- For gains b, c, what gain a satisfies $(b \leftarrow a) \sim_{\mathrm{e}} (a \leftarrow c)$?

In assuming that she can answer these questions we are making *solvability assumptions*. Solvability refers not so much to the DM's preferences, but to the set of possible gains, assuming here that monetary gains are sufficiently divisible that she can always find a to satisfy the requested indifference.

We have made a further assumption, namely given any a there exists an integer n such that she holds $a_{(n+1)} \succ a$. This assumption is known as an *Archimedean assumption* and effectively demands that no monetary gain exists of infinite value to the DM.

The sequence a_0, a_1, a_2,… is an example of a *standard sequence*. The defining property of standard sequences is that they are equally spaced in some sense.

We have laboured the discussion of measurable value functions a little to give a flavour of (axiomatic) decision theory and measurement theory (French, 1986; French and Rios Insua, 2000; Krantz *et al*, 1971; Roberts, 1979). Essentially, these investigate the necessary and sufficient assumptions – or axioms – that justify a numerical representation of preferences and beliefs. In any analysis we should always explore and justify our assumptions; and it is easy to introduce unseen assumptions in the process of quantification.

We noted that ordinal value functions were unique up to strictly increasing transformations. Measurable value functions are *unique up to positive affine transformations***.** ; i.e. *,* $v(.)$ and $w(.)$ are two measurable value functions both agreeing with the same preferences $\succeq$ and $\succeq_e$ if and only if there exist real numbers α and β with $\alpha > 0$ such that $w(a) = \alpha v(a) + \beta$ for all objects a. Because of this 'greater uniqueness', it turns out that differences in values, means, standard deviations, etc. are meaningful – at least in some senses: see French (1986).

3.5 Multi-Attribute Value Functions

So far in our discussion the set of alternatives has not had a structure: an alternative a has simply been an object that is available for choice. We have assumed implicitly that the DM associates a complete description of the underlying object with each symbol a. However, she may feel that carrying a complete picture of each in her mind is not only unnecessary, but also confusing, since it clouds the issue with many irrelevancies. Thus we now consider how she might describe the alternatives against a number of criteria or attributes.

The DM's perceptions of and preferences for possible alternatives are usually a mixture of complex, conflicting values. She wants more profit, more safety, less environmental impact, etc. – and she cannot have them all. She has to make trade-offs. The term *attribute* means one of the factors which need to be taken into account in a decision. In Section 4.2 we discuss the structuring introduces into decision modelling by the use of multiple attributes.

Suppose then that the DM may choose between alternatives described by vectors of achievement against a number of attributes: $\boldsymbol{a} = (a_1, a_2,\ldots, a_q)$. Throughout we assume that each attribute is real valued; however, this assumption is not strictly necessary. Economists may find a familiar interpretation of such structuring of alternatives in terms of *commodity bundles***,** $\boldsymbol{a} = (a_1, a_2, \ldots, a_q)$, with a_1 units of commodity 1, a_2 units of commodity 2, etc. Another economic interpretation occurs when $(a_1, a_2, \ldots, a_q)$ represents a timestream of cash-flows over q years.

How does this structuring of alternatives as vectors of attribute levels affect the assessment of value functions? What reasonable conditions might hold in some circumstances such that $v(a_1, a_2, ..., a_q) = v_1(a_1) + v_2(a_2) + ... + v_q(a_q)$, i.e. when might the q-dimensional function $v(.)$ be formed as the sum of q one-dimensional functions? We refer to such a representation as an *additive* (*multi-attribute*) *value function*. The functions $v_i(a_i)$, known as *marginal value functions*, serve as ordinal value functions on each of the attributes.

Essentially, an additive value function is justified if and only if the DMs judge the attributes to be *preferentially independent*. Preferential independence is a technical concept which formalizes the very common (but *not* universal) feature of preferences embodied in the following statements.

- All other things being equal, more money is preferred to less
- All other things being equal, greater safety is preferred to less
- All other things being equal, less environmental effect is preferred to more.

Stated formally: a subset of the attributes is *preferentially independent* of the remaining attributes if the preference between any pair of alternatives which differ *only* in their levels of achievement on attributes within the subset do not depend on the levels of achievement on the remaining attributes. Note that, because the alternatives differ *only* in terms of their achievement on the subset of attributes, they attain precisely the same levels of achievement on the remaining attributes. If this is true for all subsets, the attributes are said to be (*mutually*) *preferentially independent***.**

A little thought shows the importance of preferential independence. Without it, there is no possibility of defining – let alone assessing – the attribute value scales, $v_1(a_i)$. Only when preferential independence holds can one talk of preferences for one different levels of achievement on one attribute independently of another. In any decision analysis, therefore, it is imperative that preferential independence assumptions are checked. The facilitator and

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

analyst will generally do this while the attribute hierarchy is being constructed.

In fact, there is an element of 'chicken and egg' in the process of choosing attributes and checking for preferential independence, since the choice of attributes is closely entwined with the validity of the independence assumption. The folklore among many decision facilitators and analysts is that it is usually possible to select appropriate attributes in a problem such that preferential independence holds: see Section 4.2. Full discussions of preferential independence may be found in, e.g., French (1986) and Keeney and Raiffa (1976).

Much as in the case of measurable value functions, further Archimedean and solvability assumptions are necessary, although we do not indicate these in detail since they relate more to the richness of the decision space than specific assumption about the form of the DM's preferences. Additive value functions are unique up to positive affine transformations: again see either of the texts cited above.

It is possible to identify further assumptions so that an additive value function quantifies strength of preference and moreover, the marginal value functions, $v_i(a_i)$ are also measurable value functions on each of the attributes (Dyer and Sarin, 1979a; French, 1986, Section 9.3).

A multi-attribute value function is *linear* if $v(a_1, a_2,\ldots, a_q) = w_1a_1+w_2a_2+\ldots+w_qa_q$. The coefficients $w_1, w_2, \ldots, w_q$ are known as *weighting factors* or, simply, *weights.* Linear value functions are commonly assumed in many areas of economics, commerce and operational research. Comparing timestreams of cash-flows according to net present value (NPV) assumes a linear value structure in which:

$$v(a_1, a_2,\ldots, a_q) = a_1 + \rho a_2 + \rho^2 a_2 + \ldots + \rho^{q-1} a_q$$

However, there is extra structure here over that assumed by linear value theory: the weighting factors are related by $w_i/w_{i-1} = \rho^{i-1}/\rho^{i-2} = \rho$ for $i = 2, 3, \ldots, q$.

Cost-benefit analysis, a form of decision analysis commonly used by government agencies, assumes a linear value structure – at least, it does in its most naïve form. The distinguishing assumption of cost-benefit analysis is that every attribute of an alternative can be given a financial value, positive or negative. Alternatives are compared according to their total financial value, $w_1a_1 + \ldots + w_qa_q$,

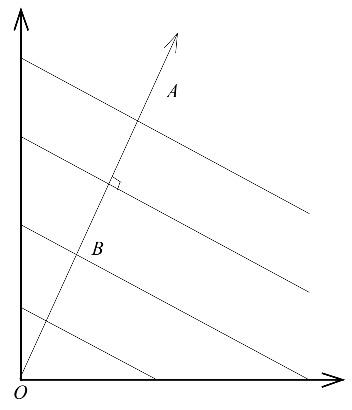


Figure 15: The necessity for a monotonicity condition.

in which w_i is the financial value of one unit of the i^{th} attribute.

Clearly linear value functions are additive. So an agreeing linear value representation can only exist when the attributes are mutually preferentially independent. But we need a further condition[4]. We shall say that there is a *constant relative trade-off* of γ_{ij}:1 between a_i and a_j if and only if $(a_1, a_2, \ldots, a_i, \ldots, a_j, \ldots, a_q) \sim (a_1, a_2, \ldots, a_i + \gamma_{ij}\varepsilon, \ldots, a_j - \varepsilon, \ldots, a_q)$ for any ε, positive or negative. If $v(\boldsymbol{a}) = \Sigma_i w_i a_i$, the constant relative trade-off between a_i and a_j is (w_j/w_i). Clearly the assumption of linearity implies that there are constant relative trade-offs between all pairs of attributes.

The constant relative trade-offs condition requires that there are constant relative trade-offs between all pairs of attributes. The essential implication is that the indifference curves are parallel straight lines or hyperplanes. This is almost, but not quite, sufficient to characterise a linear value function. Consider the two-dimensional case: Figure 15. *OA* is the common perpendicular to the indifference curves. A linear value function insists that preference either increases monotonically or decreases *monotonically* from *O* to *A*. It is not possible for, say, *B* to be the most preferred point on the line *OA*.

In summary, to use a linear value function representation of a DM's preferences requires the assumption of the constant relative trade-offs condition and monotonicity. Additive value functions are unique up to positive affine

[4] In fact, the constant relative trade-offs condition implies mutual preferential independence.

transformations: $\phi(z) = \alpha z + \beta$ ($\alpha > 0$). If we wish to maintain a linear structure, we must further insist that $\beta = 0$. This is necessary because, if $v(.)$ is linear, $v(0,0,\ldots,0) = 0$.

3.6 Concepts of Probability

So far we have discussed decisions under certainty. We now turn to decisions under risk; i.e. when the possible states are uncertain but the DM believes that she can judge some more likely than others. This means that we shall need to introduce probability into the mathematical model. In order to begin this, let us consider what we mean by probability.

Attaching a meaning to probability is not as trivial as many think. For instance, in the case of gambling, surely everyone would agree that the probability of a die landing 'six' is a sixth … or would they? Suppose it is known that someone has loaded the die; what then is the probability? And, if that question does not give pause for thought, what about the case in which it is *suggested* that someone might have loaded the die? If we turn from gambling contexts and think about the sort of problems faced by DMs in government, industry, etc., then we see that the difficulty in identifying probabilities becomes far greater. How, for instance, might the probabilities needed in evaluating investment options on the stock exchange be found? Indeed, how would you define conceptually the probability of the stock market falling over the next year?

The Classical Notion of Probability

Early writers on probability were somewhat pragmatic in approach, eschewing clear definitions. However, their intentions were summarised by Laplace (1825). He took the probability of an event to be the ratio of the number of possible outcomes favourable to the event to the total number of possible outcomes, each assumed to be equally likely. He assumed that it was possible to divide the future into *n* equally likely primitive events. In considering the throws of a die, he divided the possible future into six events: 'the die lands 1 up', 'the die lands 2 up', etc. Each of these he considered to be 'equally likely events: and few would disagree with him, until you ask questions such as those above – what happens, for instance, when the die is 'loaded'?

In a very real sense, the classical definition of probability is circular. It requires a partition of equally likely events, and surely *equally likely* is synonymous with *equally probable.* So to define the probability of one event we need to recognise equality of probability in others. However, this would not be so serious a flaw providing that we could find a method of recognising equally likely events without involving some concept of probability. We may define equally likely events as being those for which there are no grounds for favouring any one *a priori* (i.e. before the throw of the die etc.) – put another way, those for which there is no *relevant* lack of symmetry. Laplace expressed this idea in his famous, now infamous**,** *Principle of Indifference* or *Principle of Insufficient Reason*, which in modern terminology asserts: if there is no known reason, no relevant lack of symmetry, for predicting the occurrence of one event rather than another, then relative to such knowledge the events are equally likely.

There are two serious difficulties with this principle. It is seldom applicable: how would you divide the future up into events without relevant lack of symmetry to judge the probability of 'the FT100 rising by 63 points' tomorrow? More fundamentally, what does 'no relevant lack of symmetry' mean? Relevant to what? If you answer 'relevant to judging the events to be of different probability', you enter an infinite regression since you would have to recognise 'relevance to probability' without having defined what probability is. See, e.g., French (1986, Chapter 6) and Barnett (1982) for further discussion. So the Classical notion is at best inapplicable to the majority of decision making situations and at worst philosophically unsound.

The Frequentist Notion of Probability

Having discarded the classical interpretation, let us consider the or, rather, *a frequentist* approach; there is no single frequentist approach, but a family of similar ones, sharing the same 'flavour'. The common thread in these approaches is that probability can only have a meaning in the context of an infinitely repeatable experiment. The probability of an event is taken to be its long run frequency of occurrence in repeated trials of the experiment. Consider repeated throws of a die, a single throw being a trial of the experiment. Suppose that we observe the results of many throws. The results shown in Figure 16 would not seem unexceptional. The proportion, i.e. the relative frequency, of sixes might well settle down to 0.1666…= 1/6; indeed, this is what we would expect of a 'fair die'. If the die were weighted then the proportion might be 0.412, say. Of course, no die can be tossed infinitely often, however a frequentist hypothesises that it can.

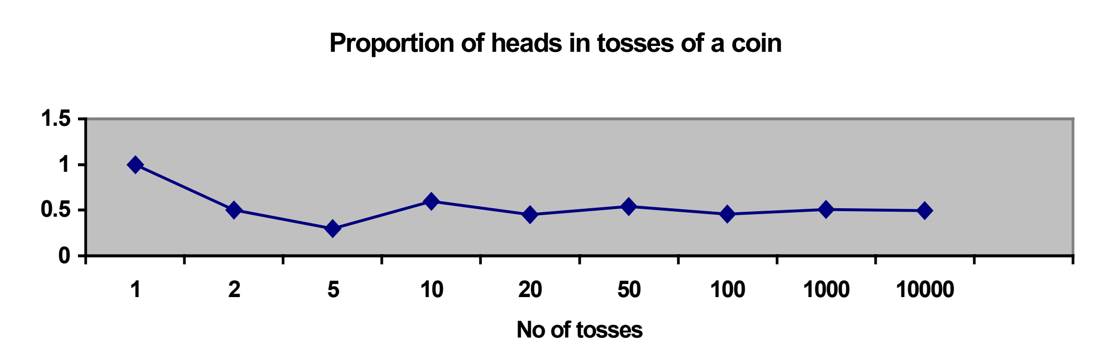


Figure 16: The proportion of sixes in repeated throws of a die.

Frequentist concepts of probability underpin much of standard statistical practice, at least as it was developed in the 1930's to 1970's. Nowadays, Bayesian approaches are becoming more common. These are based upon subjective concepts of probability which we discuss below. Here we note two points that are key to our thinking on frequentist concepts in relation to decision analysis and support. First, it is absolutely essential that the experiment should be repeatable. In many decision contexts, the situation is unique and far from repeatable, thus rendering frequentist approaches inappropriate. Second, a frequentist probability is a property of the system being observed; its value is completely independent of the observer of the system.

Subjective Probability and Degrees of Belief.

The *subjective* schools of probability are known as such because they associate probability, not with the system under observation, but with the observer of that system. For instance, consider a decision table (Table 5). Probability is taken as representing the DM's degree (or strength) of belief in what an unknown state is: $P(\theta_j)$ represents the her degree of belief in θ_j being the state that actually pertains; the stronger her belief, the greater $P(\theta_j)$. Different people have different beliefs. Thus different observers, different DMs, may assign different probabilities to the same event. Probability is, therefore, personal; it belongs to the observer.

The subjective school of probability is, like the frequentist, a family of approaches, each sharing the same 'flavour'. There are many ways in which the meaning of $P(\theta_j)$ may be defined operationally. We shall outline only in the broadest of details one of these approaches here. Others may be found in, e.g., DeFinetti (1974, 1975), Fine (1973), French (1986), French and Rios Insua (2000).

The starting point is to assume that given any two states or events[5]. A and B not necessarily mutually exclusive, the DM has an inherent feeling of *relative likelihood* and so can say whether she holds

A to be more likely than B,
A to be equally likely as B,
A to be less likely than B.

Note that we do not demand that the DM say how much more likely one event is than another, only that she rank them in order of her perception of their likelihood.

Some writers, ourselves included, feel that it is not necessary to define this intuitive ranking any further. We claim simply that anyone can meaningfully answer questions of the form: do you think it is more, less, or equally likely that it will snow tomorrow rather than rain.

We shall use the following notation.

$A \succeq_\ell B$ – the DM believes A to be at least as likely to occur as B.

$A \succ_\ell B$ – the DM believes A to be strictly more likely than B to occur.

$A \sim_\ell B$ – the DM believes the A and B to be equally likely to occur.

It is possible to make very reasonable assumptions about the consistency of the DM's judgements of relative likelihood which allow us to construct probabilities with the property:

$$A \succeq_\ell B \Leftrightarrow P(A) \geq P(B).$$

---

[5] Note that sometimes we talk of *states* and others of *events*. To a student of probability there is a serious distinction here; but for our purposes there is little difference. When we are concerned with external happenings, then it seems more natural to talk in terms of events; when we are concerned with decisions in which the external 'state of the world' is key, then state seems a more natural terminology.

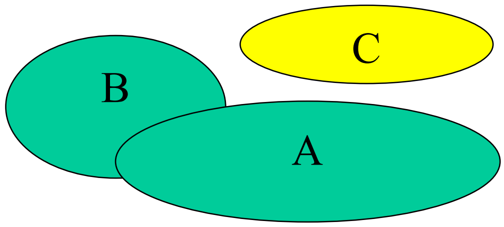


Figure 17: Consistency of belief relative to common but disjoint events

These *normative* assumptions represent the consistency that the DM should aspire to. In practice, her unaided judgements of relative likelihood might well be subject to the many inconsistencies discovered in discussed in Section 2.2. There are three key assumptions – canons of rationality – that we believe the DM should wish her judgements to obey:

1. For all events A, B and C, $A \succeq_\ell B$, $B \succeq_\ell C$ implies $A \succeq_\ell C$: i.e. if she holds A to be at least as likely as B and B to be as least as likely as C, then she *should* hold A to be at least as likely as C. The relations should be transitive.
2. If it matters to her, she can form a judgement between *any* two events
3. $\forall A, B, C$ with $A \cap C = \varnothing = B \cap C$,
   $$A \succeq_\ell B \Leftrightarrow A \cup C \succeq_\ell B \cup C.$$
   i.e. under the assumption that neither A and C can happen together nor B and C, if she holds A as likely as B then she should hold A or C as likely as B or C. See Figure 17.

The next step is the development introduces a *reference experiment* which enables her to make judgements about probability via comparisons between the evens of interest and events in a (hypothetical) experiment for which she 'knows' the probabilities[6]. Imagine, for instance, that she compares an event E of interest (e.g., rain tomorrow) with an event A based upon a probability wheel: see Figure 18. Does she think it more likely that the spinning arrow will stop in the sector A than E will occur?

If these assumptions are accepted as sensible criteria which describe the consistency expected of rational beliefs, then it can be shown that the DM should represent her uncertainty by probabilities. See, *inter alia*,

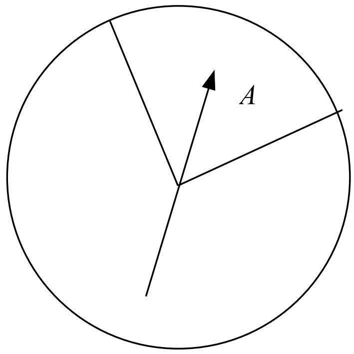


Figure 18: A probability wheel

DeGroot (1970), French (1986) and French and Rios Insua (2000) for full developments.

The reference experiment also provides an operational mechanism for assessing the DM's probability. She can be asked to compare her belief in an external event E happening with the event that the spinning arrow stops in sector A. The size of the sector A can be adjusted until she believes them to be equally likely, thus determining her subjective probability for E. Or she can be asked to choose between gambles of £100 if E (respectively, A) happens and nothing otherwise.

The subjective interpretation of probability gives us a tool for quantifying belief and uncertainty in decision analysis because it can be applied to *unique* sets of circumstances. There is no need for a repeatable experiment; we can discuss, for instance, the DM's probability that the stock market increases by more than 25 points tomorrow.

3.7 Preferences over Lotteries

In the world of gambling there are many simple, easily-understood examples in which simple decisions with risk have to be made and in which the outcomes are uncertain. Thus we continue our discussion of decisions with risk there. Suppose that the DM has to choose between a number of gambles in which the outcomes are determined solely by some simple and fair chance mechanisms. Thus the probability of any outcome is well defined and uncontroversial. We shall refer to these as *lotteries*.

We lose little by assuming that only a finite number of prizes are possible. Let $C = \{c_1, c_2, \ldots, c_r\}$ be the set of possible prizes. In particular, we assume that one of the prizes

6 The reference experiment corresponds to the standard sequence we introduced in the construction of measurable value functions (Section 3.4).

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

is the 'prize' of losing, i.e. of winning nothing. Moreover, we assume that the prizes include the deduction of any stake. Thus, if the DM pays £1 for a lottery ticket and wins a teddy bear, her prize is 'an increase in her possessions of a teddy bear and a decrease in her cash assets of £1'.

A typical lottery will be represented by

$$l = \langle p_1, c_1; p_2, c_2; \ldots; p_r, c_r \rangle$$

where $p_i \geq 0$ is the probability of winning c_i ($i = 1, 2, \ldots, r$) and $\Sigma_i p_i = 1$. It is quite possible that several $p_i = 0$, indicating that certain prizes are not possible in a particular lottery. We shall refer to such lotteries as *simple lotteries*, because, once the chance mechanism is resolved, the prize is determined. We shall also assume that the DM is prepared to consider compound lotteries. In a *compound lottery* some or all the 'prizes' may be entries into further lotteries. For instance, the compound lottery

$$\langle q_1, l_1; q_2, l_2; \ldots ; q_s, l_s \rangle$$

gives probabilities $q_i \geq 0$ of winning an entry into lottery l_i ($i = 1, 2, \ldots, s$; $\Sigma_i q_i = 1$). Such compound lotteries are often found in real life: raffles in which some of the prizes are premium bonds or national lottery tickets. We will allow lotteries to be compounded several times. Since a lottery may give rise to a prize immediately in C or to an entry into a further lottery, we shall refer to the *outcomes* of a lottery rather than prizes. A *direct outcome* is one that results from the single randomisation that governs the lottery. The *ultimate prizes* of a lottery are those members of C which may ultimately result from a compound lottery once all the chance mechanisms have been resolved.

We shall assume that the DM has to choose between lotteries in a set L. These lotteries may be simple or compound. However, we shall assume that all lotteries are finitely compounded. A *finitely compounded lottery* is one which yields prizes from the set C after a finite number of randomisations. We shall let A be the set of all possible prizes together with a set of simple and finitely compounded lotteries that contains the set L. Thus C and L are subsets of A. Note that A will contain lotteries that are not members of L; however, for the present we shall avoid specifying what these additional lotteries may be.

In considering the DM's preferences between the members of A, we shall make several reasonable assumptions concerning the consistency of her preferences if she is to be considered rational. We shall show that these assumptions imply the existence of a *utility function* $u(.)$ on C such that the DM holds:

$$c_i \subseteq c_j \Leftrightarrow u(c_i) \geq u(c_j) \text{ for any } c_i, c_j \text{ in } C,$$

and

$$\langle p_1, c_1; p_2, c_2; \ldots; p_r, c_r \rangle \succeq \langle p'_1, c_1; p'_2, c_2; \ldots; p'_r, c_r \rangle$$

$$\Leftrightarrow \quad \textstyle\sum_i p_i u(c_i) \geq \sum_i p'_i u(c_i)$$

for any pair of simple lotteries in A. The first condition shows that $u(.)$ is an ordinal value function on the set of prizes C; the second condition shows that $u(.)$ possesses the *expected utility property* on the set of simple lotteries. The assumptions will also justify choosing between compound lotteries according to the expected utility rule.

The first assumption that we make is that the DM's weak preferences, strict preferences and indifferences over A should obey the assumptions we discussed in Section 3.3, in particular, weak preference between lotteries and prizes should be comparable and transitive.

For convenience and without loss of generality, we shall label the prizes such that $c_1 \succeq c_2 \succeq_{\ldots} \succeq c_r$. Since there is little to be gained from discussing a situation in which a DM does not care which prize she receives, we shall assume that she strictly prefers c_1 to c_r.

Even though we have set our discussion in the context of simple gambling situations, we shall not allow our rational DM to enjoy gambling in certain specific senses. In very rough terms, we shall not allow her any enjoyment from watching a roulette wheel spin or dice being thrown, other than the enjoyment that she gains from any prize she might win. The chance mechanism which gives rise to the probabilities will be irrelevant.

Our next assumption, which we shall refer to as *reduction of compound lotteries*, also denies any value to an aspect of the chance mechanism itself. Consider the compound lottery $l = \langle q_1, l_1; q_2, l_2; \ldots ; q_s, l_s \rangle$ which gives as prizes entries into further simple lotteries $l_1, l_2, \ldots, l_s$, where

$$l_j = \langle p_{j1}, c_1; p_{j2}, c_2; \ldots ; p_{js}, c_s \rangle$$

for $j = 1, 2, \ldots, s$. Let l' be the simple lottery $\langle p_1, c_1; p_2, c_2; \ldots; p_r, c_r \rangle$, where

$$p_i = q_1 p_{1i} + q_2 p_{2i} + \ldots + q_s p_{si} \text{ for } i = 1, 2, \ldots, r.$$

Then the DM must be indifferent between l and l': *viz.* $l \sim l'$. To understand the import of this

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

assumption notice that p_i is the probability that the prize c_i will ultimately result from the compound lottery l. Thus the assumption is simply demanding that the DM's preferences depend only upon the ultimate prizes and the probabilities with which they are obtained; the number of chance mechanisms involved in generating these probabilities is irrelevant.

In the presence of the other assumptions, this reduction of compound lotteries assumption has an implication, which we state now but prove later. Consider the lottery $\langle 0, c_1; 0, c_2; \ldots; 1, c_i; \ldots; 0, c_r \rangle$, i.e. the lottery which gives 100% chance of receiving c_i and no chance of receiving anything else. It seems reasonable to suppose that the DM is indifferent between simply being given c_i and entering this lottery:

$$c_i \sim \langle 0, c_1; 0, c_2; \ldots; 1, c_i; \ldots; 0, c_r \rangle$$

for all $i = 1, 2, \ldots, r$.

It might be argued, indeed many have argued, that in ignoring the thrill of gambling our theory loses something. Many people do enjoy watching the spin of a roulette wheel to see whether they win, quite independently of the prize that they might win. Visitors to casinos often place imaginary bets just for the pleasure of seeing whether they would have won. Equally some may have such moral objections to gambling that each spin of the wheel is abhorrent to them. However, while these observations are undoubtedly true, they are, we would contend, irrelevant to our present argument. We are not developing a descriptive theory of decision making, and certainly not a descriptive theory of gambling. Rather we are developing a normative theory of decision making. How *should* a DM choose in the face of uncertainty? Our ultimate aim is to develop a style of decision analysis that is appropriate to problems such as the siting of nuclear power stations, budgeting decisions in industry, etc. In such context we would not think it rational for a DM to allow her enjoyment of watching chance mechanisms being resolved to influence her decision.

Our next assumption, which we call *substitutability*, says that if the DM is indifferent between two objects in A then she does not mind whether she wins one or the other in a lottery. To be precise, let b, c in A be such that the DM holds $b \sim c$. Let l in A be any lottery, simple or compound, such that:

$$l = \langle \ldots; q, b; \ldots \rangle,$$

i.e. there is a probability q that b is a direct outcome of l. Let l' be constructed from l by substituting c for b and leaving all other outcomes and all probabilities unchanged, *viz*.

$$l' = \langle \ldots; q, c; \ldots \rangle.$$

Then the DM holds $l \sim l'$. There are a number of points that should be noted about this assumption. Firstly, b, c in A, so each may be a prize or a lottery. Secondly, q is the probability that b is a *direct* outcome. It is not the probability that b is an *indirect* outcome. Similarly, the only difference between l and l' is that c has been substituted for b as a *direct* outcome. If other outcomes in l are entries into further lotteries which in turn give b as an outcome, then c is *not* substituted for b in these; i.e. c is not substituted for b as an indirect outcome.

At first sight, substitutability seems uncontroversial. If $b \sim c$, how can the DM mind whether she receives b or c as a result of a lottery? But consider. Suppose that b is a prize and c a lottery. Then substituting c for b increases uncertainty, because at least one more chance mechanism may have to be resolved before the ultimate prize of the lottery is determined. Given this extra uncertainty it is perhaps reasonable for the DM to have a preference between l and l'. However, although this argument convinces some, it fails to convince us. In holding $b \sim c$ the DM must surely already have allowed for the uncertainty in c. Does the uncertainty inherent in c change in some way when it is substituted into a further lottery? We think not.

The set A contains both the set of prizes C and the set of lotteries L between which the DM must choose. We have also indicated that it contains some further lotteries, and the time has come to explain what these are. We shall assume that the DM is prepared to consider certain hypothetical lotteries of the form

$$c_1pc_r = \langle p, c_1; 0, c_2; 0, c_3; \ldots; 0, c_{r-1}; (1 - p), c_r \rangle,$$

that is a simple lottery which gives rise to c_1, the most preferred prize in C, with probability p, and c_r, the least preferred prize in C, with probability $(1 - p)$; any other prize is impossible. Since we shall need to refer to such lotteries constantly in the next few pages, we use the shortened notation: c_1pc_r.

It is easy to see how the DM might visualise such lotteries. She need only imagine a probability wheel with the background divided into two sectors such that the angles θ and $(360° - \theta)$ are in the ratio $p:(1 - p)$: see Figure 19. The lottery c_1pc_r is visualised by imagining that the pointer is spun and that the

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

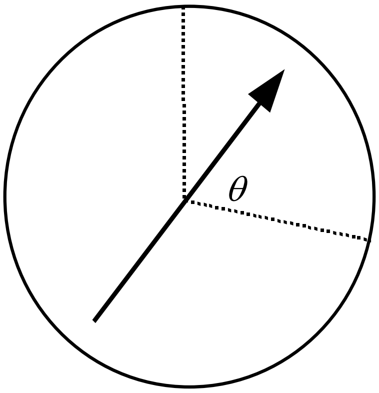


Figure 19: A probability wheel

prize c_1 awarded if it stops in the sector with angle θ and the prize c_r awarded if it stops in that with angle (360° - θ).

We shall assume that the DM is prepared to imagine and to consider her preferences for such lotteries for all possible values of p, $0 \le p \le 1$. As we shall see, by this assumption we introduce into the problem a reference scale or 'ruler' against which the DM can measure her preference. The set of lotteries $\{c_1pc_r \mid 0 \le p \le 1\}$ is known as the *reference* or *auxiliary experiment*, and a lottery of the form c_1pc_r as a *reference lottery*. We assume that all these reference lotteries lie in A: *viz.*

$$c_1pc_r \text{ is in } A \text{ for all } p,\ 0 \le p \le 1.$$

We are now in a position to state the structure of A. It comprises all the prizes in C, all the lotteries in L, all possible reference lotteries c_1pc_r together with all finitely compounded lotteries that may be constructed by substituting for an outcome of a lottery any prize or reference lottery in A that is indifferent to that outcome. This will become clearer as the discussion progresses.

The introduction of hypothetical reference lotteries is often attacked on the grounds that it requires the DM to consider her preferences between these imaginary objects and the real objects of choice. Why should she have such preferences? It is surely not rational to ask a DM to daydream about what can never be. However, this argument makes a false and emotive contrast between what is real and what is imaginary. In a sense, all the alternatives in a decision problem are imaginary until one is selected: and, at that point, the decision problem ceases to exist, because the choice has been made. The selected alternative become real and the unselected alternatives become not just imaginary but impossible, since the alternatives in a decision problem are mutually exclusive. The true difference between the reference lotteries and the lotteries in L is that circumstances have motivated the DM to consider the objects in L. She may choose one and so affect her future. Thus she is motivated to consider her preferences between the lotteries in L. She is not motivated to think about her preferences between the reference lotteries. But suppose that we provide that motivation. Suppose that we show her that by thinking about the reference lotteries she may clarify her preferences in L and help herself towards a better decision. Then surely that will motivate her sufficiently to consider seriously her preferences over the whole of A and not just over L.

Our next assumption, *monotonicity*, states something that is completely uncontroversial. We assume that the DM's preferences between two reference lotteries are such that she prefers the lottery that gives her the greater probability of winning c_1, the best prize, and, therefore, also the lesser probability of winning c_r, the worst prize: *viz.*

$$c_1pc_r \succeq c_1p'c_r \iff p \ge p'.$$

For our final assumption we return to the controversial. To introduce it we consider an example. Suppose that c_1 is £100, that c_r is £0 and that some prize c_i is £40. Consider reference lotteries c_1pc_r for different values of p. For large values of p, say $p = 0.9999$, it is likely that the DM prefers the lottery to having £40 for certain: *viz.*

$$£100(0.9999)£0 \succ £40.$$

(The parentheses in £100(0.9999)£0 have been introduced to clarify the notation c_1pc_r when numerical values have been substituted.) Similarly for small values of p, say $p = 0.0001$, it is likely that the DM prefers having £40 for certain to the lottery: *viz.*

$$£40 \succ £100(0.0001)£0.$$

Consider a sequence of reference lotteries as p increases from 0.0 to 1.0. Initially the prize £40 is preferred to the lotteries; but as p increases, this preference reverses. This argument suggests strongly that there is an intermediate value of p such the DM is indifferent between the lottery and having £40 for certain. See Figure 20.

In general, we make the following *continuity* assumption: for all c_i in X there exists u_i, $0 \le u_i \le 1$, such that $c_i \sim c_1u_ic_r$.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

We have chosen to use u_i rather than p_i to denote the probability in the reference lottery which gives indifference with c_i, because the utility function, whose existence we shall shortly show, is such that $u(c_i) = u_i$.

Note also that the continuity assumption shows the value of u_i to be unique. Suppose that there were two values, u_i and u_i', such that

$$c_1 u_i c_r \sim c_i \sim c_1 u_i' c_r.$$

Then either $u_i > u_i'$, which implies that $c_1 u_i c_r \succ c_1 u_i' c_r$, or $u_i' > u_i$, which implies that $c_1 u_i' c_r \succ c_1 u_i c_r$, both of which contradict the assumed indifference.

There are two important criticisms of continuity. Firstly, many argue that there may be prizes such that for no value of u_i does the DM hold $c_1 u_i c_r \sim c_i$. For instance, suppose that c_1 = £1, c_i = £0 and c_r is the DM's death. Then surely for any value of $u_i < 1$ the DM would strictly prefer to receive £0 for certain than to take the lottery with its risk of her death: at best the lottery can only make her £1 better off. If $u_i = 1$, then £1(1)death ~ £1 $\succ$ £0, since preferences clearly increase with monetary value. Thus there is no value of u_i such that £1(u_i)death ~ £0.

Persuasive though this argument is, it hardly bears inspection. Suppose $u_i = (1 - 10^{-20})$; the lottery then gives a 1 in 10^{20} chance of death. The argument above suggests that the DM would not take this risk just for the chance, admittedly very high chance, of making £1. But each day we all take far greater risks for far less substantial gains. For example, crossing the road brings a risk of death far greater than 1 in 10^{20}; and many people cross the road just to be in the sun. There are many things that we would refuse to do if we objected to the slightest risk of death; yet we do them. We shall allow our rational DM to do them too.

The second criticism accepts that, in principle, a value u_i exists such that $c_1 u_i c_r \sim c_i$, but argues that in practice no DM would ever have the discrimination to give it a precise value. Descriptively this is undoubtedly true. However, we are developing a normative theory and in an *ideal* world the DM should be able to give a precise value of u_i. In discussing sensitivity analysis in Section 5.3, we shall discover that practical decision analysis based upon expected utility theory does not demand such precision.

In the above we assumed that

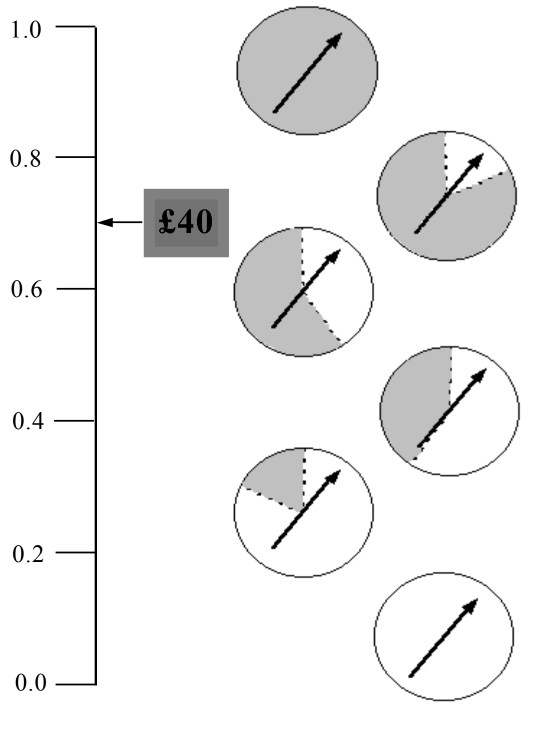


Figure 20: Illustration of the idea underlying the continuity assumption.

The DM considers her preference for a certain prize of £40 relative to a sequence of reference lotteries £100(p)£0 as p increases from 0.0 to 1.0. Probability wheels are shown explicitly for p = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0. In each case the unshaded sector yields £0 and the shaded sector yields £100. For the sake of example, the DM is assumed to hold £40 $\succ$ £100(p)£0 for $p < 0.7$ and £100(p)£0 $\succ$ £40 for $p > 0.7$. The indifference point is £40 ~ £100(0.7)£0.

$$c_i \sim \langle 0, c_1; 0, c_2; \ldots; 1, c_i; \ldots; 0, c_r \rangle$$

for all i = 1, 2, ..., r, which we have claimed, but not shown, is implied by our other assumptions. It is time to rectify that omission. Consider $\langle 0, c_1; 0, c_2; \ldots; 1, c_i; \ldots; 0, c_r \rangle$. By continuity there is a u_i, $0 \le u_i \le 1$, such that $c_i \sim c_1 u_i c_r$. Substitute $c_1 u_i c_r$ for c_i in the lottery. Thus

$$\langle 0, c_1; 0, c_2; \ldots; 1, c_i; \ldots; 0, c_r \rangle$$
$$\sim \langle 0, c_1; 0, c_2; \ldots; 1, (c_1 u_i c_r); \ldots; 0, c_r \rangle$$

by substitutability

$\sim \langle u_i, c_1; ...; 0, c_2; ...; 0, c_i; ...; (1 - u_i), c_r\rangle$

by reducing the compound lottery

$= c_1 u_i c_r$

$\sim c_i$.

Note that this result also ensures that the obvious requirements that $u_1 = 1$ and $u_r = 0$.

We are now in a position to justify the existence of a utility function. Consider a simple lottery

$$l = \langle p_1, c_1; p_2, c_2; \ldots; p_r, c_r\rangle.$$

By continuity each prize c_i is indifferent to a reference lottery $c_1 u_i c_r$, for $i = 1, 2, ..., r$. One prize at a time, substitute $c_1 u_i c_r$ for c_i in the lottery l. Substitutability and transitivity of indifference give:

$$\begin{aligned} l &= \langle p_1, c_1; p_2, c_2; ...; p_r, c_r\rangle \\ &\sim \langle p_1, (c_1 u_1 c_r); p_2, c_2; ...; p_r, c_r\rangle \\ &\sim \langle p_1, (c_1 u_1 c_r); p_2, (c_1 u_2 c_r); ...; p_r, c_r\rangle \\ &\sim \langle p_1, (\mathrm{x}_1 u_1 c_r); p_2, (c_1 u_2 c_r); ...; p_r\, (c_1 u_r c_r)\rangle. \end{aligned}$$

Remembering that each reference lottery is a simple lottery,

$$(c_1 u_i c_r) = \langle u_i, c_1; 0, c_2; ...; 0, c_{r-1}; (1 - u_i) c_r\rangle\rangle,$$

we may reduce the compound lottery, giving:

$$\begin{aligned} l &\sim \langle (p_1 u_1 + p_2 u_2 + ... + p_r u_r), c_1; 0, c_2; \\ &\quad ...; 0, c_{r-1}; (p_1(1 - u_1) + p_2(1 - u_2) + \\ &\quad ... + p_r(1 - u_r))c_r\rangle \\ &= x_1\left(\textstyle\sum_{i=1}^{r} p_i u_i\right)x_r, \end{aligned}$$

i.e. the simple lottery l is indifferent to a reference lottery which gives a probability of $\Sigma_i p_i u_i$ to the receipt of c_1.

Similarly, if $l' = \langle p_1', c_1; p_2', c_2; ...; p_r', c_r\rangle$, $l' \sim c_1\left(\sum_{i=1}^{r} p_i' u_i\right)c_r$.

It follows from our assumptions about weak preference and monotonicity that

$$l \succeq l'$$

$$\Leftrightarrow \quad c_1\left(\textstyle\sum_{i=1}^{r} p_i u_i\right)c_r \succeq c_1\left(\textstyle\sum_{i=1}^{r} p_i' u_i\right)c_r$$

$$\Leftrightarrow \quad \left(\textstyle\sum_{i=1}^{r} p_i u_i\right) \geq \left(\textstyle\sum_{i=1}^{r} p_i' u_i\right)$$

On setting $u(c_i) = u_i$ we obtain the expected utility property.

That $u(.)$ is an ordinal value function over the set of prizes is a straightforward deduction. From continuity and monotonicity we have:

$$c_i \succeq c_j$$

$$\Leftrightarrow \quad c_1 u_i c_r \succeq c_1 u_j c_r$$

$$\Leftrightarrow \quad u_i \geq u_j$$

3.8 Subjective Expected Utility

We are now in a position to start pulling things together. Consider a decision under risk (see page 21). The problem facing the DM is that she wishes to construct a ranking of the actions which reflects her preferences between the consequences taking into account her beliefs about the unknown state. We shall approach such problems via the *subjective expected utility model* (SEU). Central to this is the separation of the modelling of the DM's beliefs and preferences by, respectively,:

- a *subjective probability distribution*, $P(.)$, which represents her belief about the unknown state of the world;
- a *utility function*, $u(.)$, which represents her preferences.

These obey the following three key properties, which together define

1. The subjective probability distribution represents her beliefs in the sense that:

   $$P(\theta) > P(\theta')$$

   if and only if, after due reflection she believes state θ to be more likely to occur than θ'.

2. The utility function represents her preferences in the sense that:

   $$u(c) > u(c')$$

   if and only if, after due reflection she strictly prefers consequence c to consequence c'.

3. The SEU model asserts that to combine her beliefs and preferences coherently in order to rank the actions the DM should form *expected utilities*:

   $$Eu[a_i] = \textstyle\sum_{j=1}^{n} u\left(c_{ij}\right)P\left(\theta_j\right)$$

   Then she should rank a_i above a_k if and only if its expected utility is higher, i.e.

   $$Eu[a_i] > Eu[a_k].$$

Earlier we discussed subjective probability and saw how her beliefs about the relative likelihood of states could be represented by probabilities. In the previous Section, we showed how a utility function could be

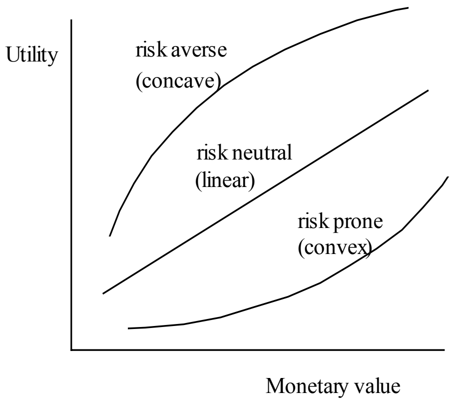


Figure 21: Utility modelling of risk attitudes

assessed to represent her preferences between lotteries. To extend the modelling of her preferences to the context of a decision table (Table 5) we need only[7] associate each action, a_i, with the lottery:

$$\langle P(\theta_1), c_{i1}; P(\theta_2), c_{i2}; \ldots ; P(\theta_n), c_{in}\rangle.$$

Finally, it may be shown that utility functions are unique up to positive affine transformations (French, 1986; French and Rios Insua, 2000).

A utility function models the DM's preferences between actions with uncertain outcomes such that it captures her risk attitude. Suppose that the actions have simple monetary outcomes: i.e. each consequence c_{ij} is a sum of money. Associated with any action a_i are two expectations: its expected monetary value $E[c]$ and its expected utility $E[a_i]$. The expected monetary value is simply the average payoff in monetary terms that results, if the DM took action a_i many, many times. However, it should be emphasised that in the following she may only take it once: there is no repetition. Related to the expected utility of an action is its *certainty equivalent*, c_c, which is the monetary value that the DM places taking a_i once: i.e. if she were offered the choice, she would be indifferent between accepting the monetary sum c_c for certain or taking a_i. Thus $u(c_c) = E[a_i]$, i.e. $c_c = u^{-1}(Eu[a_i])$. The *risk premium* of an action is $\pi = E[c] - c_c$. It is the maximum portion of the expected monetary value that she would be prepared to forfeit in order to avoid the risk associated with the action. The risk premium of an action indicates a DM's attitude to the risk inherent in *that* lottery.

A DM is *risk averse* if for any action her risk premium is non-negative. Equivalently, she is risk averse for any action if she prefers to receive a sum of money equal to its expected monetary value than to take the action itself. She is *risk prone* if for any action her risk premium is non-positive. She is *risk neutral* if for any action her risk premium is zero. Risk attitude is closely related to the shape of the utility function which represents the DM's preferences: see Figure 21. It may be shown that a concave utility function corresponds to risk aversion, a convex utility function to risk proneness, and a linear utility function to risk neutrality: see, e.g., Keeney and Raiffa (1976).

3.9 Multi-attribute Utility

In Section 3.5 we looked at the structure introduced into the form of a value function when the alternatives were modelled using multiple attributes. We noted that if an additive value function was to be a suitable representation, then the DM's preferences needed to be preferentially independent. We now ask the same question for utility functions. What conditions need to hold on the DM's preferences if we are to be able to represent her preferences with a simply structured utility function. A key concept that we shall need is that of utility independence.

To motivate utility independence, consider the following example. The prizes in four lotteries involve monetary rewards to be received now and in a year's time: (x, y) represents £x received now and £y received a year from now. The four lotteries are illustrated in Figure 22.

$l_1 = \langle ½, (100, 150); ½, (400,150)\rangle$

$l_2 = \langle ½, (175, 150); ½, (225, 150)\rangle$

$l_3 = \langle ½, (100, 250); ½, (400, 250)\rangle$

$l_4 = \langle ½, (175, 250); ½, (225, 250)\rangle$

Thus l_1 represents a 50-50 gamble giving a ½ chance of £100 this year and £150 next and a ½ chance of £400 this year and £150 next. The figure makes it clear that in the choice between l_1 and l_2 the amount received next year is guaranteed to be £150, whichever lottery is accepted and whatever happens. Similarly, in the choice between l_3 and l_4 the amount received next year is guaranteed to be £250. Moreover, if only this year's payoff is considered, it is clear that the choice between l_1 and l_2 is identical to that between l_1 and l_2. This suggests very strongly that

$$l_1 \succeq l_2 \Leftrightarrow l_3 \succeq l_4.$$

[7] Strictly, there is a little more to this than we are indicating here: see French (1986) or French and Rios Insua (2000).

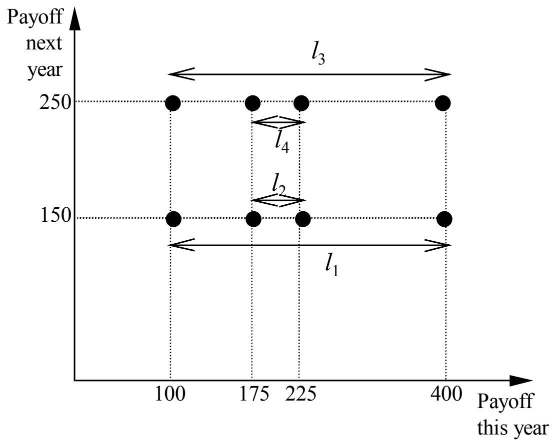


Figure 22: The four lotteries in the illustration of utility independence.

However, it should be emphasised that there is only a strong suggestion that this should be. Suppose that the payoff next year for l_3 and l_4 is increased to £250,000. Then it might be quite reasonable to hold $l_2 \succeq l_1$, since l_2 is less risky, yet to hold $l_3 \succeq l_4$, since with income of £250,000 next year the higher risk associated with l_3 is not a significant factor. Despite this reservation, the general tenor of the above argument suggests that the following independence condition might often be reasonable.

Attribute a_1 is said to be *utility independent* of attribute a_2 if preferences between lotteries with varying levels of a_1 and a common, fixed of a_2. If the DM is only concerned with two attributes and if a_1 is utility independent of a_2, it is, therefore, possible to assess a utility function for A_1 independently of a_2 . Or put another way, the DM's attitude to risk in lotteries over a_1 is independent of a_2 – and for this reason some authors use the term risk independence.

Consider now the case of q attributes and assume that the DM's preferences between consequences in conditions of certainty may be modelled by an additive multi-attribute *value* function:

$v(a_1, a_2, ..., a_q) = v_1(a_1) + v_2(a_2) + ... + v_q(a_q)$.
If, in addition, the DM holds each attribute to utility independent of all the others, then the *utility* function must have one of the following three forms:

i) $u(a_1, a_2, ..., a_q) = 1 - e^{-(v_1(a_1) + v_2(a_2) + ... + v_q(a_q))/\rho}$

ii) $u(a_1, a_2, ..., a_q) = v_1(a_1) + v_2(a_2) + ... + v_q(a_q)$

iii) $u(a_1, a_2, ..., a_q) = 1 + e^{(v_1(a_1) + v_2(a_2) + ... + v_q(a_q))/\rho}$

In both cases i) and iii), $\rho > 0$. Full details are given in Keeney and Raiffa (1976, Section 6.10).

We shall not explore multi-attribute utility theory further here, referring to the literature instead: see, e.g., French and Rios Insua (2000), and Keeney and Raiffa (1976).

4 Decision Analysis Techniques

4.1 Introduction

In this chapter we introduce some of the key techniques of decision analysis, particularly for the general and corporate strategic domains, indicating through examples several decision analytic techniques: multi-attribute modelling, expected utility, decision trees and influence diagrams.

4.2 Multi-Attribute Modelling

Consider decisions in which there are several conflicting objectives which must be balanced in order to decide upon a course of action. The decision may or may not also involve uncertainty – but we shall ignore that aspect for the present. We shall use the term *attribute* to mean one of the factors which need to be taken into account in a decision. For instance, 'cost' may be an attribute, as may be 'safety' or 'environmental impact'. Other authors sometime use the term *criterion* or, simply, *factor*. The term (*sub*)-*objective* is used by almost all writers to mean a factor which one wishes to maximise or minimise: i.e. an objective is an attribute plus a direction of preference. For instance, 'minimise cost' is an objective. In general, the most preferred point need not be at the top or bottom of an attribute scale. It may be in the middle or there may be more than one most preferred point.

An *attribute tree* or *hierarchy* essentially provides a pictorial breakdown of an overall value into the component factors of which it is comprised. There is no *objectively* right breakdown and, hence, no *objectively* right attribute tree for a problem: just one that is sufficient or *requisite* for the analysis. It is a subjective choice.

Perhaps an example will help. Lathrop and Watson (1982) consider the ways that decision analysis can help in the evaluation of risk in nuclear waste management. Their concern was not to evaluate different waste management policies in terms of all the different factors that might be considered important in deciding between them, but only in terms of their health risks. For their study, cost, non-health effects on local environment, etc. were irrelevant: only health effects mattered. As a result they decomposed their problem as shown in Figure 23. This figure may be 'read' as follows.

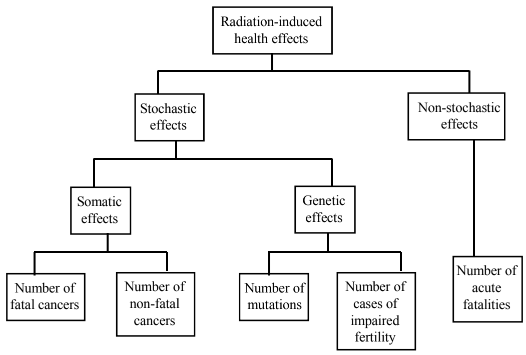


Figure 23: The attribute hierarchy used by Lathrop and Watson (1982) for nuclear waste management.

Radiation-induced health effects can be broken down into stochastic effects and non-stochastic effects. The stochastic effects may be broken down into somatic effects and genetic effects; the somatic effects, in turn, into number of fatal cancers and number of non-fatal cancers; and the genetic effects into numbers of mutations and numbers of cases of impaired fertility. Finally, the non-stochastic effects are simply taken to be the number of acute fatalities. Thus a hierarchy is formed, the lower levels of which refer to numbers of effects:

- c_1 – number of fatal cancers
- c_2 – number of non-fatal cancers
- c_3 – number of mutations
- c_4 – number of cases of impaired fertility
- c_5 – number of acute fatalities

For the purposes of Lathrop and Watson's study, these five attributes were sufficient to describe the various waste management policies in sufficient detail to evaluate their relative health effects. Each alternative can be represented by a vector or profile of five numbers, (c_1, c_2, c_3, c_4, c_5), which gives its level of achievement on each of the attributes. Here the attributes are negatively sensed: the small the number the better.

The attribute hierarchy used in an analysis does not drop out of thin air. It needs to be constructed after careful discussion between the DMs. Ideally, one would like to have procedures constructing hierarchies that reflect the views and perceptions of the DMs as they evolve in such discussion: but no simple, universally applicable procedure exists. Attribute hierarchies are constructed as much by the art of the decision analyst or facilitator as by the application of any procedure. Sometimes, the discussion proceeds 'from the top of the hierarchy to the bottom'. One can imagine a discussion between the facilitator and DMs leading to the hierarchy in Figure 5.

Facilitator: What are radiation-induced health effects?

DMs: Well, there are stochastic and non-stochastic effects.

Facilitator: What are the stochastic effects?

DMs: These break down into somatic and genetic effects.

Facilitator: And do these break down further?

DMs: Yes. Somatic into fatal and non-fatal cancers. Genetic into mutations and cases of impaired fertility.

Facilitator: What about the non-stochastic effects?

DMs: Essentially, these are simply acute fatalities.

Facilitator: Are there any other radiation-induced health effects that are significant in determining your evaluation of different nuclear waste management policies?

DMs: No. As far as radiation-induced health effects are concerned, those are the significant ones.

Of course, in reality the discussion would be longer, less directed and much more exploratory. But if the above is taken as a summary of such a discussion, then it should be clear how the hierarchy would be constructed. In this case, one says the hierarchy has been constructed by a *top-down* approach. There is also a *bottom-up* approach in which the lowest attributes are identified by 'brainstorming' and then grouped together through discussion among the DMs. In practice, as I have said, the construction of the hierarchy depends on the art of the facilitator. He or she will seldom use either the top-down or bottom-up approach entirely, but rather a mixture of both as the discussion between the DMs proceeds.

It should be realised that the attribute hierarchy used in analysis is not objective in the sense that any group of DMs would construct precisely the same hierarchy for the same problem. It is based on the perceptions of that particular group of decision makers charged with the particular decision. Having said that, it is common for different groups of DMs facing the similar problem to construct very similar hierarchies.

In a hierarchy it is usual to gather attributes into clusters (branches) that are cognitively similar. There are several requirements that attributes must meet if they are to be useful:

- all attributes must be measurable, either objectively or subjectively for each option;
- attributes should not measure the same aspect of the model to avoid double counting;
- attributes should distinguish between consequences, or else they are redundant.

4.3 Multi-Attribute Value Analysis

The simplest form of multi-attribute value analysis is very straightforward. First of all we produce a set of overall values or scores for the alternatives by:

1. Scoring each against each of the lowest level attributes.
2. Bringing each set of attribute scores to the same scale by applying weights.
3. Adding up the weighted attribute scores to give an overall score for each alternative.

$$\begin{aligned}\text{Overall value} &= \Sigma_i \text{ (weight of } i^{\text{th}} \text{ attribute)} \times \\ &\qquad \text{(score on } i^{\text{th}} \text{ attribute)} \\ &= \Sigma_i\, w_i \times x_i\end{aligned}$$

This only applies if the DMs' preferences in relation to any attribute or set of attributes is independent of performance on the other attributes. For instance, if we consider two attributes: cost and safety, it is reasonable to prefer more safety to less for the same cost, independently of that cost – and less cost to more for the same safety, independently of that level of safety. We say the attributes are (*mutually) preferentially independent* in this case. This additive form only applies when attributes are preferentially independent. It is common experience ('folklore') among decision analysts that with experience and a little luck it is almost always possible to choose the attributes in an analysis such that they are preferentially independent. In the few cases in which this is not possible, one can use more complex models.

Suppose the DM has to choose an evening's entertainment and has arrived at the possibilities and attribute tree in Figure 24.

Alternative actions:

TV	Watch TV at home.
Old_Film	Go to local cinema and see old film.
New_Film	Go to CineCentre and see new film.
Theatre	Go to Theatre and see Shakespeare
Opera	Go to Opera

Attribute hierarchy:

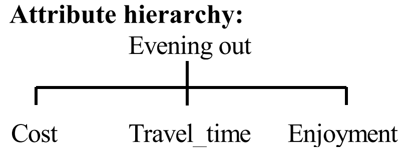


Figure 24: The 'evening out' example

Conventionally, attribute scores increase with increasing preference. Suppose the costs are:

```
TV          £0
Old_Film    £5
New_Film    £8
Theatre     £15
Opera       £25
```

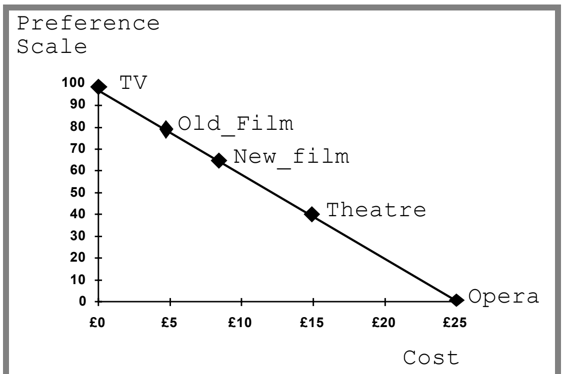


Figure 26: Value scale for cost

For small sums of money (relative to the total assets of a DM) it is usually satisfactory to take preference to be linear, here with a negative slope, as we are dealing with cost. It is also common to normalise each attribute scale so that the minimum is 0 and the maximum is 100: see Figure 26. Moreover, one only works to 2-figure accuracy. This is because human perception is not more accurate than this – arguably less accurate: maybe 5% at best.

Suppose that the travel times are:

```
TV        0  mins
Old_Film 15  mins
New_Film 25  mins
Theatre  40  mins
Opera    40  mins
```

Moreover, suppose that The DM has a strong preference not to travel at all. To me, the difference between travelling 25 mins and 40 mins doesn't seem too great; certainly not as different as the difference between 0 mins and 15 mins. Thus The DM might value travel times as in Figure 25.

To value her enjoyment of the various options, note that there is not an underlying scale to transform. The DM needs to assess her preferences directly. This can be done on a 'thermometer' scale: see Figure 27. Generally, scales for subjective attributes need to be assessed directly, whereas scales for objective attributes may be assessed directly or indirectly by a linear or non-linear transformation.

We now need to bring these attribute scores onto the same scale so they may be added up. To do this we use *swing weighting*. This takes into account *both* the importance of the attributes in determining her preference *and* the particular difference, remembering 100 points on each attribute represents the difference between the best and worst of the actual alternatives before her. See Figure 28.

Consider an imaginary alternative which scores 0 on all three scales (i.e. it is as bad as paying £25 and travelling for 40 mins to watch an old film). If the DM could improve this option up to 100 points on just one of the scales, which would the DM choose? Suppose her answer is: cost. Then each point on the cost scale is worth more than a point on the other scales. Suppose with further

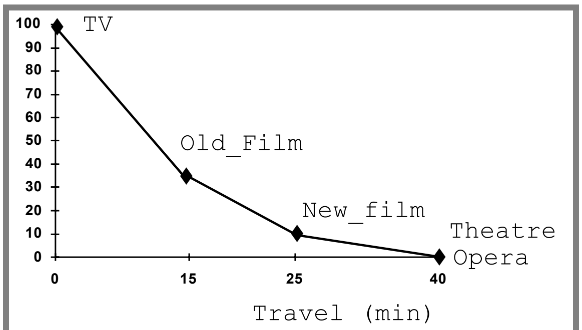


Figure 25: Value scale for travel times

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

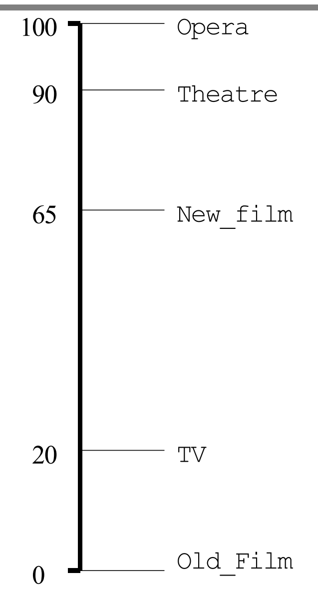


Figure 27: Value scale for enjoyment

consideration, the DM feels that 100 points on the enjoyment is worth 60 on the cost scale and 100 points on the travel time scale is worth 30 on the cost scale.

This gives overall scores:

TV:
$100\times 100 + 30\times 100 + 60\times 20 = 14200$

Old_Film:
$100\times 80 + 30\times 35 + 60\times 0 = 9050$

New_Film:
$100\times 68 + 30\times 10 + 60\times 65 = 11000$

Theatre:
$100\times 40 + 30\times 0 + 60\times 90 = 9400$

Opera:
$100\times 0 + 30\times 0 + 60\times 100 = 6000$

Normalising so that the maximum possible score is 100 (i.e. dividing by 190=100+30+60).

TV	75
Old_Film	48
New_Film	58
Theatre	49
Opera	32

Her dislike of travel and her meanness $\Rightarrow$ TV!

Some insights can be drawn from simple Pareto plots: Figure 29 illustrates one of the three possible in this example. Note that preference increases towards the top right hand corner in these plots. It is clear that in terms of Enjoyment and Cost (but ignoring Travel Time) that the Old_film option is dominated.

Sensitivity analysis is a key component of any decision analysis which utilises a quantitative model. The DM can investigate the effect of varying the weight on Cost. At present Cost contributes 100/190 = 52.6% of the weight. Thus the overall score for any alternative may be written:

$$\text{overall score} = (100/190) \times \text{Cost} + (90/190) \times (30/90 \times \text{Travel Time} + 60/90 \times \text{Enjoyment})$$

or, writing w_{Cost} for the percentage weight on Cost:

$$\text{overall score} = w_{Cost} \times \text{Cost} + (1- w_{Cost}) \times (30/90 \times \text{Travel Time} + 60/90 \times \text{Enjoyment})$$

Which gives for each alternative:

TV:
$w_{Cost}\times 100+(1-w_{Cost})\times(30/90\times 100+60/90 \times 20)$

Old_Film:
$w_{Cost}\times 80+(1- w_{Cost})\times(30/90\times 35+60/90\times 0)$

New_Film:
$w_{Cost}\times 68+(1-w_{Cost})\times(30/90\times 10+60/90\times 65)$

Theatre:

Thus we set the weights:

Cost	100
Travel Time	30
Enjoyment	60

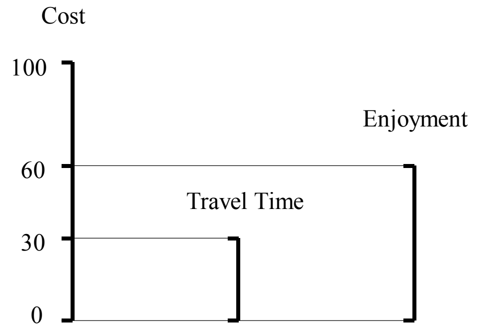


Figure 28: Swing weighting

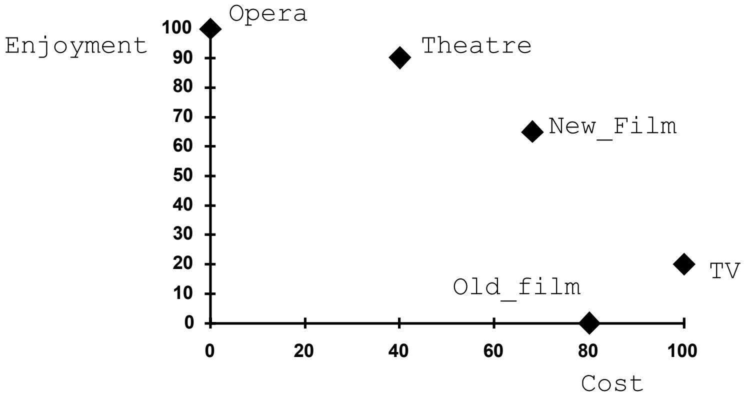


Figure 29: Pareto plot of enjoyment versus cost

w_{Cost}×40+(1- w_{Cost})×(30/90×0+60/90×90)

Opera:
w_{Cost}×0+(1- w_{Cost})×(30/90×0+60/90×100)

i.e.:

TV:	w_{Cost}×100+(1- w_{Cost})×47
Old_Film:	w_{Cost}×80+(1- w_{Cost})×12
New_Film:	w_{Cost}×68+(1- w_{Cost})×47
Theatre:	w_{Cost}×40+(1- w_{Cost})×60
Opera:	w_{Cost}×0+(1- w_{Cost})×67

This leads to the sensitivity plot shown in Figure 30.

4.4 Expected Utility Modelling

Multi-attribute value techniques provide tools for exploring trade-offs between conflicting objectives in conditions of certainty, i.e. when there is no uncertainty about the consequence of picking an alternative. How do we approach much more realistic circumstances in which the outcome of any choice has a degree of uncertainty? One of the key methodologies is subjective expected utility: see Section 3.8.

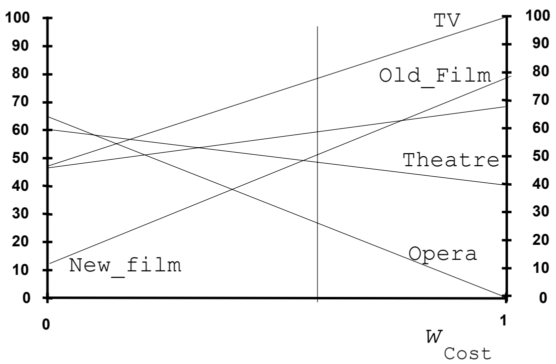


Figure 30: Sensitivity analysis on w_{cost}

A DM has £100 to invest. For simplicity, there are only three choices of investment bond open to her. Thus she has three possible actions: a_1, a_2, a_3 – buy the first, second or third bond, respectively. We shall assume that she cannot divide her money between two or more bonds each must be cashed in one year. The encashment values of two of the bonds are uncertain because of the difficulty of forecasting financial markets. Suppose that the DM is prepared to categorise the possible state of the market in a year's time into three levels of activity relative to the present: the market might fall, stay level, or rise. These are the states of the world for the problem. Suppose further that she predicts the possible consequences of her actions, i.e. the encashment values of the bonds, as being those indicated in Table 6. We shall assume that the DM employs a decision analyst (A) to help her think about the problem.

The interview between the DM and A might go as follows.

A: Consider the wheel of fortune or, as I shall call it, probability wheel shown in Figure 31(a). Which of the following bets would you prefer? I shall spin the pointed and see which sector it ends up pointing into.

Bet A: £100 if the pointer ends in the shaded area;
£0 otherwise.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

		States		
		fall θ_1	stay level θ_2	rise θ_3
Action	a_1	£110	£110	£110
	a_2	£100	£105	£115
	a_3	£90	£100	£120

Table 6: The investment problem

Bet B: £100 if the pointer ends in the unshaded area;
£0 otherwise.

DM: I wouldn't really mind. If I must choose, I'11 say *Bet A*.

A: Would you be unhappy if I made you take *Bet B*?

DM: Not at all. As far as I can see they are the same bet. Look, what is this all about? I have £100 to invest in bonds. Why are you getting me to consider silly gambles?

A: Investing in the stock exchange is a gamble, isn't it?

DM: Well, yes. ...

A: So what I am going to do to help you decide how to invest your money is to ask you to consider your preferences between a number of simple gambles. These will form a sort of reference scale against which to compare the actual choices available to you. Parts of the investment problem will be represented in each of these; other parts will be forgotten. This will mean that you may concentrate on each particular part of your investment problem in turn without being confused by the other parts. However, before I can do this I must ensure that you are thinking 'rationally' about simple gambles. I must ensure that you have reasonable conception of the probabilities generated by a probability wheel and that you do not believe, say, that the shaded area is 'lucky'. Let me go on. Just one more question about these silly gambles. Consider the probability wheels in Figure 31(b) and (c). Do you think it more likely that the spinner would end in the shaded area of (b) or of (c)?

DM: Both shaded sectors are a quarter of a circle, aren't they?

A: Yes.

DM: Then they are equally likely.

A: So you would be indifferent between a bet in which you receive £100 if the pointer ended in the shaded sector of wheel (b) and nothing otherwise, and the same bet based on wheel (c).

DM: Of course.

A: All right. Let's start looking at some parts of your investment problem. Suppose that you pay me that £100 of yours and as a result I offer you the following choice.

Bet C: At the end of the year I will spin the pointer on probability wheel (a). You will receive £90 if the pointer ends in the shaded area; £120 otherwise.

Bet D: £110 for sure (i.e. I guarantee to give you £110 at the end of the year).

Which bet would you choose?

DM: *Bet D*, the certainty of £110.

A: O.K. Now what happens if I change *Bet C* so that it is based on probability wheel (d). Thus you have the choice between:

Bet C: £90 if the pointer ends in the shaded area of wheel (d); £120 otherwise.

Bet D: £110 for sure.

DM: In this case *Bet C*, but only just.

A: The shaded sector in wheel (d) is 10% of the area of a circle. How big would it have

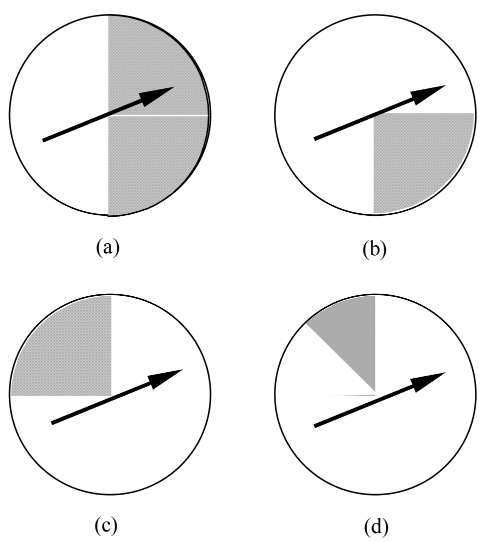


Figure 31: Some probability wheels

 15/02/2003
Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

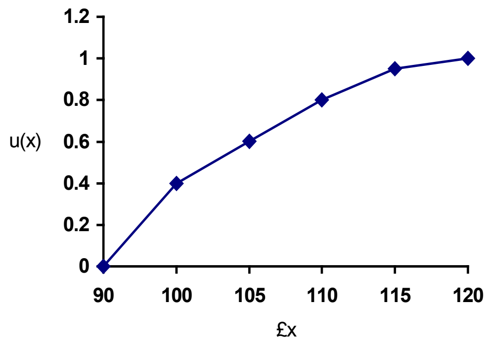


Figure 32: The DM's utility curve

to be for you to be indifferent between *Bet C* and *Bet D*? You need only give me a rough answer.

DM: About 20%. That gives me odds of 4 to 1 on winning £120, doesn't it?

A: Yes. Now we can start calculating some utilities. Let us set

$$u(£90) = 0$$
$$u(£120) = 1$$

We could choose any other numbers, as long as $u(£120) > u(£90)$. It really doesn't matter. All they do is set the unit of measurement for your preference. From your indifference between the bets we know that

$$u(£110) = 0.8 \times u(£120) + 0.2 \times u(£90)$$
i.e. $u(£110) = 0.8 \times 1 + 0.2 \times 0 = 0.8$.

DM: But I only said I was indifferent when the shaded sector was *roughly* 20% of the probability wheel. How can we say that my utility is exactly 0.8?

A: We can't. But it will serve as a working hypothesis. Later we will do a sensitivity analysis to find the significance of this assumption.

The analyst will then proceed to question the DM in the same way about her preferences between bets involving the other sums of money of interest: viz. £100, £105, £115. He would probably do this by keeping the form of *Bet C*, namely a gamble between £90 and £120, and replacing the certain reward in *Bet D* by £100, £105 and £115 in turn. Suppose that as a result of this questioning the following utilities are determined:

$$u(£90) = 0.00$$
$$u(£100) = 0.40$$
$$u(£105) = 0.60$$
$$u(£110) = 0.80$$
$$u(£115) = 0.95$$
$$u(£120) = 1.00.$$

These values are plotted in Figure 32.

Notice that the utility function is concave. This is a very common property of utility functions for monetary consequences: see notes on risk attitude. Roughly speaking, it means that the DM would be prepared to pay a small premium to reduce the uncertainty associated with each of her possible actions. Having determined the DM's utilities, the analyst would then check that these values are consistent with some of the DM's other preferences, ones that he had not elicited in the above questioning. For instance, he might ask:

A: Which of the following bets would you prefer? At the end of a year I will spin the pointer on probability wheel (b).

Bet E: £120 if the pointer ends in the unshaded area;
£110 otherwise.

Bet F: £115 for sure.

DM: I don't really mind.

A: Good. That is comforting: because, if you look at the expected utilities of the bets, you will find they are both 0.95. Thus consistency demands that you should be indifferent.

The analyst would continue questioning the DM until he was satisfied that the utility curve well represents her preferences. If an inconsistency becomes apparent, he would point it out to her, identifying the points at which her preferences were inconsistent. It would always be left to the DM to revise her preferences to resolve the inconsistency. For instance, if she had preferred *Bet E* to *Bet F* here, it would suggest that the utility function undervalues *Bet E* or overvalues *Bet F* (or both). Since $u(£120)$ takes the conventional value of 1.00, this means that the DM's earlier indifferences, which determined $u(£110)$ and $u(£115)$, are called into question. Thus she would be asked to reconsider these. Typically DMs revise their judgements when an inconsistency is pointed out to them. However, in the event that they do not, the analysis should be halted, because they are, in effect, denying the rationality of SEU analysis: it is not for them.

The next task for the analyst is to assess the DM's probabilities for the states of the market at the end of the year.

A: Consider the probability wheel shown in Figure 31(a). Do you think the event that the level of market activity falls over the next year is more, equally, or less likely than the event that the spinner ends in the shaded sector of wheel?

DM: What do you mean by 'the level of market activity falling'?

A: Exactly what we meant in the decision table: the state of the world, θ_1.

DM: Oh I see. I am sure the event on the wheel is more likely.

A: O.K. Now compare the event of the market falling, with the event that the spinner ends in the shaded sector of wheel (d). Which is more likely?

DM: The event of the market falling.

A: How big would the shaded sector have to be for you to think the events equally likely?

DM: About twice as big as that on wheel (d).

A: The shaded area in (d) is about 10% of the wheel. So you would think the events equally likely if it were about 20%?

DM: Yes. That would be odds of about 4 to 1, wouldn't it?

A: Yes. So we shall provisionally take your subjective probability $P(\theta_1)$ as 0.2.

Note that the analyst has asked the DM directly for her feelings of relative likelihood between events. If the DM felt more comfortable discussing preferences, then the analyst might have asked the DM to state her preferences between bets of the form:

Bet A.	£100	if the market falls;
	£0	otherwise;
Bet B.	£100	if the spinner stops in the shaded sector of wheel (d);
	£0	otherwise.

We shall suppose that the interview continues and that the analyst confirms that the DM's subjective probabilities are approximately:

$$P(\theta_1) = 0.2;\ P(\theta_2) = 0.4;\ P(\theta_3) = 0.4.$$

The analyst would be comforted that these summed to one, but he would not accept this alone as sufficient evidence of consistency in the DM's replies. For instance, he might ask:

A: Which do you think more likely:

Event E the market activity does not rise, i.e. it stays level or falls;

Event F the market activity changes, i.e. it rises or it falls but it does not stay level?

DM: *Event E*… I think.

A: Hm… now think carefully. *Event E* occurs if θ_1 or θ_2 happens. Event *F* occurs if θ_1 or θ_3 happens. By your earlier replies the probability of both events is 0.60. Thus they should both appear equally likely.

DM: Oh, I see what you mean. What should I do?

A: That is not for me to say really. I – or rather the theory – can tell you where you are being inconsistent. How you change your mind so that you become consistent is up to you. But perhaps I can help a bit. Both events occur if θ_1 happens. So your perception of their relative likelihood should really only depend on whether you think it more, equally, or less likely that the market stays level than that it rises. You have said that you consider these equally likely. Do you wish to reconsider your statement?

DM: No, I am happy with that assessment: θ_2 and θ_3 are equally likely.

A: Then it would appear that you should revise your belief that *Event E* is more likely than *Event F* to the belief that they are equally likely.

DM: Yes, I agree.

A: But don't worry too much about this. We will remember this conflict later when we do a sensitivity analysis.

Next the analyst would proceed to calculate the expected utilities of the three investments.

$$\begin{aligned} \mathrm{E}u[a_1] &= 0.2 \times u(£110) + 0.4 \times u(£110) + 0.4 \times u(£110) \\ &= 0.80. \end{aligned}$$

$$\begin{aligned} \mathrm{E}u[a_2] &= 0.2 \times u(£100) + 0.4 \times u(£105) + 0.4 \times u(£115) \\ &= 0.70. \end{aligned}$$

$$\begin{aligned} \mathrm{E}u[a_3] &= 0.2 \times u(£90) + 0.4 \times u(£100) + 0.4 \times u(£120) \\ &= 0.56. \end{aligned}$$

The interview might then continue.

A: So a_1 has the highest expected utility.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

DM: Which means that I should choose the first investment?

A: Well, not really. I expect that you will choose that investment, but the analysis is not yet over. Remember that your utilities were determined from replies of the form '*about* 20%'. In a sense, SEU theory assumes that you have infinite discrimination and can discern your preference however similar the bets. But you cannot, er… can you?

DM: No. I was never sure of my replies to within more than 2 or 3%.

A: So we must see how sensitive the expected utility ordering of a_1, a_2 and a_3 is to the values that we have used in the calculations. Now, if we are to consider the possibility that the expected utility of a_1 is less than the expected utility of one of the other actions, we must find a *lower* bound on the expected utility of a_1 and *upper* bounds on the other expected utilities. Let us begin by looking for a lower bound on the expected utility of a_1, i.e. on $u(£110)$. Which of the following bets would you prefer? At the end of a year I will spin the pointer on wheel (b).

Bet G: £120 if the pointer ends in the unshaded sector;
£90 otherwise.

Bet H: £110 for sure.

DM: Definitely, *Bet H*.

A: So we know that

$$u(£110) > 0.75 \times u(£120) + 0.25 \times u(£90),$$

i.e. $u(£110) > 0.75$.

DM: You have assumed that $u(£120)$ and $u(£90)$ are exactly 1.0 and 0.0 respectively. Shouldn't we check those values?

A: No. We can *set* the scale and origin of a utility function arbitrarily. If we varied those values it would be precisely like changing from measuring temperature in Centigrade to Fahrenheit. The numbers would be different, but they would place hot and cold objects in the same order. Here the utilities would be different, but the resulting order of your actions would be the same.

Suppose that the analyst questions the DM further and determines, similarly, bounds on the other utilities:

$$u(£110) < 0.45,$$
$$u(£105) < 0.64,$$
$$u(£115) < 0.96.$$

Then it follows that, if for the present the analyst takes the DM's probabilities as fixed:

$$\begin{aligned}
\mathrm{E}u[a_1] &= u(£110) \\
&> 0.75; \\
\mathrm{E}u[a_2] &= 0.2 \times u(£100) + 0.4 \times u(£105) + 0.4 \times u(£115) \\
&< 0.2 \times 0.45 + 0.4 \times 0.64 + 0.4 \times 0.96 \\
&= 0.73; \\
\mathrm{E}u[a_3] &= 0.2 \times u(£90) + 0.4 \times u(£100) + 0.4 \times u(£120) \\
&< 0.2 \times 0.0 + 0.4 \times 0.45 + 0.45 + 0.4 \times 1.0 \\
&= 0.58.
\end{aligned}$$

Thus the expected utility of a_1 has a *lower* bound of 0.75, which is greater than the *upper* bounds on both the expected utilities of a_2 and a_3. The DM should prefer a_1 to both a_2 and a_3, whatever the numerical values of the utilities within the ranges of acceptable to her.

DM: So I should pick a_1.

A: Yes. It would seem to be your most preferred investment. At least, it does if you believe that your values for the probabilities truly reflect your judgements of the likelihood of events. But we had better check that.

Just as the analyst conducted a sensitivity analysis on the DM's utilities so he must conduct one on her probabilities. Indeed, he should consider variations in *both* her utilities and probabilities simultaneously. He might continue as follows.

A: Remember that we discovered that you were slightly unsure whether θ_2 and θ_3 were equally likely. Also remember that your subjective probabilities and utilities were determined from consideration of a probability wheel in which the shaded area was, say, *about* 20%. We must see how sensitive the ordering of the expected utilities is to the values that we have used in the calculations.

For this let

$$P(\theta_2) = p,$$
$$P(\theta_3) = q.$$

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

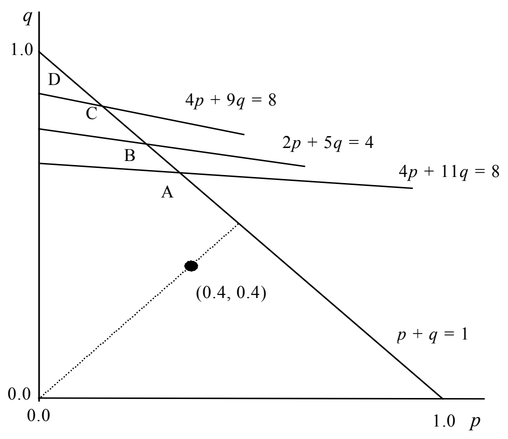


Figure 33: Plot of the permissible region for p and q.

Then, since probabilities sum to one,

$$P(\theta_1) = 1 - p - q.$$

DM: Why have you set $P(\theta_2)$ and $P(\theta_3)$ to be p and q respectively? Why not $P(\theta_1)$ and $P(\theta_2)$?

A: You have said that θ_2 and θ_3 are equally likely, or, at least, very nearly so. Thus we shall be interested in cases where $P(\theta_2) = P(\theta_3)$ and such cases will be easy to see, given the assignments that we have made. You'll see.

We shall leave your utilities as they were determined for the time being, and we shall use $Eu[a_i]$ for the expected utility of a_i ($i = 1, 2, 3$). Then

$$\begin{aligned} & \mathrm{E}u[a_1] > \mathrm{E}u[a_2] \\ \Leftrightarrow\quad & 0.8 > (1 - p - q) \times 0.40 + p \times 0.60 + q \times 0.95 \\ & 8 > 4p + 11q. \end{aligned}$$

Similarly,

$$\begin{aligned} & \mathrm{E}u[a_1] > \mathrm{E}u[a_3] \\ \Leftrightarrow\quad & 0.8 > (1 - p - q) \times 0.0 + p \times 0.4 + q \times 1.0 \\ \Leftrightarrow\quad & 4 > 2p + 5q. \end{aligned}$$

And

$$\begin{aligned} & \mathrm{E}u[a_2] > \mathrm{E}u[a_3] \\ \Leftrightarrow\quad & (1 - p - q) \times 0.40 + p \times 0.60 + q \times 0.95 \\ & > (1 - p - q) \times 0.0 + p \times 0.4 + q \times 1.0 \\ \Leftrightarrow\quad & 8 > 4p + 9q. \end{aligned}$$

Now let us plot these results. In Figure 33 we have plotted the permissible region for p and q. They are probabilities, so,

$$p \geq 0,\ q \geq 0 \text{ and } p + q \leq 1.$$

Hence (p, q) must lie in the triangle running from (0, 0) to (1, 0) to (0, 1) and back to (0, 0). Consider the line $4p + 11q = 8$. Above this, $4p + 11q > 8$, so $\mathrm{E}u[a_2] > \mathrm{E}u[a_1]$. Below it, $4p + 11q < 8$, so $\mathrm{E}u[a_2] < \mathrm{E}u[a_1]$. Similar remarks apply to the regions above and below the other lines. Hence the triangle defining the permissible region for (p, q) is divided into four subregions, A, B, C and D, such that:

in A $\quad \mathrm{E}u[a_1] > \mathrm{E}u[a_2] > \mathrm{E}u[a_3]$;
in B $\quad \mathrm{E}u[a_2] > \mathrm{E}u[a_1] > \mathrm{E}u[a_3]$;
in C $\quad \mathrm{E}u[a_2] > \mathrm{E}u[a_3] > \mathrm{E}u[a_1]$;
in D $\quad \mathrm{E}u[a_3] > \mathrm{E}u[a_2] > \mathrm{E}u[a_1]$.

In the analysis so far we have modelled your beliefs with probabilities $p = 0.4$ and $q = 0.4$. This is marked by the point • in Figure 33. Notice that it lies well within region A. Thus investment a_1 does seem to be your best choice, as we have found already. To confirm this we must check that, if slightly different, but still reasonable, values of p and q were used, then • would still lie in A. Also we must check that the upper boundary of A does not move down below • if slightly different values are used for your utilities.

DM: Shouldn't we check what happens if the other lines dividing B from C from D are moved slightly?

A: Theoretically, yes; but practically it is unnecessary. Those lines would have to move a lot further than the upper boundary of A. If that last move seems unreasonable, then so surely will the former.

O.K. let's do the sensitivity analysis. Remember that, when we checked your beliefs for consistency, you initially thought event E was more likely than F.

DM: I did change my mind, on reflection.

A: True. But, as we can see from the figure, if you changed your mind back again, it would not affect your choice of action. The dotted lines through • and the origin is the line $p = q$. On it, points represent your belief that the market activity rising is equally likely as it staying level. Below it you think it more likely that the market will stay level. In other words, had you maintained your belief that E were more

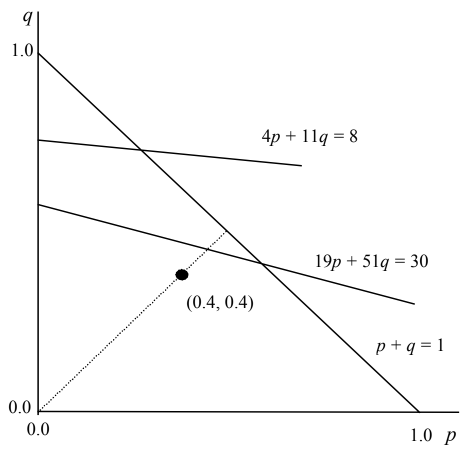


Figure 34: Effect of lowering the upper boundary of region *A*.

likely than *F*, you would have moved further into the region *A*.

DM: So it seems that I should choose a_1.

A: Wait a minute! We must consider sensitivity to your utilities again. When we did that before, we assumed that the probabilities were exact. We have considered sensitivity to utilities and probabilities independently of the other. Now we must consider sensitivity to both simultaneously remember that we obtained the bounds:

$$u(£100) < 0.45,$$
$$u(£105) < 0.64,$$
$$u(£115) < 0.96,$$
$$u(£110) > 0.75.$$

With these bounds we know for certain that

$$Eu[a_1] > Eu[a_2]$$
$$\text{if} \quad 0.75 > (1 - p - q) \times 0.45 + p \times 0.64 + q \times 0.96,$$
$$\textit{viz.} \quad 30 > 19p + 51q.$$

So let us plot $30 = 19p + 51q$ in the diagram. See Figure 34. The point • still lies in region *A* after the upper boundary has been lowered from $4p + 11q = 8$ to $10p + 51q = 30$.

It only remains to see whether you might be prepared to increase your subjective probabilities p and q above the line $19p + 51q = 30$. Are you still content that the possibilities of the market staying level and of it rising are equally likely?

DM: Yes.

A: Then we only need consider movement of the point • along the line $p = q$. Now $(1 - p - q)$ is your probability for the market falling. You have said that this is 0.20. Would you be prepared to change this?

DM: I still think 0.20 is about right. I suppose it might be an underestimate.

A: Well, if $(1 - p - q)$ increases, the point • moves down $p = q$ further into the region A. So it does seem that, however we model your beliefs and preferences, the investment a_1 comes out with the highest expected utility.

DM: So I should choose a_1, the first investment.

A: If you want your beliefs and preferences to be consistent with principles of rational behaviour assumed by SEU theory: yes. But really you should not ask me or the theory to tell what to do. Rather I would have hoped that the above analysis helped you think more clearly about your problem and brought you understanding. Now in the light of that understanding, you must choose for yourself.

DM : I suppose that you are right. I had always favoured investment but I was afraid that I did so because it was completely without risk. Now I can see that I do not believe that the likelihood of a favourable market is high enough to be worth taking the risk involved in a_2 and a_3. Before, I could not see how to weigh up uncertainties.

4.5 Decision Trees and Influence Diagrams

Decision tables such as Table 6 provide a very limited way of representing decision problems under risk. It is hard, but actually not impossible, to represent contingencies in potential strategies: e.g., if this happens, I will take action A; if that happens I will take action B. Representing dependencies between events is also difficult. For this reason, two other formats for the decision model are more commonly used: decision trees and influence diagrams. To introduce these consider the following example.

An airline has been offered the chance of buying a second-hand airliner. Categorising very broadly, such an aircraft may turn out to be very reliable, moderately reliable or very unreliable. A very reliable aircraft will make high operating profits and satisfy customers. A

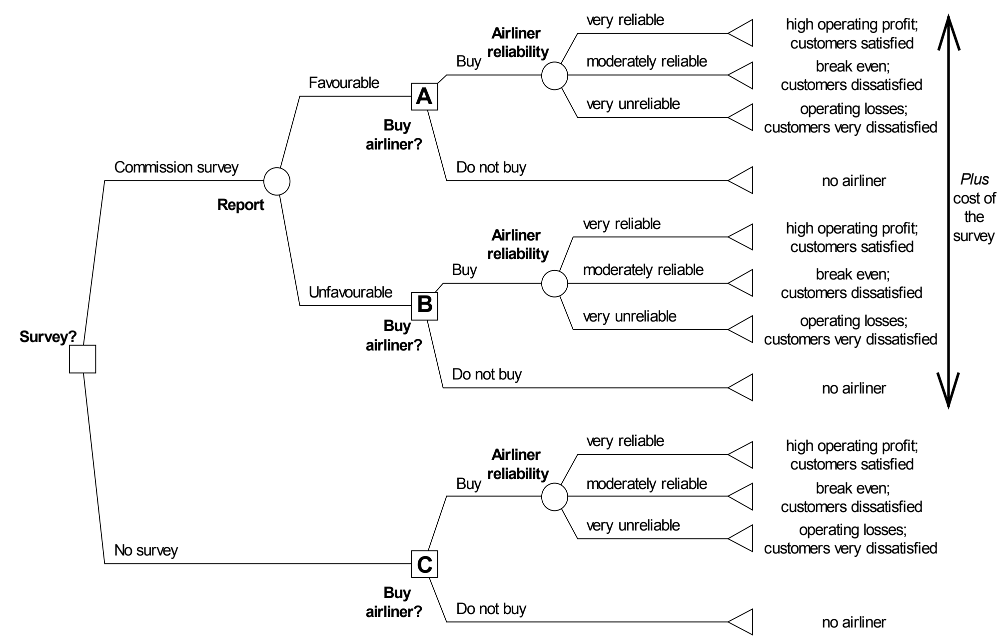


Figure 35: The airliner purchasing problem.

moderately reliable aircraft would break even on operating costs, but would lead to some dissatisfied customers. An unreliable aircraft would cost the company dear, both in terms of operating costs and in terms of customer dissatisfaction. Before making their decision the company may, if they wish, commission a firm of aeronautical engineers to survey the airliner. Of course, the airline will have to pay for the survey. Moreover, the engineers will not make explicit predictions about the aircraft's reliability. They will only couch their report in favourable or unfavourable terms. The airline must draw its own inferences about future reliability.

The problem represented as a decision tree is displayed in Figure 35. The first issue facing the airline is the decision of whether or not to commission a survey. This is represented by the square to the left of the figure. The upper branch corresponds to the decision to commission a survey and continuing across to the right, this branch divided according to the possible outcomes of the survey at the chance point representing the report. The survey may be favourable or unfavourable and in either case the airline has to decide then whether to buy the airliner. It would be wise to remark, perhaps, that there are decisions to be made at points A and B. While it is usually true that the airline should buy the airliner after a favourable report and should not after an unfavourable one, it is not always so. It depends upon the specific prior beliefs of the airline, their perception of the competence of the aeronautical engineers and their valuation of the possible consequences. We shall see how these points enter the analysis below.

The endpoints describe the consequence that accrues to the airline in the case of each set of decision choices and contingencies.

Note that, despite first appearances the decision problems at points A, B and C are not identical. Certainly at each of these points the airline must decide whether to buy the aircraft, but the information that it has to support its decision is different in each case. At A, they know that the aeronautical engineers have reported favourably on the plane; at B the report is known to be unfavourable; and at C they have no report. Thus the airline's beliefs about the aircraft's reliability will differ at each point.

Suppose that at the outset, the airline assess the reliability of the aircraft as:

P(very reliable) = 0.2
P(moderately reliable) = 0.3
P(very unreliable) = 0.5

They would assess these probabilities on the basis of their knowledge of the average reliability of airliners of the same class as the one they are considering, moderated by their knowledge of the particular aircraft's history and ownership.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

Next they need to consider how their beliefs would change in the light of the information they may receive from the aeronautical engineers. They could simply assess the probabilities:

P(very reliable|favourable report)
P(moderately reliable|favourable report)
P(very unreliable|favourable report)

and a similar set of three probabilities conditional on the receipt of an unfavourable report.

However, although these are the probabilities that they need to consider what to do at decision points A and B, they are not straightforward to assess. There is much evidence in the behavioural decision literature that DM's have difficulty in assessing the effect of evidence on their beliefs. It is better to help them construct these probabilities from a coherent set of probabilities all based upon information available at the *same* time.

The initial or *prior* probabilities were assessed before any report was received from the aeronautical engineers: indeed, before a decision whether or not to consult the engineers had been made. At the same time they will have some knowledge of the engineers – one doesn't consider taking advice of this kind without some background knowledge of the engineers' track record. Thus the directors of the airline may ask themselves, how likely is it that the report will be favourable *if* the airliner is very reliable. Ideally the answer should be 100%, but no firm of engineers is infallible. Thus assume that they assess:

P (favourable report|very reliable) = 0.9
P(unfavourable report|very reliable) = 0.1

along with the two further pairs conditional respectively on the plane being moderately reliable and very unreliable. Their assessments are given in Table 7.

Now Bayes' Theorem (see Appendix) allows the calculation of the probabilities that are really needed in the analysis: e.g.

$$P(\text{very reliable}|\text{favourable report}) = \frac{P(\text{favourable report}|\text{very reliable}) \times P(\text{very reliable})}{P(\text{favourable report})}$$

where[8]

Probability that report is:	Conditional on the airliner being very reliable	moderately reliable	very unreliable
favourable	0.9	0.6	0.1
unfavourable	0.1	0.4	0.9

Table 7: Assessed probabilities of the tone of the report given the airliner's actual reliability

P(favourable report) =
P(favourable report|very reliable) ×
P(very reliable) +
P(favourable report|moderately reliable) ×
P(moderately reliable) +
P(favourable report|very unreliable) ×
P(very unreliable).

Thus

P(very reliable|favourable report)

$$= \frac{0.9\times0.2}{0.9\times0.2 + 0.6\times0.3 + 0.1\times0.5} = \frac{0.18}{0.41} = 0.439$$

Similarly Bayes Theorem gives:

P(moderately reliable|favourable report)

$$= \frac{0.6\times0.3}{0.9\times0.2 + 0.6\times0.3 + 0.1\times0.5} = \frac{0.18}{0.41} = 0.439$$

P(very unreliable|favourable report)

$$= \frac{0.1\times0.5}{0.9\times0.2 + 0.6\times0.3 + 0.1\times0.5} = \frac{0.05}{0.41} = 0.122$$

Note that these numerical calculations can be streamlined considerably. The same denominator appears in all three cases, and is *P*(favourable report) = 0.41. Moreover, the three component products in the denominator form in turn each of the numerators.

Looking at the conditioning event that the report is unfavourable and applying Bayes' Theorem again:

P(very reliable|unfavourable report)

$$= \frac{0.1\times0.2}{0.1\times0.2 + 0.4\times0.3 + 0.9\times0.5} = 0.034$$

[8] Note that {very reliable},{moderately reliable} and {very unreliable} form a partition of the certain event in the terminology used in the formal statement of Bayes' Theorem.

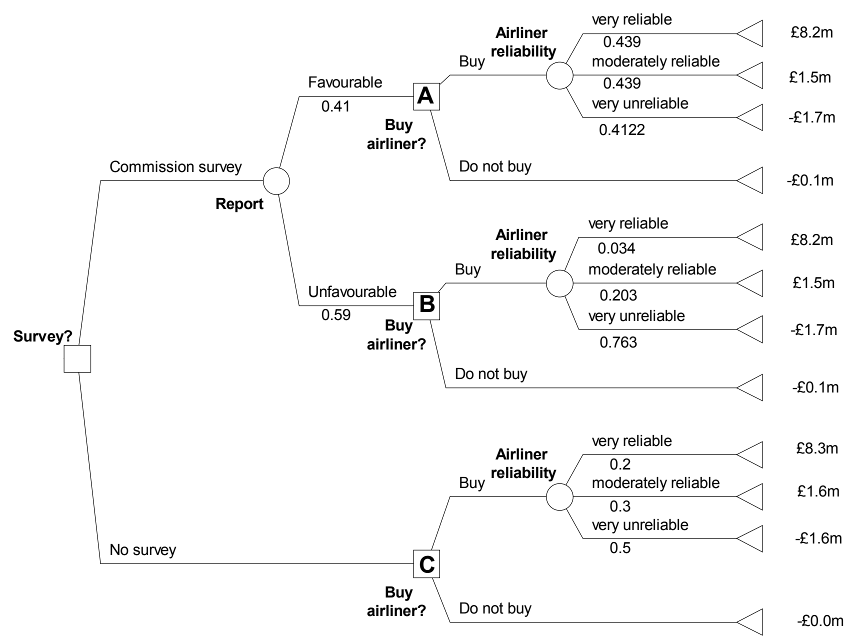


Figure 36: The decision tree for the airliner example with the probabilities and NPVs of the outcomes attached.

P(moderately reliable|unfavourable report)

$$= \frac{0.4\times0.3}{0.1\times0.2 + 0.4\times0.3 + 0.9\times0.5}$$
$$= 0.203$$

P(very unreliable|unfavourable report)

$$= \frac{0.9\times0.5}{0.1\times0.2 + 0.4\times0.3 + 0.9\times0.5}$$
$$= 0.763$$

In this case the common denominator is P(unfavourable report) = 0.59.

We have now calculated all the probabilities that we need at the chance events: see Figure 36. Note how the information in the case of a favourable report shifts the mass of the probabilities towards very reliable, whereas in the case of an unfavourable report, the shift is towards very unreliable: things are making sense!

However, we are only half-way in determining all the inputs we need for the analysis. Initially, let us suppose that the airline simply wishes to think in financial terms. We shall assume that the utility of a consequence is purely its monetary value. Assume that the net present value (NPV) over the next ten years of running a very reliable airliner, having allowed for financing of the purchase is £8.3 million. Suppose that the NPV of a operating a moderately reliable airliner is £1.6 million and that the NPV of the losses of operating a very unreliable airliner is -£1.6 million. Finally suppose that the survey would cost the airline £100 000 (= £0.1 million). These values have been attached to the endpoints in Figure 36.

Consider first the decision at A. If the airline buy the airliner they face:

Expected NPV of buying at A

$$\begin{aligned} &= £0.439\times8.2 + 0.439\times1.5 \\ &\qquad + 0.122\times(-1.7)\text{m} \\ &= £4.05\text{m} \\ &> -£0.1\text{m} \\ &= \text{Expected NPV of not buying at A.} \end{aligned}$$

So it makes sense for the airline to buy the aircraft *if* they commission a report and it is favourable. Similarly, consider first the decision at B. If the airline buys the airliner they face:

Expected NPV of buying at B

$$\begin{aligned} &= £0.034\times8.2 + 0.203\times1.5 \\ &\qquad + 0.763\times(-1.7)\text{m} \\ &= -£0.71\text{m} \\ &< -£0.1\text{m} \\ &= \text{Expected NPV of not buying at B.} \end{aligned}$$

So it makes sense for the airline not to buy the aircraft *if* they commission a report and it is unfavourable. Finally consider the decision at C. If the airline buy the airliner they face:

Expected NPV of buying at C

$$\begin{aligned} &= £0.2\times8.3 + 0.3\times1.6 + 0.5\times(-1.6)\text{m} \\ &= £1.34\text{m} \end{aligned}$$

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

> £0.0m
= Expected NPV of not buying at B.

So if they do not commission a survey, the balance would seem to be in favour of buying the aircraft. Is it worth commissioning the survey? Note that we now know what the airline should do at decision points A and B. Thus we know the expected NPV's at these: *viz*.

Expected NPV at A
= £ max{4.05, -0.1}m = £4.05m

Expected NPV at B
= £ max{-0.71, -0.1}m = -£0.1m

It follows that if a survey is commissioned:

Expected NPV of commissioning a survey
= £0.41×4.05 + 0.59×(-0.1)m
= £1.60m

We know the expected NPV if a survey is not commissioned: it is simply the expected NPV of buying at C:

Expected NPV of not commissioning a survey
= £ max{1.34, 0.0}m = £1.34m

We can now see that the airline should commission a survey because the expected NPV of doing so is greater than that of not doing so. The analysis suggests that their optimal strategy is: commission a survey; if the report is favourable, buy the airliner; if not, do not buy it.

Note that the analysis proceeded in reverse chronological order. Later decisions are analysed first, because they determine the consequences of earlier decisions. This procedure is known as *rollback* or *backward dynamic programming*. Thus the analysis is simple:

1. take expectations at chance nodes
2. optimise at decision nodes: i.e. minimise in problems concerning costs and maximise in those concerning profits.
3. calculate from right to left (*rollback*).

The analysis may be continued to investigate the value to the airline of the aeronautical engineers' report.

The survey will cost the airline £100 000, but the analysis above shows it is well worth paying this. How much would it be worth paying?

The expected NPV of commissioning a survey is £1.60m including the cost of the survey. The expected NPV of not commissioning a survey is £1.34m. Had the survey cost £(1.60-1.34)m = £0.26m more, then the expected NPV for both would have been the same. In other words, the most it is worth paying for a survey is £(0.26+0.1)m = £0.36m. The value of the information derived from the survey is £360 000. At least it is if the decision is to be evaluated in terms of expected NPV.

A decision tree displays the different contingencies in a decision well, but does not provide a clear picture of the interrelation and influences between the uncertainties and decisions. Thus Figure 35 shows the airline that their first decision is whether to commission a survey. Then in the light of the outcome of the survey, they must decide whether to buy and, only if they do, will they discover the plane's reliability. Laying out the chronology of a decision can be very useful: indeed, it may be enough to allow the DMs to see their way through the problem without further analysis (Wells, 1982). However, the probabilistic dependence of the nature of the survey report on the reliability of the airliner is implicit rather than explicit in the tree. An influence diagram is an alternative representation of decision problems, which focuses on dependencies, but, in doing so, loses explicit representation of chronological relationships and contingencies. Again squares are used to indicate decisions and circles or ovals used to indicate uncertainties. However, the arrows do not indicate a flow of time from left to right and the range of possibilities that might result from either a decision or by 'chance'. Rather the arrows indicate dependences which are reflected by the way the DM looks at the problem.

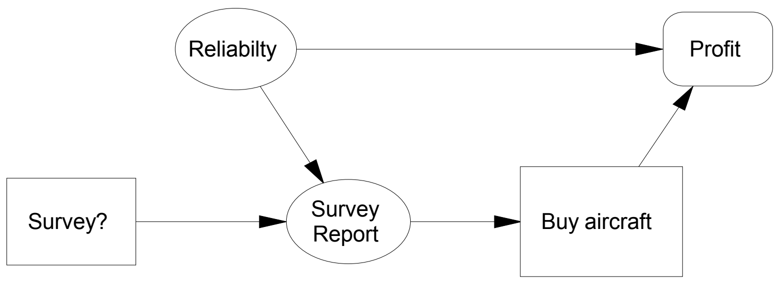


Figure 37: Influence diagram representation of the airliner purchasing problem

Figure 37 shows an influence diagram representation of the problem. This shows that the survey report depends on the aircraft's reliability and on the decision to commission a survey. The decision to buy the aircraft will be influenced by any survey report; and the profit arising from the decision depends upon both the reliability and on the decision to buy the airliner. So far the interpretation is straightforward: but there are subtleties.

Influence diagrams do not show temporal relationships unambiguously. From the tree it is clear that the aircraft's reliability is only discovered after the plane is bought. This is far from clear in the influence diagram. The influence arc from reliability to the survey indicates that the plane's reliability influences the survey report; *not* that it is know before the report is written. In influencing the profit, however, there is a necessity that the airline observe the actual reliability. There is ambiguity. Some authors resolve this ambiguity by using dotted and solid (or different coloured arrow to represent differing temporal relationships. Others, and we prefer this approach, run decision tree and influence diagram representations in parallel, each providing a complementary perspective on the decision problem.

Decision trees have a disadvantage in that for many problems they rapidly become very large: too large for the eye to comprehend as one. As such they have been described as a "bushy mess". Thus decision trees are often displayed as a series of sub-trees.

Influence diagrams are a much more compact representation. However, their advantage in this respect is in a sense illusory. Decision trees can represent asymmetric decision problems, i.e. problems in which a particular choice of action at a decision node makes available different choices of action at subsequent decision nodes to those available after an alternative choice. The airline problem is asymmetric. Such asymmetric problems are the rule rather than the exception in decision analysis. It is not possible to

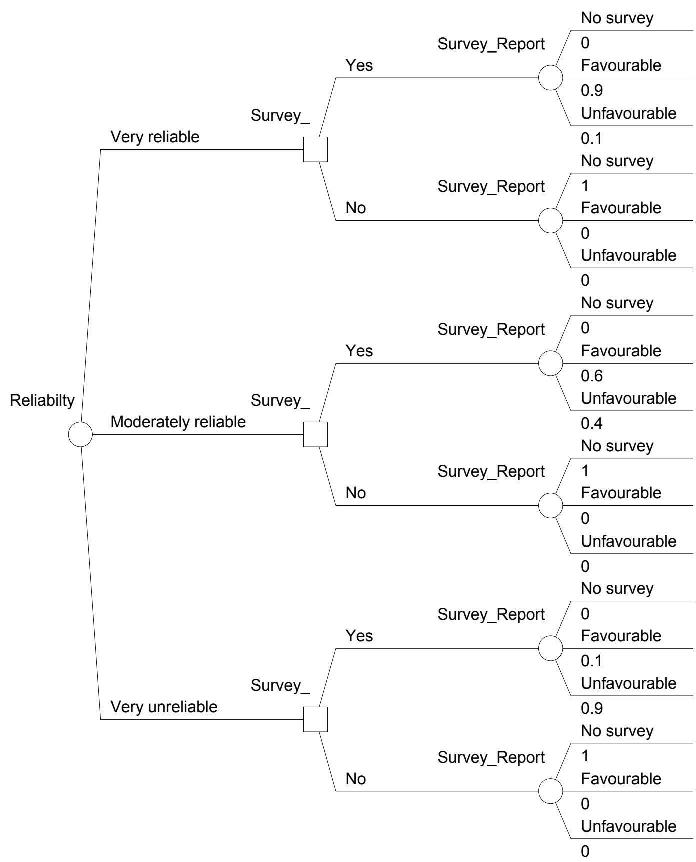


Figure 38: The introduction of an artificial outcome of no_report in order to create a symmetric tree.

represent such asymmetry within influence diagrams directly. Figure 37 hides the fact that 'mathematical tricks' have been employed to construct the influence diagram. Suppose a decision not to commission a survey is made. Then there is no survey report despite the arc from survey? to survey report in the influence diagram. To overcome this an artificial observation of 'no_survey' is introduced. This has probabilities of either zero or one depending on whether the survey is commissioned: see Figure 38.

Decision trees and influence diagrams provide complementary perspectives on a problem. A decision tree emphasises temporal contingencies between actions and possible events, whereas an influence diagram emphasises relationships between knowledge and beliefs, showing dependencies and independencies between beliefs. Decision trees have a disadvantage in that they can soon become so 'bushy' that comprehending them becomes very difficult. Splitting the tree into sub trees can mitigate this, but the difficulty remains. Influence diagrams, on the other hand, are more compact and can represent larger problems without challenging the DMs comprehension so much. However, they cannot easily represent *asymmetric decision problems*, in which a particular choice of action makes available different choices of action at subsequent decision nodes: see the airliner problem for an example. Such asymmetric problems are unfortunately the rule rather than the exception in practice.

There are many algorithms and methods available to solve problems represented either as decision trees or influence diagrams. We illustrated one of the older methods for decision trees in the airliner example, however, in general, we refer the reader to the literature for descriptions. In practice, the models are built and solved by using the very powerful software now available so the need to be able to perform the calculations for oneself is less great today than in the past. For descriptions of the algorithms see, e.g., Clemen (1996), Jensen (2001), Marshall and Oliver (1995), Oliver and Smith (1990), Raiffa (1968) and Smith (1988).

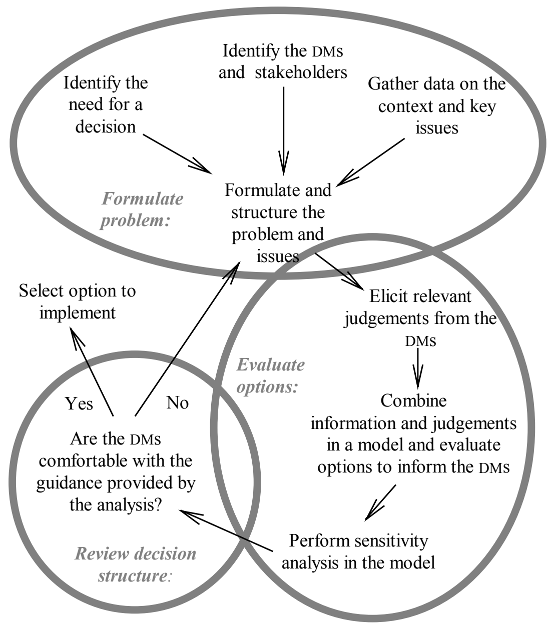


Figure 39: Overview of the Decision Process

5 Methodology of Decision Analysis

5.1 An Outline of the Decision Analytic Process

Firstly, we emphasise that the view of decision analysis presented here adopts a prescriptive viewpoint in which requisite modelling techniques are used to guide the DMs' evolving judgements and understanding. Secondly while we lean strongly to Bayesian approaches to decision analysis, the broad process described below can be used with a much wider range of analytic approaches.

In Figure 39 we present an outline of the decision analytic cycle. It is intended to show that any decision analysis cycles round three phases:

- problem formulation;
- evaluation of options;
- review of the decision models.

Each phase involves many sub-activities, the main ones of which are shown in the figure. The analysis will seldom be purely cyclic. Rather it will move backwards and forwards between phases with the predominant direction being anti-clockwise, but with many short reversals. The process is complete when the DMs are comfortable with the conclusion of the analysis; i.e. when they feel the analysis is *requisite*.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

We do not claim that our representation of the decision analysis cycle is unique: all authors have their own representation and their own words; but we do believe that it captures the essence of the process.

Problem formulation

Decision problems seldom arrive, fully formulated with a clear set of options and objectives. More often, DMs become aware of a set of issues which need addressing. Even the issues seldom arrive together, but an awareness among the DMs and their agents gradually builds-up over a period of time. Eventually, however, that awareness builds to such a level those responsible ask the question: "do we need to take some sort decision here?" At this point, the decision analysis begins.

Before any quantitative analysis can begin – indeed, before the problem can be formulated – there is a need to gather data and other material to turn the growing awareness that issues need to be faced into something more concrete. Lewis (1991) describes this phase as an *intelligence activity*, in which the DMs survey "the economic, technical, political and social environment to identify new conditions which call for new actions". In the event that such conditions are identified, the DMs become aware of the need for a decision and, again in Lewis's words, "invent, design and develop possible courses of action for handling (the) situation". In Janis and Mann's (1977) conception of decision making, this represents a move from *unconflicted adherence* to the current strategy to what they term *vigilant decision making* in which current strategy is compared with other alternatives. Indeed, requisite modelling can be thought of as a means of supporting progression from unconflicted adherence through unconflicted change to vigilant decision making: i.e. of supporting escalation through the three 'rational' patterns of coping identified by Janis and Mann. See Figure 40. In this sense, unconflicted adherence and unconflicted change are the first two cycles of requisite decision making.

The Bayesian approach, see Figure 6, being based upon SEU models, decomposes the model of a problem into two parts (French and Smith, 1997):

- a modelling of beliefs and uncertainties – i.e. the 'scientific' knowledge relevant to the analysis;
- a modelling of preferences and values – the value judgements of the players.

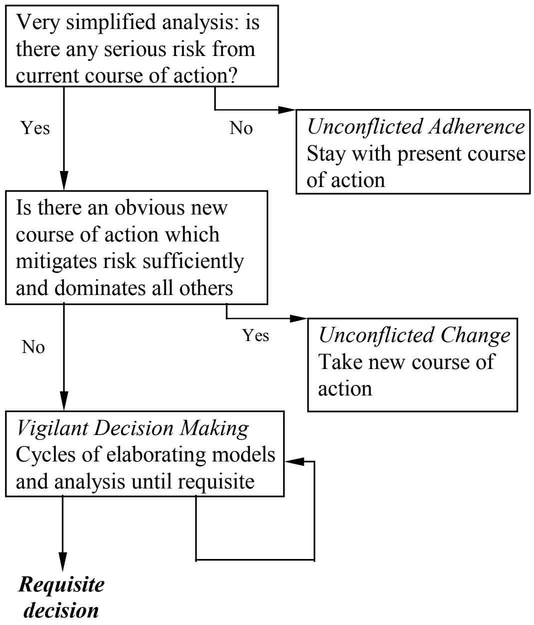


Figure 40: An indication of the progression through unconflicted adherence, unconflicted change and vigilance supported in the cycling of requisite modelling

Experts contribute the Science and knowledge on which forecasts of outcomes are based: the left hand side of Figure 6. Stakeholder values provide input to the variety of factors which help form the DMs value judgements and which are modelled by multi-attribute value or utility models: the right hand side of Figure 6. Analysts oversee the whole process, without, in principle, contributing contextual knowledge; while the DMs are responsible for taking the understanding provided by the process and interpreting it into action.

Note also that often the left or right hand sides of the process have very unequal emphases. Thus some analyses may concentrate on exploring value judgements with little or no analysis of uncertainty. In such cases, the Science provides *deterministic* descriptions of the possible consequences of the alternatives. In other analyses, exploration of the uncertainty may dominate with little modelling of complex preferences: perhaps relying purely on expected monetary values based upon fairly simple cost models.

Remember the distinction between data and information. The very early stages of a decision analysis are perhaps much more data gathering than information gathering, since until some analysis has been undertaken it is not clear what data are relevant. Information comes into being during an analysis and ceases to have value at the end of the analysis (except in terms of reporting and justifying the analysis).

Methods of information gathering will vary from context to context. Experts will be interviewed and maybe asked to complete questionnaires or fact sheets. In some cases the interactions with experts might be very formalised; there are many methodologies for this (see, e.g., Cooke, 1991; French and Rios Insua, 2000, Chapter 4). There is likely to be a variety of brainstorming sessions during which the bounds of the problem will be defined. Some such sessions may be formal meetings with a facilitator structuring the discussion, but the informal thinking should not be discounted: discussions over cups of coffee, reading reports of related projects and quiet periods doodling have a part to play.

Two points should be noted. First, in the very first cycle of a decision analysis (Figure 39), the data and information gathering may be very superficial. Only in later cycles when models have some substantive structure and much thinking will have gone into their construction will the questions be sufficiently clear that they may be answered with some assurance. Second, Keeney (1992) has argued that that one should adopt *value focused thinking*: namely, "first deciding on what you want and then figuring out how to get it". Thus an early stage in any decision analysis – long before one starts to identify alternatives – should be to explore the DMs' values. We do not pretend that value focused thinking is easy to adopt. There are times when one is presented with a set of alternatives and one simply has to choose. In such circumstances the analysis is necessarily led into alternatives first mode, although we would argue that the question "what are we trying to achieve?" should be considered as soon as possible.

Problem formulation is an art rather than a science; and, whilst there are guidelines that can help construct workman-like analyses, the best and most professional analysts draw upon a mix of experience and flair. Generally, the area of problem formulation is well developed in the UK, where it derives from our strength in the field of *soft OR* (Rosenhead and Mingers, 2001). These methods help one develop perspectives on the issues, identify objectives and values, key uncertainties and possible actions. We discuss these methods in Section 5.2 below and indicate how they help the process of structuring quantitative models such as multi-attribute value models, decision trees, or influence diagrams.

Evaluate alternatives

Once the structure of the model[9] has been developed, the next step is to populate it with numbers. Some will come from data gathered earlier; others will need to be elicited from experts, stakeholders or, of course, the DMs themselves.

Once the models have been developed and numerical inputs determined, the analyst can 'run the analyses' which today usually involves little more than clicking a 'button'. We emphasise the value of exploring a *series* of models. By looking at a series of models rather than one complex, all-inclusive model, one can see better the importance of each issue. What does including this or that particular effect do to the analysis?

The exploratory process should also be supported by *sensitivity analyses*. Sensitivity analyses can focus individuals attention on those issues that are key, i.e. do much determine the guidance offered by the analysis. They also aid group communication, by allowing each member in turn to see where other members 'are coming from'. For further details, see Section 5.3.

Review decision models

Decision analyses guide the evolution of the DMs, analysts and others perceptions. During an analysis everybody's perceptions evolve – or at least have the potential to do so. Their perceptions evolve because of the analysis: indeed, it is the purpose of the analysis that they should. Thus it is vital to see the modelling process involved in representing their perceptions as creative, dynamic and cyclic. The DMs', experts and stakeholders beliefs and preferences are assessed and modelled; the models are explored leading to insights and a revision of their judgements, and thence revision of the models used. The process cycles until no new insights are found. Phillips (1984) describes this evolution, referring to the process as *requisite modelling*, the final model being requisite or sufficient for the inference or decision faced.

Being aware that one needs more analysis is relatively straightforward; but knowing when to stop is harder. Requisite decision modelling

[9] Note that in addition to the decision analytic model of a form such as those discussed in Chapter 4, there will also be consequence models which predict the consequence of particular actions under particular states. These models incorporate scientific knowledge of the context.

requires that the modelling process cycles until no new insights on the part of the DMs and analysts are discovered. Inevitably this rather vague 'termination condition' annoys those schooled in the hard sciences. How does one tell that there are no new insights? How can it be known for sure that one further round of modelling will not uncover a major new insight? What is an insight anyway; and what makes an insight significant in terms of keeping the analysis going? Not easy questions to answer – and certainly they are impossible to answer in any manner that leads to a clear-cut, unambiguous test of when to end an analysis. It is a matter of skill and judgement on the part of the analyst and agreement on the part of the DMs. That being said, we have never been involved in a decision analysis during which it was not clear that the time had come to finish modelling and decide. There are a number of possible indicators:

- the DMs, experts and the analysts cannot find any significant assumption to doubt;
- sensitivity analyses show that the indicated choice is robust to reasonable variations;
- everyone concerned is comfortable with the analysis and feels that he or she have sufficient understanding to decide;
- the time is up – external factors may demand that a decision *is* made.

Finally, decision making is not about avoiding risk entirely. Risk is part of life. "No pain: no gain" is a common adage; "No risk: no gain" is a better one. Decision making is, in part, about evaluating risk and identifying courses of action which balance the risks involved with potential benefits (or avoidance of potential harm). The decision making process itself involves risks such as:

- have all potential uncertainties, advantages and pitfalls been anticipated?
- have all stakeholders been identified?
- are the data and judgement used in the analysis in the right 'ball-park'?
- are all the assumptions justified?

No matter how good the analysis, some of these risks will remain. So in the end DMs will need to accept the remaining level of risk. They simply have to choose and implement a course of action.

5.2 Soft OR Methodologies and Problem Formulation

Introduction

All analysis, problem solving and decision making has to be sensitive to context, both internal and external to the DMs' organisation: *c.f.* the concept of *appreciation* (Lewis, 1991). The DMs must maintain, revise and reflect upon the ideas and notions which shape their and their organisation's understanding of itself and its environment. Thus within the information gathering phase of a cycle of decision making, one of the first things that the DMs must do is to discuss and explore context and eventually set the boundaries for their decision making.

The easiest way to open up and begin modelling is simply to ask the open question: "What are the issues and concerns that are drawing your attention?" As the discussion flows in answer to this question, the key points can be noted down in a list or, perhaps, a number of lists: external factors, opportunities, threats, constraints, ... If one is working with a group, it is often useful to write each point on a *Post-It* and stick it to a board. This allows one to construct lists and groups of related concepts as the discussion continues.

This process can be made more effective by using formal brainstorming techniques which seek to generate ideas in a manner which overcomes intra- and interpersonal barriers to creativity (Rickards, 1999). The simplest approaches to brainstorming do little more than we suggested above: ask the participants to list uncritically all the issues that seem vaguely relevant to the general problem they face, but there are many variants which introduce more structure to the process to catalyse thinking without biasing perspectives. The key thing is to be spontaneous and non-evaluative: to get the ideas out for later consideration.

Brainstorming draws out 'top of the head' ideas from the DMs, but there is no guarantee that it draws out *all* the ideas they need, nor does it help organise them. There are a much broader range of techniques of varying degrees of formality which seek to pull out and arrange informatively the issues and concerns that they must address in their decision making along with the environmental context – including constraints, threats and opportunities. These techniques are known variously as *soft-modelling, soft systems* or *soft-OR*.

We shall consider four categories of soft modelling:

- check-lists
- simple two-dimensional plots
- trees and networks
- rich pictures

A further technique, scenario planning, is also important in that it catalyses creative thinking about possible futures, but it does not lead to 'pictorial' representations of the issues in the same way as soft modelling techniques do. While soft modelling helps set the context so that the analyst may build appropriate decision models to reflect the DM's evolving perception of the issues that she faces, it can often be useful to build rough, outline decision models in the early problem formulation phase of decision analysis. Thus, although attribute hierarchies, decision tables, decision trees and influence diagrams are seldom thought of a soft modelling tools, they can be so used.

The key thing in using any soft modelling technique is not to apply any of the methods too rigidly. They are tools to help the DMs think – or rather to get them thinking. Thus their role is to stimulate not constrain discussion. In developing any particular soft representation of the issues, any differences of judgement between the DMs will stimulate discussion and enhance understanding. For example, disagreement as to whether a stakeholder should be counted as having high or low power can prompt clarification of exactly what options are available to the stakeholder, and what effects they would have. Note that these discussions should not be thought of in isolation. Thinking about and discussing the roles of stakeholder can, for instance, provide insight into the DMs' objectives or key uncertainties.

These notes have been framed in terms of decision making – not surprisingly given that they have been written to support discussion of decision analysis and decision support systems. However, it should be noted that soft modelling techniques have a much wider application. They are valuable in supporting – or rather setting the context for – many forms of analysis: from information systems design and implementation to product innovation[10]. In developing an analysis or a design there is a need to be clear on context, on what is being assumed and what the objectives are. It is worth remembering that many information systems projects flounder because too little attention is paid to the human and organisational aspects of the system. Using soft modelling methodologies can identify and focus attention on key issues in these areas, leading to more successful projects.

Finally, beware of some of the 'branding' present in the soft-OR literature. Many of the methods have been developed, marketed and championed by particular decision analysts and management scientists with an apparent suggestion that *their* methods obviate the need to apply anyone else's. Our view is much more catholic: use any method that helps the DMs think and formulate their problem. Moreover, don't be rigid: stick to the basic motivation of a technique but modify it as discussion develops to fit with the DMs' evolving thinking.

Good general discussions of soft modelling techniques may found in Daellenbach (1994), Eden and Radford (1990) and Rosenhead and Mingers (2002). See also French *et al* (1998).

Check-lists

Check-lists are a very simple development of brainstorming in which the DMs are prompted with key words focusing on generic aspects of the context of their problem and then asked for their immediate and specific thoughts on these with little evaluation or reflection. There are many check-lists suggested in the literature. We only give a few here to illustrate the range of possibilities. Also note no single list will serve to remind them of all the issues they may need to consider. Nor are they mutually exclusive: there are many overlaps. The analyst will need to choose to use those that seem natural for the problem in hand.

PEST and 7'S

A helpful checklist for external context is *PEST*. Factors that need considering may be categorised under the headings:

- Political
- Economic
- Social
- Technical

For internal context, the *Seven S's* gives a useful list of factors that contribute to the context:

- Strategy
- Structure
- Systems
- Style

[10] A perfectly tenable position is to argue that design and innovation are no more than an application of decision making: design and innovation require many decisions on details as well as decisions on the general direction of the development.

Argument-based prompts	Strategy-based prompts
1. What do you think might be the cause of this problem or set of issues?	6. Can you think of any factors which would make this proposed action fail?
2. Can you think of any similar situations that might help in thinking about this matter?	7. Under what scenarios would this action work?
3. What class of risk issue do you think we are looking at and why?	8. Under what scenarios would this action *not* work?
4. Have you heard anything recently which seems relevant in some way?	9. Why do you favour/dislike this action?
5. Can you think of any indications which would provide evidence of an event or its absence?	10. Why might others disagree with you about the suitability of this action? What eventualities might concern them?

Table 9: Prompts which may help in identifying uncertainties

- Shared values
- Skills
- Staff

Note that the first S refers to the higher level strategy that forms the context for more detailed tactical and operational decisions. At least three of the S's – style, shared values and staff – relate strongly to organisational culture, which may embody many cultural aspects relating to the national identity.

SWOT

Check-lists need not be linear. SWOT – or **s**trengths, **w**eakness, **o**pportunities and **t**hreats – prompts thinking via a simple tableau: see Table 8. It requires that the DMs identify precisely those four sets of issues insofar as they, their organisation and their environment are concerned. SWOT analyses help DMs explore the context in which they operate. In particular, the strengths and weaknesses refer to their internal context: they relate to the Seven S's. Similarly the opportunities and threats refer to the external context and relate to PEST.

Prompts to identify uncertainties

In order to build a decision tree or influence diagram, one needs to be clear on the key uncertainties. Some will be obvious, but others may only become apparent after prompting. Useful questions (*c.f.* Browne *et al*, 1997) are given in Table 9.

CATWOE

Many soft OR techniques have within them checklists, i.e. lists of concepts that any picture should address if it is to provide a full perspective on the concerns and issues of the DMs. A particularly useful one is known by the mnemonic CATWOE (Checkland, 1989; Checkland and Howell, 1998; Rosenhead and Miongers 2001,). See Figure 41 for an explanation. Whereas we would not argue that one need use all or indeed any of these particular mapping techniques to explore the DMs' concerns and issues in a particular problem, we would argue that prescriptive support was flawed if it did not, at least implicitly, address and identify the factors which constitute CATWOE.

Checkland (1989) defines *Weltanschauung* as "the stocks of images in our heads, put there by our origins, upbringing and experience of the world, which we use to make sense of the world and which normally go unquestioned." It is important that an analyst recognises the importance being in tune with the DMs' worldview and does not seek to impose his own or some expert's worldview on an analysis – at least without substantial and open discussion with the DMs. Much of the recent

Strengths:	**Weaknesses:**
• …. • ….	• …. • ….
Opportunities:	**Threats:**
• …. • ….	• …. • ….

Table 8: Format of a SWOT table

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

Formulation of Root Definitions

Consider the following elements: CATWOE

C	customer	Who would be the victims/beneficiaries of the purposeful activity?
A	actors	Who would do the activities?
T	transformation process	What is the purposeful activity expressed as: input → $\boxed{T}$ →output
W	Weltanschauung*	What view of the world makes this meaningful?
O	owner	Who could stop this activity?
E	environmental constraints	What constraints in the environment does this system take as given?

Example

A professionally-manned system in a manufacturing company which, in the light of market forecasts and raw material availability, makes detailed production plans for a s defined period.

CATWOE analysis:

C	people in the production function
A	professional planners
T	need for a production plan → need met; or information → plan
W	rational planning of production is desirable and is a possibility; there is a degree of stability needed to make rational planning feasible.
O	the company
E	staff and line roles; information availability

Figure 41: CATWOE as defined by Checkland (1989, p87)

debate about health, food and environmental risks has centred around the difference in worldviews between government scientific advisors and those of the public, often without the politicians conducting the debate realising that there is a difference. In many decision making contexts, albeit much less significant ones, there is still a need to ensure that all the participants' worldviews are understood and explored. Thus to support discussion there may need to be several soft models of the same context, each characterised by a different worldview. In that way debate of the real issues can be brought into the open.

2-Dimensional Plots

Much can be done by drawing two axes with suitable if informally defined dimensions and getting the DMs locate aspects of the problem, e.g. stakeholders or uncertainties, against these.

Stakeholder identification

It is possible to use simple plots to identify and prioritise attention on relevant stakeholders and unknowns. For stakeholders, the relevant dimensions are 'power/influence' and 'stake' (respectively, how much that party *can affect* what happens, and how much they *are affected by* what happens). This affects how one might consider dealing with the relevant stakeholders, as suggested in Figure 43. *Players* are stakeholders with high power and influence as well as high stake. They are the people that the DMs need to work with or identify as competitors: whichever be the case, the DMs must be careful to manage and, ideally, try to control the agenda of interactions with players. *Victims* and *Beneficiaries* are those stakeholders with high stake, but little influence or power. How the DMs react to them depends on their feelings of altruism and responsibility to the respective stakeholders. Governments and their agencies may have a 'legal' obligation to be altruistic; others may simply have to react to or ignore a moral obligation. *Loose cannons* have high power, but little stake. The unpredictability of their actions can add significantly to the uncertainties facing the DMs, thus one tactic is to try to change the environment and external systems so that the loose cannons find that they do have a stake and so have to become more rationally involved. Finally, there are bystanders who have neither influence or a stake. However, it should be remembered that they may become more involved in the future; so it might be wise to monitor some of the bystander groups if this is likely.

An example of a stakeholder plot is given in Figure 42. This was developed in a training exercise within the UK Department of Health on health risk management. The hypothetical scenario concerned some evidence from an epidemiological study which suggested – inconclusively – a possible increased risk of laryngeal cancer within groups of asthmatic children who were taking a particular drug for their asthma. The balance of risks was still in favour of maintaining the children's treatment with the drug, since asthma itself is a life threatening condition if left uncontrolled. Health managers in the Department of Health

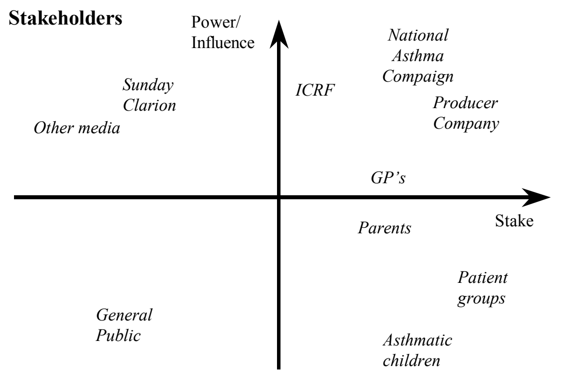


Figure 42: Stakeholder plot in a hypothetical scenario concerned with risks relating to an asthma drug.

were considering their possible strategies for advising the public without creating a 'health scare' (Bennett *et al*, 1999). A complicating factor was that one Sunday newspaper was known to be aware of the issues and might place the information in the public domain independently of the Department of Health. Thus the actual underlying decision problem concerned the timing and tenor of any government press release.

One way of developing this plot and also using the process to structure group discussions is:

1. A pair of axes are drawn on a blank flip-chart. Meanwhile the DMs, working individually, note down their first thoughts on the identity of stakeholders on 'post-its', up to a set maximum, say five each.
2. They then stick the labels in what they think is the most appropriate position on the chart.
3. Useful points for discussion (which should be noted) include the extent to which different individuals provide the same items, and their agreement or otherwise as to where they should be placed.

Alternatively, instead of asking each DM to work individually and locate a few stakeholders on the plot, the analyst may lead a quick brainstorming session to identify stakeholders and then guided by open discussion and evaluation between the group members locate each stakeholder on the plot. Either way the process is catalytic in that it draws out from the group the identities of potential stakeholders and their importance in the analysis.

Uncertainty identification

A similar plot of the key uncertainties may also be useful. See Figure 44. The unknowns can be classified, firstly, according to their importance either in affecting what may be done or their impacts upon the DMs and, secondly, according to the degree of unpredictability. Here 'degree of unpredictability' is not a simple probability scale. A certainty arises when one is fairly sure that an event will *not* happen just as much as when one is. Thus, in a very informal sense, the underlying probability scale 'doubles back' on itself. The reasoning behind this classification is to identify important external factors which are currently highly unpredictable and, hence, which may be well worth further investigation. Efforts to obtain further information should, in the first instance, be focused on uncertainties which are plotted in the upper right hand quadrant. Other

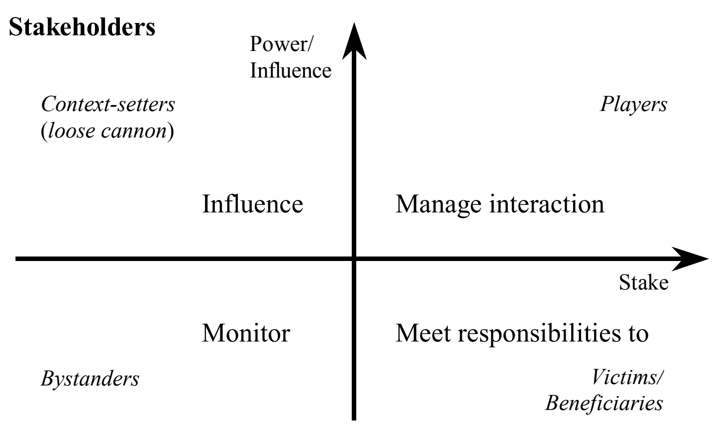


Figure 43: Stakeholder Identification

 15/02/2003
Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

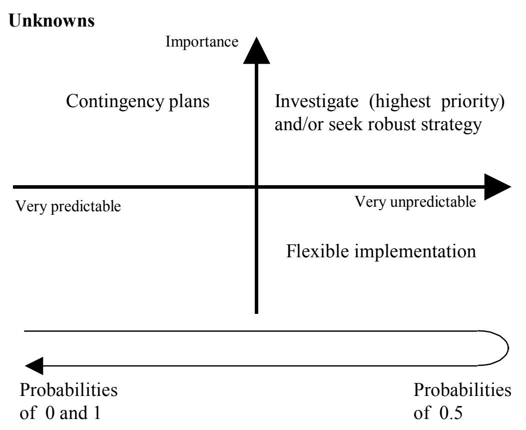


Figure 44: Identifying Uncertainties

uncertainties may be allowed for by developing contingency plans or implementing any strategy in a flexible manner.

Networks

Networks seek to elicit, display and stimulate discussion of the relationships between concepts.

Mindmaps are very simple plots that enable one to connect and begin to associate ideas. Figure 45 provides a mindmap of some of the issues relating to the communication of risks to the public. As can be seen, mindmaps do no more than take a major concept or set of issues in a problem and 'join' to this related concepts, braking down problem to enable the DMs see the issues more clearly.

Cognitive Maps

Figure 46 shows part of a cognitive map that arose in a study by Belton, Ackermann and Shepherd (1997) to develop a strategy for the supplies department in a Trust Hospital. Without going into the notational conventions used in cognitive mapping – this is a particularly simple example – the picture that this paints of the issues facing the Supplies Department is intuitively simple to follow. See Eden and Ackermann (1998) for a full description of cognitive mapping. Issues are related to each other by network showing the associations in the perceptions of the DMs. More sophisticated cognitive mapping notes the direction of association between concepts: i.e. does the presence of one factor make the other more or less likely? It is also possible to categorise the concepts in a cognitive map into objectives, external factors, key uncertainties, etc. Moreover, software tools allow one to look at a map from a number of viewpoints and to compare the maps of several stakeholders. But even with such sophistication the maps are intuitive and help capture and convey perceptions.

Rich Picture Diagrams

Soft systems methodology emphasises the value of *rich pictures*. These seek to explore and summarise issues much more pictorially than the methods we have discussed so far. They can be extremely intuitive and can also compress a lot of information into a single picture. Rich pictures can be very useful in forming a backdrop to subsequent analysis, acting as an *aide memoire* to allow the group continually to refer back and check they are addressing all the relevant issues. Figure 47 shows the many interacting systems and issues that relate to the hole in the ozone layer which has developed over Antarctica. Within a single picture all the key points and issues have been captured and may be communicated to those concerned in deciding what action to take.

Many analysts are scared of using rich pictures, because they feel their artistic skills are not adequate. However, with modern drawing tools in Office tools and inventive use of clip art, this is now much less of a problem.

Scenario planning

Another technique, although not really a soft modelling one, which may be used to open up DMs' thinking to a range of possible futures is

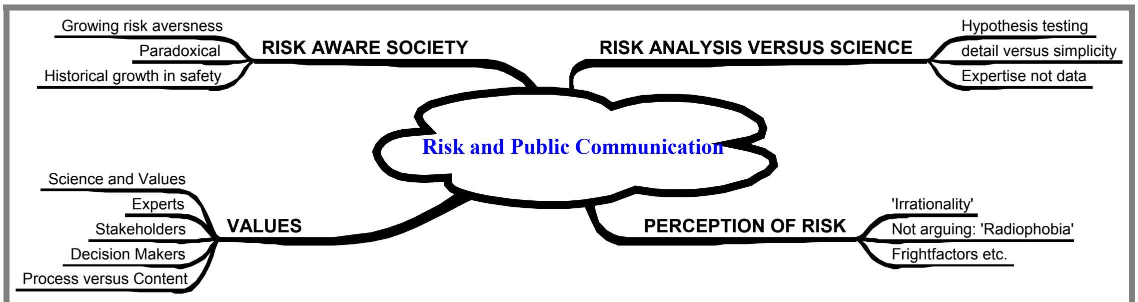


Figure 45: Mindmap of some of the issues in public risk communication

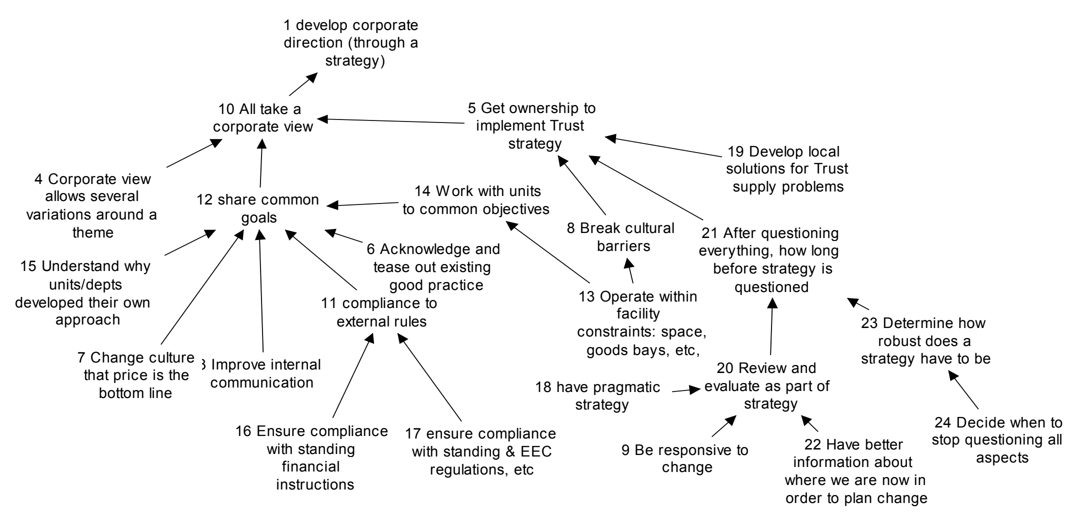


Figure 46: Cognitive map of issues arising in defining a supplies strategy for a Trust hospital (from Belton, Ackermann and Shepherd (1997)

scenario planning. This encourages the DMs to explore how the world may evolve and how other stakeholders may react to their actions. Scenario planning helps the DMs think about not how things are, but how things might be; and, in particular, it stimulates contingent thinking.

Broadly speaking the approach has eight stages.

1. Identify the key variables in the decision. This may be done using some or all of the tools described above. It is important to define the time frame for the decision and the time frame over which each variable is important.
2. Identify the key stakeholders in the decision. Again this may be done by any of the methods above.
3. Identify and list potentially important economic, environmental, political, technological and social trends which may affect the consequences of the decision – or constrain the choice available.
4. Identify the key uncertainties.
5. Construct two ‘extreme’ preliminary scenarios by assuming all good outcomes of the key uncertainties in the first and all the bad outcomes in the second.
6. Assess the self-consistency and plausibility of these extreme scenarios, and modify the scenarios if necessary so that they are self-consistent and plausible, maintaining as much as possible of the optimism of the first and the pessimism of the second.
7. Consider the likely response of the stakeholders identified in step 2 to the outcomes in the two scenarios, and again modify the story in each to make it more plausible and self-consistent, while preserving its essential optimism or pessimism.
8. Now create a number of less extreme, plausible and self-consistent scenarios.

Note the emphasis on self-consistency or, in other words, contingent thinking. How might some actions on the part of some players interact with the DMs’ own actions. What would their response be?

Usually one builds scenarios by working with a group of DMs, experts and stakeholders. Moreover, rather than get the whole group to take part in the development of each scenario, there are advantages in using assigning the tasks of building scenarios to subgroups, so that a variety of scenarios are developed in parallel with a great deal of independent and free thinking.

The essential secret to good scenario planning is to tell a good story – and the greatest risk in scenario planning is to tell a good story! When used properly the method opens up the DMs’ minds to a range of possible futures. It sensitises them to the uncertainties that face them and makes them think about their and other stakeholders responses to events as they unfold. The danger is that stories can be

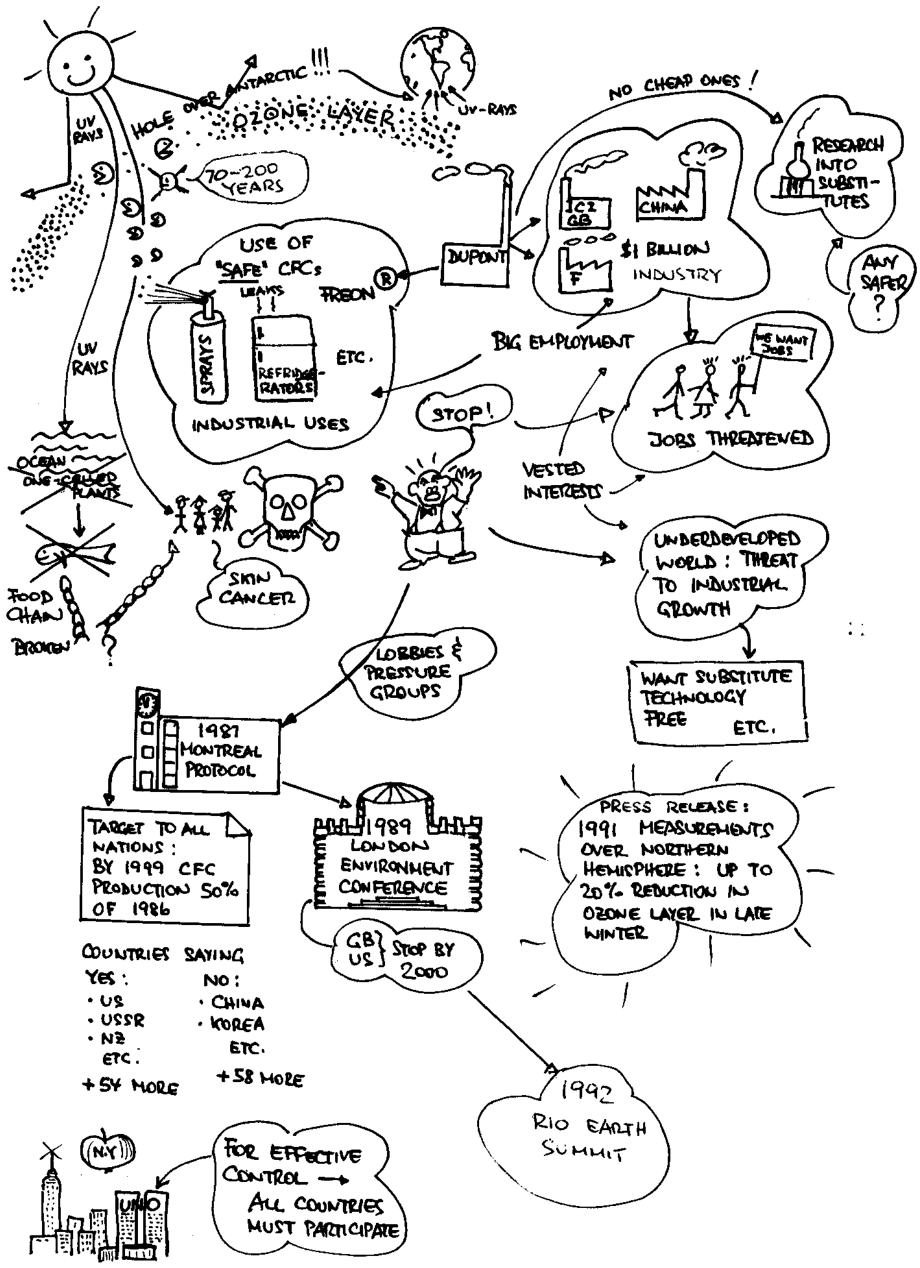


Figure 47: A rich picture diagram of the issues relating to the hole in the ozone layer. From Daellenbach (1994, p55)

believed, and psychologists have shown that plausible stories often are. Thus it is vital that several scenarios spanning a wide range of possibilities are considered simultaneously to keep the DMs aware that there is no claim that any one of these scenarios is *the* future.

Used effectively, scenario planning helps DMs

- understand the impact of key uncertainties
- think through responses of key stakeholders
- devise contingency plans and robust strategies – indeed, it is the ability of scenario planning to help DMs identify and explore contingencies which is arguably its strongest feature.

and, generally,

- understand all the issues better.

Scenario planning came to the fore in the early 1970's when it was used at Royal Dutch Shell to explore potential shocks to the world oil prices. The method was spectacularly successful in that Shell weathered the oil crisis in 1973 better than most of its competitors. Its managers had thought through the possibilities

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

and identified robust strategies with contingency plans. Its success has perhaps been over emphasised with some consultants seemingly recommending it not as a part of the tool-kit of decision analysis, but as a complete toolkit for thinking through possible futures and identifying strategy. The interested reader may find discussions in Kleindorfer *et al* (1993) and Schoemaker (1993). Wright and Ayton (1999) discuss the inclusion of scenario planning into the decision analytic process, dubbing the approach *future-focused thinking*, perhaps over emphasising the contribution of scenario planning relative to other soft approaches to creative problem solving, on the one hand, and the importance of clarity on objectives stemming from value focused thinking, on the other.

5.3 Sensitivity Analysis

In all quantitative procedures, it is a truism that the output depends upon the input. The significance of this dependence varies from input to input. Moreover, one is generally more confident in some inputs than in others. Thus a wise analyst varies inputs and observes the effect on the outputs: he performs a *sensitivity analysis*. Doing so advances his understanding and that of his client. Note that sensitivity analyses in Sections 4.3, 4.4 and 4.5

There are two ways that the inputs may be varied:

- Deterministic variation. Here the analyst has no guidance to the confidence in the input value available in a probabilistic form. Maybe there are bounds upon the 'permissible' values and thus deterministic calculations may be undertaken to see the range of outputs which may arise from input values in the permissible range. This is an archetypal example of sensitivity analysis.
- Stochastic variation. Here the analyst's confidence in an input value, *viz*. data point or parameter, can be encoded by means of probabilistic measures: ideally a full distribution, but often only quantiles or variances (and other moments) are available. In this case the inputs are varied stochastically in Monte Carlo analyses and the resulting distribution of the outputs studied: see Saltelli *et al* (2000) for a recent survey of these techniques.

Within decision support, the circumstances for stochastic sensitivity analyses tend to arise in relation to inputs related to the physical models, whereas those for deterministic sensitivity analyses tend to arise in relation to judgemental inputs such as preference weights or parameters in subjective probability distributions. Quite simply, if the analyst asks his client or an expert to capture her confidence in a judgement by means of a subjective (judgemental) probability distribution, he enters an infinite regress of eliciting judgemental uncertainty in judgemental uncertainty in judgemental uncertainty …. French and Rios Insua (1991) discuss a general framework for sensitivity analysis which is applicable to decision analyses: see also French (1995, 2001).

To appreciate the way in sensitivity analysis enters decision support we need three perspectives: a technical, or model-based, perspective; a cognitive, or individual, perspective; and a social, or group, perspective (Belton and Stewart, 2002).

The Technical Perspective. Technically, sensitivity analysis is the investigation of the effect of changes in the data input to a model on the suggested ranking given by that model. In this case the inputs are: the scores which indicate the evaluation of alternatives with respect to criteria at the bottom level of the hierarchy; the weights which indicate acceptable trade-offs between criteria; and, perhaps, subjective probabilities. From a technical perspective, sensitivity analyses enables the analyst and DM to explore whether there is a clearly preferred alternative, or whether there are several strongly competing alternatives. Sensitivity analyses enable the identification of dominated and near-dominated alternatives; that is, those alternatives which would never be preferred whatever weights were assigned to criteria.

The Cognitive or Individual Perspective. From this perspective, sensitivity analysis is the process of interactively exploring the effects of changes in the inputs to and structure of a model in order to learn about the problem and about the DM's values and priorities. The viewpoint here is the growth in understanding of each individual DM of her perceptions. See especially Keeney (1992) and Phillips (1984).

The Social or Group Perspective. Sensitivity analyses can help groups of DMs focus on their *real* differences. Often a heated discussion of issues on which individuals fundamentally disagree can be completely defused by a sensitivity analysis which shows that, despite differences on the appropriate weights or scores, the decision should be the same.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

5.4 Group Decision Making

> *...we cannot solve present day major political and organisational problems simply by grinding through a mathematical model or computer algorithm. What we require besides is the design of better deliberation and judgement.*
>
> *C.W. Churchman & H.B. Eisenber*

Most decisions, at least in organisations and society, are the responsibility of groups rather than individuals. Yet many of the theories we have discussed are based upon the concept of an individual DM. The SEU model, in particular, is built upon a conception of the beliefs and preferences of an individual decision maker. Can it be generalised to group decision making contexts? Unfortunately, no. There have been many explorations of the mathematics of group consensus probability and utility functions and all have discovered that any formalism for building a single model of group preference is subject to irrational or undemocratic behaviours. Impossibility theorems abound, suggesting that the endeavour is ill-formulated mathematically. A classic impossibility theorem due to Arrow suggests that not only is it impossible to extend SEU to group decision making, but the very idea of democracy may be an ill-defined concept. If this is so and if most decisions have to be made by groups, why have we been studying the normative formalisms that we have? How can they help in the majority of real circumstances?

The trick to 'squaring this circle' is to change our perspective on groups as decision making bodies. They are not. Groups are social processes which translate the decisions of its individual members – which way to vote – into implemented courses of action. This perspective does not contradict the assignment of accountability, authority and responsibility to a group. Groups can have duties and powers. What we do insist is that any cognitive activity must take place in the brains of individual group members. Only individuals can think, reason, analyses, hold beliefs and preferences, judge, evaluate and decide.

If we adopt this perspective on 'group decision making' as a social process, the task of supporting a group becomes one of supporting that process as well as supporting the decision making of its individual members. Thus, remembering the potential for dysfunctional or aberrant group behaviour, we shall seek to develop support techniques which:

- foster effective communication between the members;
- explore the issues in a creative, effective manner;
- reduce unproductive tensions and disagreements;
- protect the group from dysfunctional activities such as groupthink;
- build a shared understanding;
- build a commitment to implement the selected course of action;
- record and report their discussions

and, at the same time,

- support each member's own thought processes, judgements and decision making.

Decision conferencing is one way of providing groups with such support. This draws upon three key methods:

- facilitation, in which a facilitator who has no responsibility or accountability for the consequences of the decision, joins the group to structure, smooth and enhance the deliberative processes;
- decision analytic models to help the DMs understand themselves, the context and the issues before them;
- interactive IT to explore and dispay the implications of the models.

Groups may fall prey to various biases: we have noted groupthink. One of the key roles of a facilitator is to counter these biases and other dysfunctional behaviours such as(Barron *et al*, 1992):

- Status effects: some members views may be over or under valued simply because of their status, within the group *per se*, and perhaps in the wider environment.
- Group size: more people mean more experience, more knowledge, more intellect, but also more difficulties in communication and managing the group discussion.
- Lack of commitment and free-riding: sometimes members 'buy-out' and do not commit themselves to the group problem solving and discussion, perhaps because of inherent laziness or because they feel undervalued by or inferior to the other members.

- Differences of opinion: arguments may arise because of differences of opinion (which may be irrelevant to the issues at hand) and burn up a lot of emotion and good-will before they are defused.
- Overconfidence: group discussion may reinforce the assumptions behind a forecast and so lead to underestimation of the uncertainty therein.

The manifestation of these behaviours may vary from group to group and from issue to issue, but the result is the same: whereas a group should be at least as effect as any of its individual members, it can become less effective.

A *facilitator*. is skilled in the process of group discussion, but seldom does he have any expertise in the context of the issues at hand, and even more seldom would he use such expertise in the discussion. The facilitator's role is to smooth the group's work, to help the process and make the team more productive, more creative. Phillips and Phillips (1993) summarise the key functions of a facilitator as: observing, attending, maintaining awareness of feelings and intervening. The content of the discussion comes entirely from the group themselves. They 'own' the problem, have knowledge of it, access to relevant data and experts, and are responsible for its resolution.

In a sense, a facilitator is no more than an impartial chairperson or group leader, but in practice his 'distance' from the group is far greater. Because he does not share in the ownership of the problem, he may concentrate on:

- encouraging member of the group to contribute ideas and listen to those of others;
- assuming responsibility for accurate communication between the members, perhaps cutting through jargon or simply making sure that no one is too shy to say "I don't understand";
- protecting minority views and ensuring they are debated fairly;
- being sensitive to unexpressed feelings and views and helping them enter the discussion;
- calming conflict by keeping the group task oriented rather than personality oriented;
- summarising the position at appropriate points in the discussion;
- generally, keeping the discussion moving and focused on the task in hand.

Facilitators concentrate their attention on the *process*, leaving the members of the group free to contribute, explore, shape and understand *content*. It is important that the group provide the content of the discussion, identify opportunities, create and evaluate the options to exploit these and generate the action lists to implement their decisions. Through their total involvement in the creation of strategy they become fully committed to its implementation. They 'own' the strategy. Moreover, because of their shared understanding of the reasons behind its adoption, they can explain the policy to others.

All the above suggests *what* a facilitator should do: *how* he should do it is another matter. How should he intervene to enhance the work of the group? Well, there are some tricks of the trade. Generally, the facilitator should raise issues neutrally, asking open questions. Although sometimes, when the group is drawing together behind a single viewpoint, he may play Devil's advocate and press an alternative view to test whether groupthink is rearing its ugly head. Ignorance is a great advantage: a facilitator can often move a group forward by asking a very simple question and uncovering hidden, perhaps contentious assumptions or misunderstandings between group members. Because he is an outsider, he often questions jargon and so clarifies discussion for all the group. Not everyone in the group may be *au fait* with all the jargon used in an organisation, but some may lack the courage to indicate their ignorance. For the facilitator there is no loss of face..

Not all interventions require the facilitator to distance himself calmly from the issues. Occasionally, there may be benefit in his pressing a point forcibly, particularly if a member of the group is trying to take control of the process or if the general level of stress has fallen below that needed for productive activity. But generally a facilitator is wise to hold his temper and intervene gently, catalysing rather than directing.

The process of decision analysis and its tools can of themselves provide very effective interventions. The prescriptive decision analysis cycle organises the general flow of discussion, moving through issues, one by one, concentrating on each in turn. This avoids the confusion that can be caused by simultaneously considering many issues and darting between them: the structure protects the group from hypervigilance Soft OR methods help foster creativity, providing a framework in which to brainstorm effectively. The model structuring, elicitation, evaluation and sensitivity analysis cycle helps move the discussion forward productively, focusing on one issue at a time. Sensitivity analysis can defuse heated but irrelevant debate, but concentrating attention on the issues that matter to the problem at hand. The developing model provides a very effective vehicle for communication.

In Section 1.6, we remarked that, although we would be developing normative models of *individual* decision making, we would use them in prescriptive *group* decision support in ways that would build shared understanding and help the group to a decision. Here we begin to see how that is accomplished.

In a *decision conference* or *facilitated workshop*, the group responsible for the decision meet together for a day or more, ideally away from their normal working environment, to discuss and explore the issues. Their deliberations are supported by a facilitator, who in turn may be supported by one or more decision analysts and, perhaps, a secretary to record the discussion.

A decision conference is generally a two-day event. Other time-scales are possible; but the inclusion of a night is considerable advantage. In the evening the group are able to relax together and reflect on the progress and discussion so far. This reflection, together with the distance from the previous day's deliberations that a night's sleep brings, helps members acquire a more mature perspective on the issues that concern them. Without the overnight break some may have second thoughts soon after the conference ends, perhaps on the journey home, and much of the value of the event will be dissipated as their commitment to its conclusions evaporates.

The entire group responsible for a decision should take part in the conference, which concentrates entirely upon the issues that led to it being called. There are no time-outs to consider peripheral matters 'while the team are together'. For that reason it is sensible to hold a decision conference away from their normal place(s) of work: perhaps a country hotel or a purpose built decision conferencing suite. All members must make themselves unavailable to other demands on their time: they must clear their diaries for the conference. Ideally, too, they should deny themselves the use of the 'phone. In short, the decision conference should focus entirely on the issues which have to be resolved.

The facilitator leads the meeting, guiding the discussion forward in a constructive fashion. He may be assisted by a decision analyst and, possibly, a recorder. The analyst builds the models and runs the software, generating decision analytic representations of the issues as they arise, which help the group gain insight into the situation facing them. The recorder uses a word-processor to record the development of the debate and the reasoning behind the judgements and decisions made. Because of the presence of the recorder, at the conclusion of the conference the group are able to take a record of important conclusions and an action list with them. A full report follows in a matter of days. More and more, the roles of recorder and analyst are becoming identified. In the early days of decision conferencing, the recorder and analyst needed a computer each; but with the advent of multi-tasking windowing environments it is possible for one person to fulfil both roles; and, moreover, there are advantages if a single person does. Inevitably, an analyst is far more closely involved with the process than a recorder and so better placed to record, for instance, the reasoning underlying a particular model.

Each decision conference is different. It evolves according to the needs of the group and not according to some fixed agenda. There are, however, common themes and patterns. The facilitator is always careful to ensure that the opening discussion is as wide-ranging as possible. It is a rare decision conference in which a single focus for discussion emerges in the opening few minutes. During the initial phase the facilitator may simply allow the discussion to develop or he may use any of the soft OR modelling techniques alluded to in the previous chapter. These can provide useful structure to the perspectives being developed. Sometimes the development of a (set of) soft OR models can occupy the entire event, particularly those that focus on setting the strategic intent and mission of an organisation: see e.g. Eden and Ackermann (1998).

With the issues and context defined, the next step is to develop a more formal decision model. The decision analytic models commonly used in decision conferences are very simple. More often than not, they are based upon additive value theory. The important attribute of these models is that they can be explained quickly and effectively to the team. The mathematics involved is no more than addition and multiplication. Within a short time of encountering the models, the team can draw valuable insights from their use. While it is possible to use more complex models, the time span of the event may need extending or, perhaps, a series of event need be run with some aspects of the model being built 'off-line' between meetings. Surprisingly, the analysis in decision conferences needs much less hard data than one would, at first, think. Strategies have to be costed: that is clear. But the costings need only be rough. It is a broad brush picture that the event seeks to create. Detail can be added at a later date. Such rough data is usually within the power of the group to judge upon the spot.

The decision model should now be well enough defined to give a first indication of 'an optimal course of action'. With good scheduling by the facilitator, this will occur just before the overnight break, giving the DMs the opportunity of an extended period to reflect upon the analysis to date. By the next morning they will be usually be clear on one thing: whatever the model is doing, it is *not* reflecting an optimal course of action. They will note flaw upon flaw with the analysis: forgotten issues, ill-formed judgements and so on. The associated feelings of frustration – "did we waste yesterday entirely?" – will provide added impetus to revitalise their discussion. Remember the importance that Janis and Mann (1977) place upon the catalytic power reasonable levels of stress. The facilitator and his team will rebuild the model, adding additional features as necessary and explore it via sensitivity analysis. The decision analytic cycle continues until the model is requisite.

The final stage of a decision conference is to work with the group to provide a summary of the conclusions and strategy that they can take with them, along with an action list to implement the decision. One of the reasons for the success of decision conferences is that the DMs are highly motivated to implement their strategy which has been developed using their model. This final stage of summarising and allocating tasks ensure that they may return to their usual working environment and colleagues able to communicate the conclusions succinctly. Decision conferences ensure that the participants own the strategy, are committed to its implementation and can communicate their arguments and enthusiasm effectively.

6 Operational Research

6.1 Introduction

The process of decision analysis described in the previous chapter is focused primarily on supporting decisions in the corporate strategic domain. As we noted, see e.g. Figure 2 page 6, decision making there is unstructured and deals with long time-spans of discretion. Soft modelling and problem formulation becomes an integral part of most decisions. As we move into the tactical domain, problems tend to be more structured, but also more detailed and with clear objectives. Thus while some decision analytic techniques are also applied her other methodologies are more appropriate, typically the methods of operational research (OR). We shall not describe these in detail here, only indicate their broadest structure: see, e.g., Daellenbach (1994), Ragsdale (2001) or White (1985) for more details.

6.2 Linear and Mathematical Programming

OR based decision support techniques focus primarily on operational decisions which are easily structured, usually involving short to medium time-spans of discretion: and, most particularly, in which clear measures of effectiveness could be defined by the DM. Thus OR is applied to problem areas such as: inventory and stock-control; logistics and routing; repair and maintenance; production planning; queuing; scheduling; and, generally, the (detailed) allocation of scarce resources. In simplistic terms the general approach is to build a model of a system and optimise its working with respect to some straight-forward objectives: e.g. maximise profit and minimise time from factory to market. But OR gains its subtlety in that it does not take the optimal solution within the model as more than a guide as to what to do in the real world. OR analysts are well aware that their models are simplifications of reality which seek to capture the key factors that affect the decision. Thus the optimal solution in the model provides no more than a guide to a direction for change that may lead to an improvement in the effectiveness of the real system. Moreover, there is a cycle of adjusting the model until it represents the reality of the system *requisitely*. The process of OR is described and discussed in White (1985).

The variety and the details of the many OR models and techniques need not concern us too

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

much. Most OR texts have chapters describing some or all of the following:

- mathematical programming including linear, quadratic, non-linear, combinatorial and dynamic programming;
- queuing and stochastic process models;
- reliability and maintenance modelling;
- scheduling including critical path analysis (CPA) and programme evaluation and review techniques (PERT);
- simulation methods.

All essentially solve the same archetypal constrained optimisation problem[11], though few would recognise this generic structure in many typical OR applications.

optimise $\pi(\boldsymbol{a})$ with respect to $\boldsymbol{a}$
subject to $\boldsymbol{a}$ being in A,

where:

$\boldsymbol{a}$ – are the *decision variables*, i.e. those quantities in the model which represent the things the DMs may choose or control in the real world. Within decision analysis the decision variables correspond to the actions in the action space. In many, but far from all, optimisations problems considered within OR, there are an infinite number of possible values of the decision variables; whereas in the majority of decision analytic problems the action space is finite.

$\pi(\boldsymbol{a})$ – is the *objective function*, i.e. the quantity in the model which the represents a measure of effectiveness that concerns the DMs in the real system. Although one speaks of an objective function in the singular, it may be a vector valued function, i.e. one which models a number of measures of effectiveness as separate components. The imperative 'optimise' then needs to be interpreted as seeking a solution which performs well against all the objectives rather than optimises any single one of them. This leads to a growing body of methods known generally under the headings: multi-objective decision making (MODM) or multi-criteria decision making (MCDM). We discuss these below in Section 6.3.

A – is the *constraint set*, i.e. a set which limits the possible values for a. A need just not provide specific bounds and ranges, but will also encode interrelationships between the different decision variables.

A mathematical programme may not look like a 'model' of a system, but to an OR analyst it is. It encodes all relationships that the analyst and DMs believe to be important in defining the working of the system.

The algorithmic details of optimising $\pi(\boldsymbol{a})$ and the precise specification of $\boldsymbol{a}$ and A, vary according to the class of problem being solved. For instance, logistical problems arising in supplying several outlets from a few warehouses can often be modelled by a special form of linear programme, called the transportation problem. Thus when faced by such a logistical problem the OR analyst will recognise the characteristics of a transportation problem, identify and gather the relevant data, input it into a standard software package for solving transportation problems, run the program and interpret the results with the DMs to set the operational details of the solution to the original logistical problem. Another example is that when faced with improving a service process seeking to meet a random stream of customers, the analyst will work with the DMs to define a queuing model of the system, then using algebraic, Markov decision process or simulation methods he will seek a service policy which optimises a performance measure and interpret this policy into operational details with the DMs.

Although the algorithms used to solve particular OR models are designed to take advantage of particular characteristics of the form of $\pi(\boldsymbol{a})$, $\boldsymbol{a}$ and A, in all but a few cases, they all have the same generic iterative structure.

First, an arbitrary trial solution $\boldsymbol{a}_0 \in A$ is chosen. Then an improved solution $\boldsymbol{a}_1 \in A$ is generated such that $\pi(\boldsymbol{a}_1) - \pi(\boldsymbol{a}_0) > 0$ – for definiteness, we assume that the problem is a maximisation one. In general, mathematical programming algorithms are distinguished primarily by the mechanism used to generate an improved solution $\boldsymbol{a}_1$, from the many possible. The simplex method of linear programming looks for an adjacent vertex of the constraint set, which is a convex polyhedron in this case.

[11] Or: "maximise profit by varying what you can control subject to constraints that you cannot surrmount" … or "minimise cost …" or whatever.

Simulation and Monte Carlo methods generate solutions at random, but often according to a distribution which is biased to give a higher chance of improvement than pure randomness. Next a second improved solution $\boldsymbol{a}_2 \in A$ is generated such that $\pi(\boldsymbol{a}_2) - \pi(\boldsymbol{a}_1) > 0$. The procedure iterates in a natural way, building a sequence of trial solutions until satisfactory convergence is obtained, i.e. when the difference between $\pi(\boldsymbol{a}_n)$ and $\pi(\boldsymbol{a}_{n-1})$ is suitably small.

Such iterative numerical methods are the lifeblood of computation and there are many software packages available which implement the procedures for particular classes of problem.

In applying mathematical programming methods as we have described them here, data are collected and judgements elicited; the model built and, in particular, the objective function is fully defined; then and only then is the model optimised. However, there are a class of algorithms which interlace the data collection and elicitation of judgement with the iterations in the optimisation algorithms.

6.3 Interactive Multi-Objective Programming

Suppose that the decision variables or actions in the problem have been represented against q attributes. Thus the constraint set A is a subset of q-dimensional space and a typical action has the form $\boldsymbol{a} = (a_1, a_2, \ldots, a_q)$. We shall consider only real-valued attributes. Moreover, we shall assume that the DM's preferences are strictly increasing in each attribute. Each attribute defines an objective (and that is why we speak of interactive *multi-objective* programming). In passing, note that this latter assumption implies that each attribute is preferentially independent of the rest.

In a decision analysis based upon multi-attribute value theory, the approach would be to assess a value function, $v(\boldsymbol{a})$ or in our notation here $\pi(\boldsymbol{a})$, and maximise this:

$$\begin{aligned} &\text{maximise } \pi(\boldsymbol{a}) \\ &\text{subject to } \boldsymbol{a} \text{ remaining in } A. \end{aligned} \tag{1}$$

We have tended to think of A as finite, but in many practically-occurring problems it is infinite, often a convex q-dimensional polyhedron. We shall assume that it is here.

Several criticisms have been levelled against techniques which seek first to assess and then to optimise a value function (Goicoechea *et al.*, 1982; Starr and Zeleny, 1977). First, the assessment procedure is very time-consuming. Moreover, it almost inevitably involves asking the DM to make hypothetical choices between alternatives which can often have no practical reality. Motivating her to consider and evaluate these choices is difficult; and, indeed, she may see their introduction into the analysis as an unnecessary confusion. Once the value function has been determined, the decision is implicitly made. All that remains is to optimise $\pi(\boldsymbol{a})$ on a computer. The optimisation process seldom, if ever, involves the DM; yet she often sees this stage as the point where the decision is made. Thus she feels excluded from the very part of the analysis that she believes she should be central to. Last and, perhaps, the most important of the criticisms is that the determination of $\pi(\boldsymbol{a})$ is an analytic process. Marginal value functions are determined separately, then combined. However, many writers on multi-criteria decision making argue that choice is a Gestalt process, in which alternatives are considered holistically. The value of an object is something more than the sum of its components parts: see, especially, Goicoechea *et al* (1982) and Duckstein *et al* (1975).

Looking back at the generic iterative optimisation algorithm described in the previous section, notice that the value function $\pi(\boldsymbol{a})$ is used only in two places: first, in defining an improvement at each iteration, and, second, in checking for convergence. Furthermore, it is possible that much of the information implicit in $\pi(\boldsymbol{a})$ is not used. Thus much of the time-consuming introspection and analysis used in determining $\pi(\boldsymbol{a})$ might be unnecessary.

Suppose, therefore, that we abandon the determination of $\pi(\boldsymbol{a})$, and instead begin simply with a trial solution $\boldsymbol{a}_0$. By an interactive dialogue with the DM, determine an 'improvement' $\boldsymbol{a}_1$ such that she prefers $\boldsymbol{a}_1$ to $\boldsymbol{a}_0$. Repeat the process again and again, each time by interaction with the DM determining a point $\boldsymbol{a}_n$ such that she prefers $\boldsymbol{a}_n$ to $\boldsymbol{a}_{n-1}$. Declare that the process has converged when she feels that $\boldsymbol{a}_n$ is satisfactory.

This, in essence, is an archetypal interactive multi-objective programming method. Notice how the criticisms raised above are avoided. The preference information obtained from the DM is precisely sufficient to find a satisfactory solution and no more. No, or little, redundant information on her preferences is obtained. She is not asked hypothetical questions; all her choices are between real alternatives. She is involved throughout the procedure and, indeed, central to it. Lastly, her judgements are made holistically. It is left to her to compare $\boldsymbol{a}_n$ with $\boldsymbol{a}_{n-1}$; no analysis into component preference orders is made.

Naturally, no interactive multi-objective programming method is as simple as this description. Usually much subtlety is used to suggest possible improvements to the DM. Some methods require much interaction with the DM, using preference information to check for improvement and also to ensure that the search for the next potential solution is computationally efficient. To avoid placing excessive demands upon the time of the DM, many authors, e.g. Zionts and Wallenius (1976), build two particular features into their algorithms. First, they limit the search to the Pareto boundary[12] of the constraint set, thus making use of the assumption that the value function $\pi(\boldsymbol{a})$ is monotonic in each of the objectives, $\boldsymbol{a}$. Second, they assume a particular functional form for the unknown value function. Often they assume that $\pi(\boldsymbol{a})$ is a known function and that only certain parameters are unknown. Many writers assume a linear value function with unknown coefficients, viz. $\pi(\boldsymbol{a}) = \sum_i w_i a_i$. The assumption that $\pi(\boldsymbol{a})$ is known up to some unknown parameters means that the preference information gained in constructing the sequence of improvements $\boldsymbol{a}_0, \boldsymbol{a}_1,\ldots,\boldsymbol{a}_{n-1}$ may be used to place limits on these parameters and these limits, in turn, may be used to restrict the search for an improvement $\boldsymbol{a}_n$.

Although the majority of interactive multi-objective programming methods assume that the DM makes all her choices consistently with an implicit, but unknown, value function within her; not all do. Some intend simply to help her explore the efficient set, confining her attention to the region in which she seems most interested at any particular stage. However, in all but a very few cases, the bare skeleton of interactive multi-objective programming outlined above underlies the methods proposed in the literature.

Now remember our discussion of behaviour decision studies in Chapter 2. There we noted that DMs' unaided judgements are often incompatible with the underlying assumptions of multi-attribute value models, e.g., they may be intransitive. Given this, interactive multi-objective methods risk building their analyses on inconsistent – and therefore inappropriate – data (French, 1984, Korhonen and Wallenius, 1996). Thus there is a need to build a prescriptive methodology for applying interactive multi-objective programming methods which draws together behavioural findings with the normative underpinnings of the methods: *cf*. Figure 5. This still has to be done, so for the present analysts and DMs should take great care in applying the methods and reflect hard on the input judgements at each iteration.

6.4 Management Science, OR, information Systems Engineering, …

OR began just before the Second World War. In 1937 A.P. Rowe instituted an activity at the Air Ministry Research Station in the UK directed at improving military capability. During the War OR matured into an essential and effective part of the Allied war effort, both in the UK and the USA. The name operational research – or operations research, as it is called in American English – shows its military origins: OR used multi-disciplinary teams of *scientists* to advise on how to make operations more effective. Its success was such in the years after the War that OR moved into civilian life and sought to bring the same improvement to industrial operations.

Originally there was a clear intention that the OR process should be scientific. A defining quality of the Scientific Method, especially as it was understood in the late 1940's and early 1950's, is that a scientist must be a dispassionate, detached, objective observer of a system. He must not interfere. This scientific imperative, pervasive throughout OR until the 1980's, meant that the OR process took a very different perspective on prescriptive decision analysis and support than we have. We have emphasised that the intention is to help DMs explore and understand both the system and *themselves*. There is an explicit intention to help them form and evolve their judgements. Decision analysis recognises, works with and seeks to evolve subjective inputs. Thus in reading the OR literature, particularly that older than 10 – 20 years, it is possible to encounter very strong arguments against the tack we have been

[12] The Pareto boundary of a constraint set is based upon precisely the same idea as the Pareto plots of Section 4.3, e.g. Figure 29. Namely, if preference increases with each attribute, then any optimal point in terms of the DM's preferences must lie on the upper right boundary. This boundary is called the Pareto boundary or efficient set.

taking here.; and the reader should be aware of that in synthesising his or her understanding.

However, that picture of OR painted is somewhat dated. Firstly, the need to support the subjective side of decision making has grown apparent to the majority of the OR profession. This has meant, particularly in the US, that the methods of decision analysis have become as much a part of the OR analyst's tool-kit as the techniques described above. Secondly, the new area of *soft OR* has grown up, see Section 5.2; and this takes a far more subjective perspective on helping DMs evolve their understanding than was common in the early days of OR.

OR is always described as a multi-disciplinary activity. A full OR study has inputs from many disciplines to describe the physical systems, the human interactions and the economic in addition to the application of complex mathematical and computer models. Usually there is a need for extensive data analysis and hence for statistical input. But OR is not the only multi-disciplinary activity which supports decision making. Management science is essentially multi-disciplinary. Information systems engineers need to multi-disciplinary in their outlook, because they provide computer systems which interact with the members of an organisation to provide the information that they need in making their decisions. Systems science would distinguish itself from OR, but in seeking to understand from many perspectives how complex systems work and evolve it provides valuable insights for decision making. Thus all these activities and professions are multi-disciplinary: they require skills and knowledge associated with a variety of scientific, engineering, human and social science disciplines to achieve their ends. Whether it helps to distinguish between all these multi-disciplinary activities is a moot point, but not for us to decide. We simply draw the moral that in supporting decisions we must be 'Jacks[13] of all trades'.

[13] And Jills!

7 Artificial Intelligence

7.1 Human vs. Artificial Intelligence

We now move forward to discuss AI approaches to supporting decision making. These, we argue, tend to be appropriate more to the operational and hands-on work domains.

Over the years many definitions of intelligence have emerged. According to Stenberg (1985), intelligence is the ability to adapt, shape and select environments. Along the same lines, Turban and Aronson (2001) define intelligence as the degree of reasoning and learned behaviour and argue that it is usually task or problem solving oriented. Intelligence is better understood and measured in terms of performance on novel cognitive tasks or in terms of the ability to automate the performance of familiar tasks. As Stenberg (1985) stresses, intelligence has three facets: analytical, creative and practical thinking.

An interesting test of whether a machine is intelligent was designed by Turing and is widely known as the Turing test[14] (Turing, 1950). According to the test, a machine is considered to be intelligent when a third party, who converses with both the machine and a human being without seeing them, cannot conclude which is which based on their responses.

AI is a broad term encompassing many definitions. Its goal is to develop machines that can mimic human intelligence. There are two main philosophies or schools of thoughts in AI. According to the first philosophy, AI is the study of understanding human intelligence by 'modelling the brain'. This approach is also known as connectionism and is being applied in research domains such as distributed processing and neural networks. The second philosophy aims at 'making a mind' through the representation of processes of human thinking in machines e.g. computers or robots. Research along these lines has focused on incorporating intelligence into computer-based systems.

There are tangible benefits arising from the use of AI, as opposed to human intelligence, in organisational settings:

[14] There is a suggestion that Turing rather had his tongue in his cheek when he designed the test, but be that as it may, it is now a recognised test in AI to identify when artificial intelligence has been created.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

Knowledge is more permanent. Employees might leave a company and take with them knowledge about a domain and about how to perform a task. AI however, can permanently store all this knowledge. A restriction according to our discussion earlier (Section 1.5) is that only explicit knowledge can be captured and used in *programmed* AI systems. Whether *tacit* knowledge can either now or at some time in the future be captured by AI systems that learn is a moot point.

Knowledge becomes easily accessible. It is not easy to transfer knowledge and experience from one person to another. AI allows the development of knowledge bases that can be accessed by all employees eliminating the need for data duplication.

Performance is improved. Computers, unlike human beings, can be easily switched on and off and transferred over to a new working environment. They do not have feelings and are not subject to stress or fatigue. Their performance therefore is always consistent. They can also automate tasks that people find tiring or unsatisfactory.

Reasoning and solutions can be documented. A computer can draw a conclusion or take a decision while documenting the rules or facts that contributed to its output. Alternatives that were used in the past to solve problems can be proposed again in similar cases. Human beings however, often find it difficult to articulate the reasoning behind their decisions or forget over time the arguments that lead to a decision.

Efficiency and effectiveness. AI can often reduce the time needed to perform a task. It can also help machines execute tasks better than people at a fraction of the cost required when using human assistants.

Despite the much recent progress in AI, many human intelligence characteristics are very difficult to mimic. Human beings are creative, have instincts, sense their environment and are repositories of vast quantities of tacit knowledge. Tasks, such as pattern recognition, performed so naturally by humans can prove to be difficult when undertaken by a machine. Even though AI is very powerful in narrow and well-defined domains it cannot be used to provide support in a wide range of problems.

7.2 AI technologies

There are many technologies which come under the general heading of AI.

Expert systems are computer-based systems that assimilate the reasoning and knowledge of experts to solve problems. We consider these at greater length in Section 7.3.

Natural language processing allows computers to communicate with their users in their native language rather than through menus, forms, commands or graphical user interfaces. It consists of two sub-fields:

- *Natural language understanding* that investigates methods for allowing a computer to comprehend instructions given by its computer users in ordinary English or any other language.
- *Natural language generation* that investigates how computer programs can be made to produce high-quality natural language text from computer-internal representations of information (Hovy, 1998).

Neural computing or *artificial neural networks* (ANNs) emulate the way that neurons work in our brains. They are based on studies of the nervous systems and brains of animals. ANNs consist of nodes and connections.

ANNs is a sub-field of *machine learning*. Machine learning encompasses AI mechanisms that allow a computer to identify patterns in example data that is important for modelling a problem and therefore learn from past experience and examples. The patterns identified can be used for monitoring, making predictions, classifying problems and providing decision support. We discuss ANNs further in Section 7.5. Other machine learning methods are *data mining* (see Section 7.6), *case-based reasoning, inductive learning, genetic algorithms and statistical methods*.

Robotics encompasses methods for controlling the behaviour of a robot. This involves:

- *Mechanical motion* that allows the robot to move. This requires knowledge of statics and dynamics to control the robot's movement (Tracy and Bouthoorn, 1997).
- *Sensory systems* that give machines the ability to sense their environment. Combined with mechanical motion, sensory systems allow robots to undertake repetitive or hazardous activities.
- *Vision and pattern recognition* that allows a robot to interpret patterns by processing visual information. A sensor such as a camera usually collects this information.

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

- *Planning* that allows a robot to devise a plan i.e. a sequence of actions to achieve a goal, often within a limited period of time.

Theorem proving focuses on proving mathematical theorems. It strives to make computers reason automatically and draw conclusions from already known facts. Most AI systems possess the ability of reasoning which allows them to deduce facts even in the face of incomplete and erroneous information. One of the disadvantages of current theorem provers is their slowness (Tracy and Bouthoorn, 1997).

Computer games. Nowadays, artificial intelligence is considered to be one of the main components of computer games. Artificial intelligence can be used to control the behaviour of the game opponents e.g. soldiers, aliens, tanks, armies and monsters. In more sophisticated computer games, artificial intelligence techniques are used to give characters beliefs, intentions and desires and make them learn from past experience.

7.3 Expert Systems

Introduction

Expertise is the knowledge that is necessary to efficiently perform a task or solve a problem. It encompasses domain knowledge, information about particular tasks, heuristic rules that provide easy ways to solve a problem and meta-knowledge i.e. knowledge about knowledge. Expertise can be gained from training, reading and experience.

The main components of an expert system are outlined below:

- A *knowledge base* which is the most important element of an expert system. This is where knowledge concepts and relationships related to a problem are stored.
- An *inference engine* that provides problem-solving skills to a system by determining how and when to apply appropriate knowledge. It uses inference mechanisms that are based on techniques such as semantic networks or procedural code to represent knowledge.
- A *user interface* to communicate with the user. Special care must be taken to ensure that it is effective. Cognitive forms can be used to interact with the DMs and make the interface easy and natural. Results are often displayed in graphical forms.

Other components can be: *explanation systems* that justify the reasoning of the expert system and *refining systems* that evolve the knowledge representations encoded in the expert system.

The aim of developing an expert system is to transfer expertise from one or more *experts* to a computer system and then to untrained *users*. A *knowledge engineer* interacts with one or more experts to build the (explicit) knowledge base of the expert system. She therefore facilitates the transfer of expertise from the expert(s) to the system.

Other participants in the development of expert systems are: a *system builder* who builds the expert system using various tools and languages, a *tool builder* who develops the necessary tools, a *vendor* who sells the expert system development products and provides advice and *support staff* who provide clerical and technical support.

There are two main expert system development strategies: develop in-house or outsource. In house development is usually preferred in organisations that have the necessary skills and resources or when an application contains sensitive data. Outsourcing can be performed by hiring a consultancy firm, joining a consortium or getting a partner e.g. a university.

Expert system shells are often used to develop expert systems. They include all the major expert system components e.g. user interface and inference engine but not the knowledge base. Modern expert system shells provide a rule set builder to help users construct rules.

More details about the development of expert systems can be found in Turban and Aronson (2001) and Waterman (1986).

Knowledge engineering and acquisition

A list of definitions is given below- the definitions have been modified from Turban and Aronson (2001):

Knowledge acquisition is the elicitation, transfer and transformation of problem-solving expertise from experts or documented knowledge sources to a computer program. Sources of expertise include not only human experts but also textbooks, reports, manuals, information available from the World Wide Web and multimedia documents.

Knowledge representation is the representation of expertise, facts and other information in knowledge representation schemes e.g. frames, rules and semantic networks.

Knowledge inferencing is the manipulation of the data structures contained in the knowledge base of an expert system using search and pattern-matching methods to draw conclusions, answer questions and perform intelligent tasks.

Knowledge transfer involves the transfer of expertise from the expert system to the user.

Again note that knowledge must be – or become – explicit if it is to be acquired, represented and transferred.

Search space

In AI, the search space is the set of all potential solutions to a problem. It corresponds to the action space in decision theory and the space of decision variables in OR. The main aim of many AI tools is to search for a solution. There are some problems in which it might be difficult to identify one solution and other problems where the search space is very large. AI techniques such as *constraint satisfaction* and *tabu-search* can be used to generate all possible solutions and identify those that are *satisfactory* i.e. satisfy some predefined constraints. These techniques are often based on *heuristics*, i.e. rules of thumb. Optimisation techniques can be employed to identify the *optimal* solution.

Expert system roles

According to Turban and Aronson (2001) an expert system can assume one of four roles depending on the type of the end-user. More precisely, an expert system can be an:

- Advisor when the user is a non-expert and wants some advice.
- Instructor when the user is a student or a novice and wants to learn more about a domain or process.
- Partner when the user is a system builder and wants to change or expand the knowledge base.
- Assistant when the user is an expert who seeks a second opinion to validate her own judgement or collect more information about a domain.

Edwards *et al* (2000) identify two expert-system roles in decision-making applications:

- Advisory i.e. support and advise a DM.
- Replacement i.e. replace a DM.

In studies Edwards *et al* (2000) discovered that, as we have suggested, expert systems that replace experts are quite effective in taking operational and tactical decisions but are not so useful at the strategic level. Advisory systems cannot eliminate all the limitations of their users and their performance is user-related.

Common expert system tasks

Marakas (1999) and Turban and Aronson (2001) outline several types of tasks that expert systems undertake.

Interpretation. Interpret sensor data by making inferences e.g. speech understanding, image analysis, signal interpretation.

Prediction. Forecast based on past and present data e.g. weather forecasting, marketing and financial forecasting, demographic predictions and traffic forecasting.

Diagnosis. Identify the cause of faults and malfunctions by observing and interpreting data e.g. medical, electronic, mechanical and software diagnoses.

Prescription. Prescribe solutions for malfunctions or provide recommendations that can help correct a problem.

Planning. Devise plans and actions to achieve given goals e.g. project management, routing and product development.

Design. Configure specifications of objects to satisfy specific requirements/constraints e.g. building design and plant layout.

Monitoring. Compare observations to expected outcomes, e.g. air traffic control.

Control. Manage the behaviour of a system, i.e. analyse the current situation, make predictions, identify the causes of anticipated problems, formulate a plan to correct/improve the situation and monitor the execution of the plan.

Instruction. Diagnose, prescribe and guide user behaviour e.g. build the profile of a student, identify her weaknesses and devise a tutorial to address her specific needs..

Benefits of expert systems

There are numerous business advantages arising from introducing expert systems:

Increase efficiency. Expert systems improve efficiency in a variety of ways:

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

- Reduce downtime, work 24 hours a day, produce results faster than people, reduce losses and increase profits;
- Replace workers, reduce operating costs and operate in hazardous environments;
- Provide a variety of outputs, reduce error rates, cope with uncertainty, solve complex problems and offer feedback.

Improve effectiveness. Expert systems allow users to gain access to relevant data, tools and other integrated systems. DMs get advice and feedback, which allows them to consider a plethora of information, understand the problem and take better decisions.

Make knowledge more accessible. Experts might not be available at all times or their expertise might be scarce. Experts systems can encode the expertise of one or more experts and provide accurate and consistent recommendations at all times.

Examples of expert systems

A few examples of expert systems are outlined below:

- XCON was developed at DEC to configure computer orders. It increased computer sales and reduced configuration errors.
- Credit assessor at American Express approves decisions in less than 5 seconds.
- Drilling advisers (Elf and BP) can diagnose oil-ring faults eliminating the need for experts.
- TARA (Hanover Trust) is an intelligent assistant for foreign exchange currency traders. It examines historical data and makes predictions.
- Siggi-Plus (University of Illinois) helps students decide what courses to take.

And two less prosaic examples:

- *Eliza* was an early artificial intelligence programme that appeared in mid 1960's. It amazed people because it was able to converse in English about any subject by storing information about the subject (i.e. interviewee) in data banks and picking up speech patterns. (But it still failed the Turing test!)
- AARON is an expert system that creates original paintings using AI techniques. Its creator, Harold Cohen, spent nearly 30 years to create it. AARON generates its drawings autonomously. It takes all the decisions e.g. how to draw lines, what colour to apply etc. It is not possible for a human to intervene and change the drawings as they emerge.

Limitations of expert systems

Several limitations of expert systems are outlined below:

- Domain knowledge might not be available and human expertise may be difficult to extract, e.g. because the tacit knowledge of the expert is essential.
- Expert systems provide support in narrow domains only.
- Expert systems, unlike human experts, lack common sense and instincts when solving a problem.
- Expert systems cannot sense their environment.
- Experts adapt to new environments and adjust to new situations whereas expert systems need to be updated.
- Systems cannot communicate as effectively as humans. Therefore, users might not trust the advice of expert systems and dismiss their results.
- An expert system gives advice no better than the expert whose expertise was transferred to the expert system.

7.4 Intelligence in Decision Support Systems

Intelligent DSSs are interactive computer-based systems that use data, expert knowledge and models to help DMs in organisations solve semi-structured problems by incorporating AI techniques (Sarma, 1994). Intelligent DSSs differ from expert systems; the emphasis is on enhancing the DMs' capabilities by focusing on their strengths while compensating for their weaknesses (Sarma, 1994).

Silverman (1995) reviews intelligent decision systems that combine mathematical modelling with ESs. These systems use a variety of techniques such as influence diagrams, Bayesian networks, risk and reliability analysis, knowledge-based systems for forecasting and model-based reasoning in decision making. Intelligent decision systems have been applied in a variety of applications (Blair *et al*, 1997; Silverman 1995) such as design, forecasting, risk management, operations planning, network routing, legal reasoning and estimation of software-development work-effort.

Goul *et al* (1992) argue that AI can increase the impact of DSSs on organisations by incorporating machine-based expertise. After reviewing the literature, they make three observations. Firstly, the focus of research on DSSs has shifted from highlighting the differences between the disciplines of AI and DSSs to promoting a synergy. Secondly, a user interacting with a ES-based DSS is automatically placed in a group decision setting with machine-based decision counterparts. Thirdly, adding knowledge into organisational DSSs has the potential to eliminate bureaucratic procedures by giving the personnel a wider access to organisational knowledge. This reduces the personnel's reaction times while lowering the cost and simplifying their interactions.

Holinagel (1987) stresses that the main purpose of an intelligent DSS should be to improve the quality of the information conveyed to the user. This involves determining the meaning and content of the information (what), the framing of the information (how) and the timing of displaying information as well as whether this display should be user-driven or automatic (when). An intelligent DSS can take three different roles depending on the type of information provided:

- A constant guard preventing any fallacies (Reason, 1987) occurring when the users have to deal with incomplete information.
- An intelligent assistant that anticipates what the needs of the users are and carries out all the necessary computations when there is insufficient information.
- An information filter that removes any redundant or superfluous data when information overflow occurs.

Turban (1993) and Turban and Aronson (2001) study the collaboration between DSSs, which consist of data, model and dialogue modules, and expert systems and identify three approaches to designing intelligent DSSs.

The first approach is to develop a multi-purpose expert system that supports several aspects of the decision making process such as problem formulation and criteria modelling and make it a central component of the DSS.

The second approach is to develop an expert system as a separate DSS component. The expert system can be used to integrate the database and a model base in an intelligent manner, provide input to the DSS (e.g. determine the most important factors to consider, classify the problem, generate alternatives), interpret the results of the DSS (when it is faster and cheaper to obtain explanations from an expert system than from a human expert or when the quality of the explanations provided is superior).

The third approach is to develop several expert systems to support the functionalities of the DSS modules and the interaction between the DSS and the DM. In this architecture an expert system can undertake one of the following roles (Turban and Aronson, 2001):

- Database intelligent component
- Intelligent agent for managing the model base
- System for improving and customising the user interface
- Advisor/consultant to the users (for example inform the users about the decision context and the feasibility of the alternatives under consideration as well as how to use the DSS and how to interpret its results)
- Advisor/consultant to the DSS developers (for example give advice on how to structure the DSS and assemble the various part together)

7.5 Neural networks

ANNs are well suited to pattern recognition, classification and prediction problems. They have been applied successfully to many applications such as risk evaluation for mortgage application, fraud detection in insurance claims and sales forecasting. The output of ANNs is indicative but not always accurate. It depends on the network's structure, the node computations and the weights attached to the links. These weights represent the importance given to the input data and training is required to adjust their values using already known examples. ANNs have been criticised mainly because they cannot justify their reasoning and require frequent training and large quantities of test data.

As we noted, ANNs consist of nodes (or process elements or neurons) and connections. The nodes are grouped in layers and may have multiple input and output connections. There

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

are three types of layers: input, intermediate (or hidden) and output. There might be several hidden layers (usually no more than three) between the input and output layers.

Each input node corresponds to an attribute or characteristic that may be 'sensed'. We can have different types of input e.g. data, voice, picture. In some cases we might have to process the input data and convert it to a meaningful form. Any time the input connection of a node is stimulated, a computation is performed which produces an output or 'fire'. Connections transfer data from one layer to another. Each connection carries a weight that expresses the importance given to the input data (i.e. the relative importance of each input to another node). It essentially indicates how much an input node that represents an attribute contributes to an output. The resulting pattern of states in the output layer nodes contains the solution to a problem. For example, in a loan approval example the answer can be 'yes' or 'no'. The ANN assigns a numeric value e.g. '1' for yes and '0' for no.

The output of ANNs is indicative but not always accurate. It depends on the network's structure, the node computations and the weights attached to the links/connections. Training is required to adjust the weight values using already known examples.

ANNs can be used in a variety of business applications. For a complete list see Turban and Aronson (2001; section 16.2). Here we describe just two cases.

Example: Bankruptcy prediction

Turban and Aronson (2001) give an example of a neural network that uses financial ratios to predict bankruptcy. It is a three-layer network with five input nodes that correspond to the following well-known financial ratios:

- Working capital/ total assets
- Retained earnings/ total assets
- Earnings before interest and taxes/ total assets
- Market value/ total debt
- Sales/ total assets

A single output node classifies a given firm and indicates a potential bankruptcy (0) or nonbankruptcy (1) based on the input financial ratios of the firm. The data source consists of financial ratios calculated for firms (129 in total) that did or did not go bankrupt between 1975 and 1982. The data set was divided into a training set (74 firms; 38 bankrupt, 36 not) and a testing set (55 firms; 27 bankrupt, 28 not).

The neural network accurately predicted 81.5 % of the bankrupt cases and 82.1% of the non-bankrupt cases. An accuracy of about 80% is usually acceptable for applications of neural networks. The performance of a neural network should be compared against the accuracy of other methods and the impact of an erroneous prediction.

Example: Loan Repayment.

Braincell is a neural network. It can be embedded in Excel as another menu. The Braincell Excel add-in and loan repayment demo can be downloaded (free) from www.promland.com.

The process of developing an ANN e.g. credit authoriser is the following (Turban and Aronson, 2001):

Step 1: Collect data from past loan applications e.g. applicant's monthly income and expenses.

Step 2: Separate data into training and test sets.

Step 3: Transform data into network inputs. For example the 'Home owner' entries are converted to '1' if the applicant is a home owner and '0' otherwise.

Step 4: Select the right network configuration (this impacts on the network's performance and influences the accuracy of its results), train the network until the error in calculating the output in the training data (i.e. the difference between an output and the output calculated by the neural network) has reached a given level (e.g. 5%- i.e. the error is less than 5%) and test it to check the predictions of the network in the new test cases.

Step 5: Deploy the network i.e. integrate it into the credit approval process and use a friendly user interface to make enquiries and see the results.

7.6 Data mining

Whether the topic of data mining, which has its roots within database design and statistics, should be within a chapter on AI is something we might debate, but we shall not. Whatever its past, its present involves many AI algorithm to dig patterns out of large datasets.

The growth of information systems and their penetration into all workings of organisations and society that nowadays DMs have vast

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

quantities of data available to help them in their monitoring performance and supporting their decision making. At least in principle they have; in practice extracting the appropriate data and organising them informatively may be very difficult.

Organisations seldom have a single database: they have tens or hundreds, each built to handle the data relevant to one task or activity. For many organisations, their history and geographical dispersion may mean that many of their key databases exist in a variety of database management systems distributed across the globe. The value of the potential information shared between these databases can be immense to a company, e.g., in comparing customer buying patterns in different countries. Gathering and working with these is difficult but not impossible. A *data warehouse* is a system in which all the organisations databases are brought together and archived in one place together with the software tools to enable detailed querying and report generation from all the data present. Data warehouses seldom archive all the data that has passed through an organisation, but rather snapshots of the data at intervals sufficiently short to capture a useful history of its processes. For a supermarket this might be daily; whereas for a small foundry it might be weekly or monthly. Originally data warehouses were single systems, essentially located in one physical place on one physical system, but technological advances mean that they may be distributed across many systems: indeed, they may soon be virtual, accessing summaries 'on the fly' from full archives of the original databases.

Using the snapshots of its history contained in a warehouse – often together with current data – an organisation may explore past patterns of behaviour and use these to underpin its decision making. Among the tools available to help management explore and extract information from the vast volume of data held in data warehouses are a family of techniques for recognising possible patterns in the data known as *data mining*.

In a very real sense data mining is no more than exploratory statistical analysis. Most of the data mining techniques currently being used have a long pedigree within the discipline of statistics – although the developers of the algorithms may not know this. Many wheels have been re-invented within AI, not all of them round! But subject ownership is relatively unimportant; what matters is what benefits data mining techniques bring and these can be substantial.

Data mining has allowed organisations, *inter alia*, to:

- increase profits and efficiency by exploiting general trends across their sub-units;
- increase profits and efficiency by exploiting local differences between their sub-units;
- personalise their dealings with clients and customers so building stronger relationships.

Data mining techniques are many and varied. Firstly, there are the long established methods of statistical analysis (see, e.g., Mignon and Gamerman, 1999), such as multivariate analysis (Krzanowski and Marriott, 1995,1995), regression analysis (Gelman *et al*, 1995), time series analysis and forecasting (West and Harrison, 1989). These are best suited to finding global or near global patterns across the data: trends, common consumer behaviour, etc. While they may draw on established statistical models, such methods may, nonetheless, use modern and fast AI algorithms to fit the models and extract the patterns. Then there are series of new methods which find local patterns which only hold for small subsets of the data – in statistical terms, conditional correlations and dependencies. A promising research area in this respect is the automatic constructions of Bayesian belief nets by exploiting the empirical correlations in very large databases.

Related to data mining systems are *executive information systems* (EIS). These are computerised systems which "provide executives with easy access to internal and external information that is relevant to their critical success factors" (Watson *et al*, 1997). EISs are essentially very clever querying and reporting software with a very intuitive human computer interface (HCI) that senior managers can use to query, draw together and summarise data from many databases. In that they are similar to data mining tools in data warehouses, but usually they extract the data without the need for the intermediary of a data warehouse and also they are concerned with much more high level, broad brush exploration of an organisation's data.

EISs allow senior management to continually review the progress of their company against its planned strategy. The information

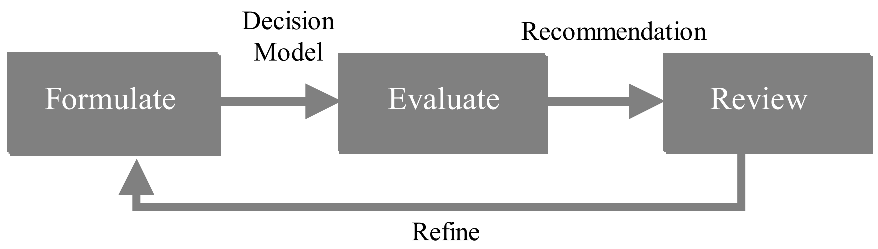


Figure 48: The main stages in the decision analysis process

displayed to them in response to their requests is a high level summary of some aspects of their company's performance drawn from databases and other information, often stored in locations distributed across the world. The summaries are typically displayed using simple graphics and other intuitive plots and tabulations. Through such monitoring of the 'big picture', DMs are protected – somewhat – against defensive avoidance.

8 Decision Support Systems

8.1 Introduction

Our aim in this section is to draw together all that has gone before and discuss how DSSs implement the various decision support processes and models that we have considered.

In Figure 39 we presented an overview of the decision analytic process, the key stages of which we indicate in Figure 48.

Decision analysis can be seen as a consultation process that attempts to focus a DM's attention on the important aspects of a decision problem. As with any other consultation, it starts with the definition of a decision problem and it ends with a DMs' commitment to a real action (Regan and Holtzman, 1995). In order to help the DMs gain insight into the decision problem and clarify their preferences, guidance is given in three stages (Holtzman, 1989):

- *Formulation* of the decision model that reflects the decision problem. This involves generating alternatives and identifying evaluation criteria.
- *Evaluation* of the decision model. This involves computing the implications of the decision model, evaluating it using a formal decision method and producing a recommendation.
- *Review* of the recommendation. This involves analysing the recommendation and presenting the interpretation in a natural language form.

A feedback or refinement path is provided to give the opportunity to the DMs to re-evaluate the decision model or modify its formulation. The decision model is progressively refined until the DMs are confident that the components, structure and values of the decision model accurately represent the decision problem (McGovern *et al*, 1994). (Philips, 1982) argues that these final decision

Domain of Activity	Characteristics of decision making	Examples of decision support tools and systems
Corporate strategic	Very long time-spans of discretion; highly unstructured; need to evolve strategic direction.	Executive information systems; problem formulation tools (soft-OR); decision analysis.
General	Medium to long time-spans of discretion; relatively unstructured tactical issues, although some structure provide by strategic direction; need to articulate long term values in the form of short term goals.	Data mining; problem formulation tools (soft-OR); decision analysis, operational research techniques.
Operational	Planning for implementation of tactics and strategy provided by higher domains; allocation of resources to meet given objectives.	Data mining; operational research techniques; expert systems; neural nets; general KB-DSS systems.
Hands-on work	Repetitive, routine decision making within a constant format. Values provided by higher level.	Rule-based expert systems and simple heuristic techniques; general KB-DSS systems.

Table 10: Decision support tools appropriate to different domains.

models are requisite because they are detailed enough to allow the DMs make a decision without considerable effort.

These stages hold in supporting decisions in all domains, although with different emphases. For instance, for unstructured decisions in the corporate domain, the all tree stage of the process should get considerable emphasis. The issues will need formulating, careful modelling and analysis to evaluate the alternatives, and then support for reflective discussion as the decision is reviewed. On the other hand, for instance, some scheduling problems in the operational domain will be so standard that little formulation is need. The data can simply be input into a standard OR or AI scheduling package, the optimisation run and the schedule implemented with little, if any review.

Table 10 summarises the characteristics of decision making in the various domains of management activity: see also Figure 7, page 13, which categorises DSS tools by domain of activity and level.

8.2 DSS in the various domains

DSS for the corporate domain

The corporate strategic domain covers very long term decision making, often extending a several years into the future. DMs operating at this level face problems of shaping society or their organisation. It also covers major projects which will take years to come to fruition: e.g. the construction of a major bridge or, more exotically, the development of a colony on the moon. In a very real sense, they may need to be visionary in their thinking.

The Level 0 tools that they require are essentially highly summarised tabulations and plots of evolving trends. Many *executive information systems* (EIS) can provide these. EISs allow senior management to continually review the progress of their company against its planned strategy. The information displayed to them in response to their requests is a high level summary of some aspects of their company's performance drawn from databases and other information, often stored in locations distributed across the world. Through such monitoring of the 'big picture', DMs are protected – somewhat – against defensive avoidance.

Rich picture diagrams, soft system modelling, cognitive mapping and similar 'soft' methods also provide Level 0 support (see Checkland and Howell, 1998; Rosenhead, 1989. They organise and encapsulate the DMs' perceptions of the external and internal environments. If they develop qualitative forecasts of how systems may develop, as, for instance, in scenario planning, then they also provide Level 1 and 2 support. Soft techniques are implemented in a number of software systems, although the value of flipcharts and whiteboards should not be overlooked..

More quantitative examples of Level 1 and 2 support may be found in statistical forecasting systems. There are many statistical software packages available, many of which have technologies to link them to large corporate data bases. Also general modelling tools may be used. For instance, large spreadsheet models may be used to predict the effects of different financing policies on the profitability of different options.

Generally the level 3 support is provided by the sorts of decision analytic techniques discussed in Chapter 4, particularly MAVT methods. There are many software implementations of these, all characterised by easy to use, intuitive interfaces which allow the results to be studied and explored in senior management meetings and boardrooms.

DSS for the general domain

Decision making in this domain has time-spans of discretion of perhaps one to two years. It is at this level that the broad strategic intent and mission is interpreted into a strategy with shorter term goals.

In some ways the decision support needed in this domain differs little from that required in the corporate strategic domain. The differences are more of degree than of qualitatively different DSS tools, particularly at Levels 0, 1, and 2. Thus EIS, soft OR, statistical forecasting and general modelling will all be applied, perhaps with some general OR tools. The application will differ in detail, type and volume of data used, but the methodologies will be much the same. At Level 3, the decision models may require more detailed data and judgements and their structure may be more complex, but the methodologies and software used will be largely similar, if not the same.

One difference may be that less effort is put into problem formulation (Levels 0, 1, 2) and more into evaluation models (Level 3) simply because, although not routine, the issues encountered will be similar to those that have arisen in the past or elsewhere. Moreover, an outline structure of the decision model may be cascaded down from analyses in the corporate strategic domain, with instructions to ‘fill in the details’.

DSS for the operational domain

The operational domain has time-spans of discretion extending from a few months to a year or so. Here the plans developed in the general domain of management are fleshed out with the details necessary to implement them. Staff and resources are allocated, activities scheduled, logistics arranged and so on. This tasks are the very ones that classical OR methods were designed to support. It was also in this domain that information systems first made their impact on management.

Management information systems (MIS) were developed to provide the information that management needed in running the business: i.e. implementing and monitoring strategy. In our terminology, MIS provide Level 0 support. In general terms one might argue that the difference between MIS and EIS is more in terms of sales-pitch than in the function it performs. Both venture into databases, extract and organise data, presenting it in ways that support their uses understanding of the current position. But they differ in degree: EIS need to draw data from a much wider range of data sources distributed over the organisation; MIS draw on data much closer to the manager, usually from fewer databases. The summaries presented by EIS are much less detailed and much more broad brush than those produced by MIS. Managers operating in this domain need to know that they are five widgets short, that there is a worrying delay on production line four, that of the seventy patients who have had their treatment delayed by more than a year forty six were waiting for ENT surgery, or whatever. Managers need more detail than executives.

Classical OR methods (see Chapter 6) will provide much of Level 2 and 3 support. Linear and other mathematical programming methods, together with more modern optimisation techniques such as genetic algorithms and tabu search, offer guidance both on what the different alternatives may lead to and to which is ‘best’ under certain circumstances. Other OR models such as inventory, maintenance and reliability, or scheduling provide similar guidance for other contexts. Sometimes, as in the case of simulation, the support is limited to Level 2 rather than Level 3. Knowledge based expert systems with their ability to learn to cope in relatively repetitive situations are also useful: see Marakas (1999).

It is also possible in this domain to apply any of the decision support tools mentioned for the corporate strategic or the general domain, but often the effort of tailoring such tools to these circumstances and eliciting judgements to populate them with parameters is disproportionate to the benefit gained.

DSS for the hands-on domain

When we get to short time-spans of discretion, the Level 0 DSS essentially monitor the current activity and success in meeting targets. They provide up to date, short term, factual information. Higher level decision support may be provided by heuristic methods or rule based expert systems, tools which may be tuned to repetitive contexts and left to run semi-automatically. At this level it might be thought that decision support did not help DMs to form and explore the implications of their judgements and hence to make a decision based upon understanding. It is all a matter of degree, however. A bank official assessing a customer's credit worthiness for a mortgage may be guided by an expert system, but he will also pay attention to other information that does not 'fit' the systems template and, if he thinks it is appropriate from his knowledge of how the system has functioned in the past, either over-ride the advice or seek guidance from a superior. Decisions will still be made through understanding.

8.3 Human Computer Interfaces

We have emphasised that prescriptive decision support requires that we guide DMs towards the ideals of behaviour encoded in normative models, mindful of there cognitive capabilities. Clearly the models embedded in DSSs encode the normative ideals; their calculations ensure that their output is compatible with the canons of rationality which the DMs would like to adopt. The decision analytic process as we have described it ensures a reflective use of these software tools which is sympathetic to the DMs' cognitive capabilities. In addition, the design of a DSS's interface should recognise potential lacunae and biases in the DMs' perceptions.

For instance, the input screens of a DSS in eliciting the DMs'. experts' and stakeholders' judgements should:

- check for and guide the users away from biases, such as
  - anchoring,
  - availability,
  - over-confidence,
  - loss-gain framing effects;
- ask questions that are cognitively meaningful to the users using, if possible, *their* language – e.g., it is inappropriate to use technical language such as decision theoretic, statistical or economic concepts in seeking judgemental values;
- ask about real observable quantities, not modelling constructs so that the users can relate their responses to their experience;
- ask for cognitively possible precision; judgement can only be made to an 'accuracy' of 1% or 2% at best;
- use sensitivity analysis techniques to focus the users' efforts and reassure them that greater accuracy is not necessary;
- recognise that the unfamiliarity of the tool and judgemental effort required may exhaust the users.

Equally the output screens which will guide the DMs choice of action should:

- use the DMs' language;
- use the DMs' metaphors, not those of the analysts or experts;
- beware of misleading the DMs by poor framing of the output;
- not over-clutter the screen – too much detail can hide or distract from the key points;
- give numeric output to appropriate and useable precision;
- watch wording for undesired associations, e.g. fright factors;
- watch colour for associations, e.g. red danger.

In addition, of course, the system should adopt the best HCI practice for all information systems as described in, e.g., Laudon and Laudon (2001).

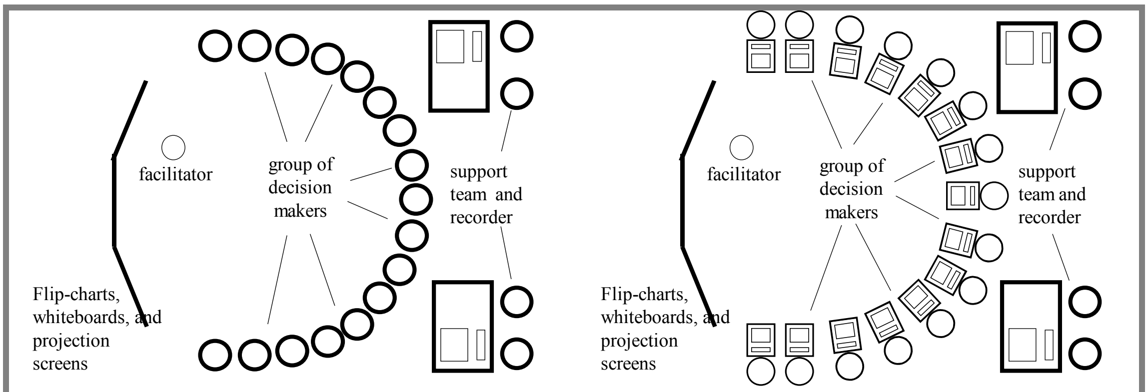
8.4 Group Decision Support Systems

All our comments on DSS need shading for their design and use within a group. Computer supported co-operative work (CSCW) or groupware are used as generic terms for software designed for such use.

Groupware may be distinguished by whether the group 'meets'

- at the same time in the same place;
- at the same time, but in different places;
- at different times.

We have implicitly been assuming that decision support is provided to the group meeting together in one location and at the same time. But modern information and communication technologies mean that



(a) Group decision support arrangement for plenary projection of common model

(b) Group decision support in which each participant has access to own decision model

Figure 49: Group decision support environments

assumption is unnecessary. With video conferencing or web based 'net meetings' they may meet together while being physically located at different places. Moreover, it is relatively trivial to allow them to interact with the same decision analytic model through distributed software. Video techniques mean that members of the group may see each other's reactions at the same time as hearing their comments. It is possible to use large scale monitors or wall projection to create the impression that the group are 'in the same room'. It is also possible to meet in virtual space, although I am unclear how that helps decision making (except in the context of game playing!).

Equally technology can record members comments and interactions with a model for transmission to other members at a later time; and, in turn, record their reactions to transmit to the former members at a yet later time. Whether this 'stop-go' meeting takes place at the same place or in different locations is essentially irrelevant. Obviously such interactions will not be as effective as same time interactions but, if a simultaneous meeting is impossible, then the technology does provide an improvement over postal communication.

Whatever the temporal and spatial relationship of the participants, one advantage – or disadvantage? – of groupware is that some interactions may be anonymous. In brainstorming, general discussion, the elicitation of judgements and so on, it is often hard for some members to dissociate the status of the member offering a comment from its content. If the interaction take place via a keyboard and screen, it is possible to enforce anonymity. Whether such anonymity is valuable is a moot point (see, e.g., Cooke, 1991). But the potential to ensure anonymity *when* it is appropriate is one of the strengths of groupware. Nunamaker and his co-workers make much use of this in their software systems (Nunamaker *et al*, 1988).

A more worrying concern with some groupware is that it automates the evaluation stage, for instance, by providing group members with a simple voting scheme. If the group are aware of all the paradoxes that may possibly arise from voting, then perhaps there is no harm in this. But it seems to me wiser to include an element of facilitation in order to help the group reach a consensus based upon a shared understanding.

For the corporate strategic and general domains, decision support is often provided via decision conferences or facilitated workshops, which may take place in purpose built group decision support rooms.

8.5 Group Decision Support Environments

Figure 49 shows two layouts for group decision support rooms. In Figure 49 (a) the participants sit in a semi-circular arc. This arrangement means that they can se each other during discussions and also a screen, flip-charts and whiteboards at the end of the room. The facilitator would operated in the space in the middle, sometimes sitting down when he does not wish to dominate the proceedings. At the end of the room the support team and recorder sit. In such room all the discussion would be conducted in open plenary form. The soft OR and decision models are built in front of the whole group, either with pen on the

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

whiteboards and flip-charts or projected on the screen from the support team's computer(s). Whatever the case, model building is done by the whole group with constant plenary discussion.

In Figure 49 (b) the set up is different in that each member of the group has his or her own computer, in addition to those used by the support team. All the computers are networked together. This allows that each group member can either build his or her own personal decision model which captures his or her perceptions or that they can work jointly on a communal model.

Although the differences between the two formats may seem less then their similarities, the two environments represent very different approaches to group decision support. Facilitators who favour the arrangement in Figure 49 (a) feel that it vitally important to maintain group discussion at plenary level, building a common model in their midst. There are no machines in front of people so that each can see the others' body language and maintain eye contact. All this builds shared understanding and a common commitment to the developing strategy. On the other hand, facilitators who prefer the arrangement in Figure 49 (b) feel that much is gained by allowing periods in which the individual members can explore the models and their perceptions on their own to understand things for themselves (Phillips and Phillips, 1993; Nunamaker *et al* 1991). At present, although the debate is strong between the proponents of these two approaches, there is little empirical evidence to suggest whether one approach or the other supports DMs better. Both arrangements have been used with success to support many groups clients who subsequently reported themselves well satisfied with the arrangement and methods used.

On point in common between the two views is that there is a distinct advantage in a semicircular arrangement with a space in the middle. Apart from allowing the facilitator to move around among the group, the space in the middle seems more conducive to calmer non-confrontational meetings than sitting around a large table.

The schematic arrangements in Figure 49 (a) and (b) are no more than schematic. Real arrangement differ in many details. I have been in group decision support rooms with oak tables, oil paintings and plush carpets "to make boards of directors feel at home", high tech 'pods', rooms that felt like and possibly doubled as teaching laboratories and, most often, hotel conference suites into which portable computers or even portable computer networks have been installed. See Hickling (1990) for a wide discussion of decision support rooms.

References

F. Ackermann (1996) 'Participants' perceptions on the role of facilitators using group decision support systems' *Group Decision and Negotiation.* **5**, 93-112.

M. Allais and O. Hagen (Eds.) (1979) *Expected Utility Hypothesis and the Allais Paradox.* D Reidel, Dordrecht.

H.R. Arkes and K.R. Hammond (Eds) (1986) *Judgement and Decision Making: an interdisciplinary Reader*. Cambridge University Press.

K.J. Arrow (1951) *Social Choice and Individual Values*. 1st Edn. John Wiley and Sons, New York. (The 2nd edition published in 1963 contains important corrections.)

M. Bacharach (1975) 'Group decisions in the face of differences of opinion' *Management Science* **22**, 182-191.

M. Bacharach and S. Hurley (Eds.) (1991) *Foundations of Decision Theory.* Basil Blackwell, Oxford.

V. Barnett (1999) *Comparative Statistical Inference*. John Wiley & Sons, Chichester. Third Edition.

R. Baron, N. Kerr and N. Miller (1992) *Group process, group decision, group action*. Open University Press.

M.H. Bazerman (2002) *Managerial decision making*. 5th Edn. John Wiley and Sons, Chichester.

T. Bedford and R.M. Cooke (2001) *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge University Press

D.E. Bell, H. Raiffa, and A. Tversky (Eds.) (1988) *Decision Making.* Cambridge University Press.

V. Belton (1990) 'Multi-criteria decision analysis: practically the only way to choose'. In L. Hendry and G. Rand (Eds.) *Operational Research Tutorial Papers 1990.* Operational Research Society, Seymour House, 12 Edward St., Birmingham, B1 2RX.

V.Belton, F. Ackermann and I. Shepherd (1997) 'Integrated support from problem structuring through to alternative evaluation using COPE and VISA' *Journal of Multi-Criteria Decision Analysis* **6**, 115-130.

V. Belton and T.J. Stewart (2002) *Multiple Criteria Decision Analysis*: *an Integrated Approach.* Kluwer Academic Publishers, Dordrecht.

P.G. Bennett and K Calman (Eds) (1999) *Risk Communication and Public Health*. Oxford University Press.

P.G. Bennett, S. French, J. Maule, D.G. Coles and A. McDonald (1999): "Improving Risk Communication: a Programme of Work in the Department of Health" *Risk Decision and Policy* **4**, 47-56.

J.M. Berrnardo and A.F.M. Smith (1994) *Bayesian Theory*. John Wiley and Sons, Chichester.

R. Boaden and G. Lockett (1991) 'Information technology, information systems and information management: definition and development.' *European J. of Information Systems*. **1**, 23-32.

M.H. Boisot (1998) *Knowledge Assets*: *Securing Competitive Advantage in the Information Economy*. Oxford University Press.

D. Bouyssou, T. marchant, M. Pirlot, P. Perny. A. Tsoukias and P. Vincke (2000) *Evaluation and Decision Models: a Critical Perspective*. Kluwer, Dordrecht.

G.J. Browne, S.P. Curley and P.G. Benson (1997) 'Evoking information in probability assessment: knowledge maps and reasoning-based directed questions' *Management. Science.* **43**, 1-14.

S. Brownlow and S.R. Watson (1987) 'Structuring multi-attribute value hierarchies' *J. Operational Research Society* **38**, 309-317.

P. Checkland (1989) 'The stages of soft systems methodology'. In Rosenhead (1989), 71-100.

P. Checkland and S. Holwell (1997) *Information, Systems and Information Systems: making sense of the field*. John Wiley and Sons, Chichester.

R.T. Clemen (1996) *Making Hard Decisions*. 2nd Ed. PWS-Kent, Boston.

R.T. Clemen and R.L. Winkler (1999) 'Combining probability distributions from experts in risk analysis' *Risk Analysis*. **19**, ???-???.

R.M. Cooke (1991) *Experts in Uncertainty*. Oxford University Press.

R.M. Cooke, L.H. Goossens and B.C.P Kraan (1995) ???

H.G. Daellenbach (1994) *Systems and Decision Making*. John Wiley and Sons, Chichester.

B. De Finetti (1974) *Theory of Probability*. Volume 1. John Wiley & Sons. Chichester.

B. De Finetti (1975) *Theory of Probability*. Volume 2. John Wiley & Sons. Chichester.

M.H. DeGroot (1970) *Optimal Statistical Decisions*. McGraw-Hill, New York.

E.V. Denardo (2002) *The Science of Decision Making*. John Wiley and Sons, New York.

J.S. Dyer and R.K. Sarin (1979a) 'Measureable multi-attribute value functions' *Operations Research* **28**, 810-822.

J.S. Dyer and R.K. Sarin (1979b) 'Group preference aggregation rules based upon strength of preference' *Management Science* **25**, 822-832.

M.J. Earl (2000) 'Every business is an information business'. In Marchand *et al* (2000) 16-22.

C. Eden and J. Radford (Eds.) (1990) *Tackling Strategic Problems: the Role of Group Decision Support.* Sage, London.

C. Eden and F. Ackermann (1998) *Making Strategy: the Journey of Strategic Management*. Sage, London.

J.S. Edwards, Y. Duan, P.C. Robins (2000) 'An analysis of expert systems for business decision making at different levels and in different roles' *European Journal of Information Systems*, **9**, 36-46.

W. Edwards (1954) 'The theory of decision making'. *Psychological Bulletin* **51**, 380-417.

S.B. Eom, S.M. Lee, E.B. Kim and C. Somarajan (1998) 'A survey of decision support system applications (1998-1994)' *J. Operational Research Society* **49**, 109-120.

T.L. Fine (1973) *Theories of Probability*. Academic Press, New York.

P.C. Fishburn (1972) *The Mathematics of Decision Theory*. Mouton.

S. French (1984) "Interactive multi-objective programming: its aims, applications and demands" *J. Operational Research Society*. **35**, 827-834.

S. French (1985) Group consensus probability distributions: a critical survey. In Bernardo, J.M., DeGroot, M.H., Lindley, D.V. and Smith, A.F.M. (Eds) *Bayesian Statistics 2*, 183-201.

S. French (1986) *Decision Theory: an Introduction to the Mathematics of Rationality*. Ellis Horwood, Chichester.

S. French (Ed.) (1988) *Readings in Decision Analysis*. Chapman and Hall, London.

S. French (1992b). "The use of decision conferencing to determine intervention levels and other counter measures after a nuclear accident". In Proceedings of the International Seminar on Intervention Levels and Countermeasures for Nuclear Accidents Cadarache October 1991 CEC (EUR14469) 586-594.

S. French (1995) "Uncertainty and Imprecision: Modelling and Analysis" *Journal of the Operational Research Society*. **46**, 70-79.

S. French (1996) "Multi-attribute decision support in the event of a nuclear accident." *Journal of Multi-Criteria Decision Analysis*. **5**, 39-57.

S. French (1997) 'Uncertainty Modelling, Data Assimilation and Decision Support for Management of Off-site Nuclear Emergencies' *Radiation Protection Dosimetry* **73**, 11-15.

S.French (2001) "Modelling, making inferences and making decisions: the roles of sensitivity analysis". Submitted to *TOP*.

S. French, J. Bartzis, J. Ehrhardt, J. Lochard, M. Morrey, N. Papamichail, K. Sinkko and A. Sohier (2000) RODOS: Decision support for nuclear emergencies. In S.H. Zanakis, G. Doukidis and C. Zopounidis (Eds). *Recent Developments and Applications in Decision Making*, Kluwer (in press).

S. French, E. Halls and D.C. Ranyard (1997) 'Equity and MCDA in the event of a Nuclear Accident'. In G. Fandal and T. Gal (Eds) *Multiple Criteria Decision Making*. LNEMS 448, Springer Verlag, Berlin. 612-621.

S. French, M.T. Harrison and D.C. Ranyard (1997) 'Event conditional attribute modelling in decision making on a threatening nuclear accident.' In French and Smith (1997) 131-150.

S. French, R. Hartley, L.C. Thomas and D.J. White (Eds.) (1983) *Multi-Objective Decision Making*. Academic Press, London.

S. French and Y. Liang (1993) Decision support systems: a decision analytic perspective. In Norman, J. (Ed.) *Developments in Operational Research 1993*. Operational Research Society, Seymour House, 12 Edward St., Birmingham, B1 2RX. 21-40.

S. French, K.N. Papamichail, D.C. Ranyard and J.Q. Smith (1998) "Design of a decision support system for use in the event of a radiation accident" F. Javier Girón and M. L. Martínez (eds) *Decision Analysis Applications* Kluwer Academic Publishers, Dordrecht. 3-18.

S. French and D. Rios Insua (1991) "A Framework for Sensitivity Analysis in Multi-Objective Decision Making. *European J. Operational Research* **54**, 176-190.

S. French and D. Rios Insua (2000) *Statistical Decision Theory*. Edward Arnold, London.

S. French, L. Simpson, E .Atherton, V. Belton, R. Dawes, W. Edwards, R.P. Hamalainen, O. Larichev, F.A. Lootsma, A.D. Pearman and C. Vlek (1998) 'Problem Formulation for Multi-Criteria Decision Analysis: Report of a Workshop' *J. Multi-Criteria Decision Analysis* **7**, 242-262.

S. French and J.Q. Smith (Eds.) (1997) *The Practice of Bayesian Analysis*. Edward Arnold, London.

S. French and M. Vassiloglou (1983) 'Examinations and multi-attribute utility theory'. In S. French *et al* (1983) 307-322.

S. French and Z. Xie (1994) A perspective on recent developments in utility theory. In Rios (1994) 15-31.

A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin (1995) *Bayesian Data Analysis*. Chapman and Hall, London.

P. Goodwin and G. Wright (1999) *Decision Analysis for Management Judgement*. 2nd Edn. John Wiley and Sons, Chichester.

A.G. Gorry and M.S. Scott-Morton (1971) 'A framework for management information systems' *Sloan Management Review* **13**, 55-70.

M. Goul, J.C. Henderson and F.M. Tonge, (1992), "The Emergence of Artificial Intelligence as a Reference Discipline for Decision Support Systems Research", *Decision Sciences*, Vol. 23, No. 6, pp. 1263-1276.

M. Hammer and J. Champy (1993) *Reengineering the Corporation* Nicholas Brearley Pub. Ltd., London.

P.J. Hammond (1976) 'Equity, Arrow's conditions and Rawls' difference principle' *Econometrica* **44**, 793-804. J.C.

J. Harsanyi (1977) *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge University Press.

A. Hickling (1990) '"Decision spaces": a scenario about designing appropriate rooms for group decision management'. In Eden and Radford (1990) 169-177.

E. Hollinagel (1987), "Information and Reasoning in Intelligent Decision Support Systems", *International Journal of Man-Machine Studies*, **27**, pp. 665-678.

S. Holtzman (1989), *Intelligent Decision Systems*, Addison-Wesley, Reading, MA.

R.J. House and J.V. Singh (1987) Organisational behaviour: some new directions for I/O psychology. *Annual Review of Psychology*. **38**, 669-718.

E. Hovy (1998), "Overview", in R.A. Cole, J. Mariani, H. Uszkoreit, G. Battista Varile, A. Zaenen, A. Zampoli and V. Zue (Eds.), *Survey of the State of the Art in Human Language Technology*, available from http://cslu.cse.ogi.edu/HLTsurvey/, chapter 4, section 4.1.

R.M. Hogarth (1987) *Judgement and Choice*. 2nd Edition. John Wiley & Sons, Chichester.

E. Jacques (1989) *Requisite Organisations*. Cason Hall, Arlington.

I.L. Janis (1972) *Victims of Groupthink*. Houghton Mifflin, Boston.

I.L. Janis and L. Mann (1977) *Decision Making: a Psychological Analysis of Conflict, Choice and Commitment*. Free Press, New York.

F.V. Jensen (1996) *An introduction to Bayesian Networks*. UCL Press, London.

F.V. Jensen (2001) *Bayesian Networks and Decision Graphs*. Springer-Verlag, New York

D. Kahneman, P. Slovic and A. Tversky (Eds.) (1982) *Judgement under Uncertainty*. Cambridge University Press.

D. Kahneman and A. Tversky (1974) 'Judgement under Uncertainty: Heuristics and Biases' *Science* **185**, 1124-1131.

D. Kahneman and A. Tversky (1979) 'Prospect theory: an analysis of decisions under risk' *Econometrica* **47**, 263-291.

P. Kanellis, M. Lycett and R.J. Paul (1999) Evaluating business information fit: from concept to practical application. *European Journal of Information Systems* **8**, 65-76.

A Karaoglou, G.Desmet, G.N.Kelly and H.G. Menzel (Eds) (1995) *The Radiological Consequences of the Chernobyl Accident*. EUR 16544 EN. CEC, Luxembourg.

R.L. Keeney (1976) 'A group preference axiomatisation with cardinal utility' *Management Science* **23**, 140-145.

R.L. Keeney (1992) *Value Focused Thinking*, Harvard University Press.

R.L. Keeney and H. Raiffa (1976) *Decisions with Multiple Objectives: Preferences and Value Trade-offs* John Wiley.

D.A. Klein (1994) *Decision-Analytic Intelligent Systems - Automated Explanation and Knowledge Acquisition*, New Jersey: Lawrence Erlbaum Associates.

M.R. Klein and L.B. Methlie (1995) *Knowledge-based Decision Support Systems*. John Wiley and Sons, Chichester.

P.R. Kleindorfer, H.C. Kunreuther and P.J. Schoemaker (1993) *Decision Sciences.* Cambridge University Press.

P. Korhonen and J. Wallenius (1996) 'Behavioural issues in MCDM: neglected research questions'. *J. Multi-Criteria Decision Analysis*, **6**, 178-182.

D.H. Krantz, R.D. Luce, P. Suppes and A. Tversky (1971) *Foundations of Measurement*. Vol 1. Academic Press, New York.

W.J. Krzanowski and F.H.C. Marriott (1994) *Multivariate Analysis Vol 1*. Edward Arnold, London.

W.J. Krzanowski and F.H.C. Marriott (1995) *Multivariate Analysis Vol 2*. Edward Arnold, London.

P.S. Laplace (1825) *Essai Philosophique sur les Probabilities*. 5th Edition, Paris. Translation published by Dover, New York (1952).

J.W. Lathrop and S.R. Watson (1982) 'Decision analysis for the evaluation of risk in nuclear waste management' *J of the Operational Research Society*. **33**, 407-418.

K.C. Laudon and J.P. Laudon (2001) *Management Information Systems*. 7th Ed. Prentice Hall, Upper Saddle River, New Jersey.

P.J. Lewis (1991) 'The decision making basis for information systems: the contribution of Vickers' concept of appreciation to a soft systems perspective' *European Journal of Information Systems* **1**, 33-43.

S. Lichtenstein, B Fischhoff and L.D Phillitps (1982) 'Calibration of probabilities: the state of the art to 1980.' In Kahneman *et al* (1982) 306-334.

C.E. Lindblom (1959) 'The science of muddling through' *Public Administration Review.* **19**, 79-99.

D.V. Lindley (1973) *Making Decisions*. John Wiley and Sons, Chichester.

J. Lochard, T. Schneider and S. French (1992). "Summary Report of Decision Conferences held in the USSR, October-November 1990". EC Report EUR 14543 EN (ISBN 92-826-4631-9).

M. Machina (1991) 'Dynamic consistency and non-expected utility'. In Bacharach and Hurley (1991) 39-91.

B.T. Maguire, B.J. Henderson, J. Potts and R. Wright (1997) 'Applying Janis and Mann's theory of decision making to the characterisation of the people faced with genetic testing' Paper presented at SPUDM, Leeds, August 1997.

E.G. Mallach (2000) Decision Support and Data Warehouse Systems. McGraw Hill, Boston.

G.M. Marakas (1999) *Decision Support Systems in the 21st Century*. Prentice Hall, Upper Saddle River, New Jersey.

K.T. Marshall and R.M. Oliver (1995) *Decision Making and Forecasting*. McGraw-Hill, New York.

D.A. Marchand, T.H. Davenport and T. Dickson (2000) *Mastering Information Management*. Financial Times Prentice Hall, London.

E.F. McClennen (1990) *Rationality and Dynamic Choice.* Cambridge University Press.

J. McGovern, D. Samson and A. Wirth, (1994), "Using Case Based Reasoning for Basis Development in Intelligent Decision Systems", *European Journal of Operational Research*, **74**, pp.40-59.

J. Milnor (1954) 'Games against Nature'. In R. Thrall, C. Coombs and R. Davis (Eds.) *Decision Processes*. John Wiley and Sons, New York. 49-59.

H.S. Migon and D. Gamerman (1999) *Statistical Inference: an Integrated Approach*. Arnold, London.

L. Narens and R.D. Luce (1983) 'How we may have been misled into believing in the intercomparison of utility'. *Theory and Decision* **15**, 247-260.

J.F. Nunamaker, L.M. Applegate and B.R. Konsynski (1988) 'Computer-aided deliberation: model management and group decision support' *Ops. Res.* **36**, 826-248.

J.F. Nunamaker, A.R. Dennis, J.E. George, J.S. Valacich and D.R. Vogel (1991) 'Electronic meeting systems to support group work: theory and practice at Arizona' *Comm. ACM.* ???, 40-61.

R.M. Oliver and J.Q. Smith (Eds) (1990) *Influence Diagrams, Belief Nets and Decision Analysis.* John Wiley and Sons, Chichester.

A. O'Hagan (1994) *Bayesian Inference. Kendall's Advanced theory of Statistics Vol 2b*. Arnold, London.

K.N. Papamichail (1998) 'Explaining and Justifying Decision Support Advice in Intuitive Terms'. *Proceedings of the 13th European Conference on Artificial Intelligence*, 102-103, Brighton.

K.N. Papamichail and S. French, (1999) 'Generating Feasible Strategies in Nuclear Emergencies - A Constraint Satisfaction Problem', *Journal of the Operational Research Society* **50**, 617-626.

I.A. Papazoglou and M.D. Christou (1997) 'A decision support system for emergency response to major nuclear accidents' *Nuc. Tech.* **118**, 97-122.

P.K. Patternaik (1978) *Strategy and Group Choice*. North Holland, Dordrecht.

J.W. Payne, J.R. Bettman and E.J. Johnson (1992) 'Behavioural decision research: a constructive processing perspective' *Annual Review of Psychology* **43**, 87-131.

J. Pearl (2000) *Causality, Models, Reasoning and Inference*. Cambridge University Press.

P. Perny and D. Vanderpooten (1998) 'An interactive multi-objective procedure for selecting medium term countermeasures after nuclear accidents' *J. Multi-Criteria Decision Analysis* **7**, 48-60.

L.D. Philips (1982), "Requisite Decision Making: A Case Study", *Journal of the Operational Research Society*, Vol. 33, pp.303-311.

L.D. Phillips (1984) 'A theory of requisite decision models.' *Acta Psychologica* **56**, 29-48.

L.D. Phillips and M.C. Phillips (1993) 'Facilitated work groups: theory and practice' *Journal of the Operational Research Society* **44**, 533-549.

C.R. Plott (1976) 'Axiomatic social choice theory: an overview and interpretation' *American J. Poliitcal Science* **XX**, 511-596.

Sir Karl Popper (1963) *Conjectures and Refutations*. Routledge and Kegan Paul, London.

C.T. Ragsdale (2001) *Spreadsheet Modelling and Decision Analysis*. South Western College Publishing, Cincinnati.

H. Raiffa (1968) *Decision Analysis*. Addison Wesley, Reading, Mass.

P Reagan-Cirincione (1992) 'Combining group facilitation, decision modelling and information technology to improve the accuracy of group judgement' ???

J. Reason (1987), "Cognitive Aids in Process Environments: Prostheses or Tools?", *International Journal of Man-Machine Studies*, Vol. 27, pp. 463-470.

P.J. Regan and S. Holtzman (1995), "R&D Advisor: An Interactive Approach to Normative Decision System Model Construction", *European Journal of Operational Research*, Vol. 84, No. 1-4, pp. 116-133.

T. Rickards (1999) 'Brainstorming' *Encyclopaedia of Creativity* Volume 1, Academic Press.

F. S. Roberts (1979) *Measurement Theory*. Addison Wesley, Reading, Mass.

K.W.S. Roberts (1980) 'Interpersonal comparability and social choice theory' *Rev. Ecom Stud.* **XLVII**, 421-439.W. Robson (1994) *Strategic Management and Information Systems*. Pitman, London.

L. Robertson and C. Persson (1993) 'Attempts to apply four-dimensional data assimilation of radiological data using the adjoint technique' *Radiation Protection Dosimetry* **50**, 333-337.

J. Rosenhead and J. Mingers (Eds.) (2001) *Rational Analysis for a Problematic World Revisited*. 2nd Ed. John Wiley and Sons, Chichester.

A. Saltelli, K.Chan and E.M. Scott (Eds.) (2000) *Sensitivity Analysis*. John Wiley and Sons, Chichester.

V.V.S. Sarma (1994), "Decision Making in Complex Systems", *Systems Practice*, Vol. 7, No. 4, pp. 399-407.

V. Sauter (1997) *Decision Support Systems*. John Wiley and Sons, New York.

L.J Savage (1954) *The Foundations of Statistics*. John Wiley and Sons, New York. (2nd Edition Dover, 1972).

P. Schoemaker (1993) 'Multiple scenario development: its conceptual and behavioural foundation' *Strategic Management J.* 193-213.

M.S. Silver (1991) *Systems that Support Decision Makers: Description and Analysis*. John Wiley and Sons, Chichester.

B.G. Silverman (1995), "Knowledge-Based Systems and the Decision Sciences", *Interfaces*, Vol. 25, pp. 67-82.

H. Simon (1960) *The New Science of Decision Making.* Harper and Row, New York.

P. Slovic and A. Tvesky (1974) 'Who accepts Savage's Axiom?' *Behavioural Science* **19**, 368-373.

J.Q. Smith (1988) *Decision Analysis: a Bayesian Viewpoint*. Chapman and Hall, London.

S. Smithson and R. Hirschheim (1998) Analysing information systems evaluation: another look at an old problem. *European Journal of Information Systems* **7**, 158-174.

A. Srinivasan, D. Sundaram and J. Davis (2000) *Implementing Decision Support Systems*. McGraw Hill, London.

R.J. Sternberg (1985), *Beyond IQ: A Triarchic Theory of Human Intelligence*, New York: Cambridge University Press.

S. Strasnick (1976) 'The problem of social choice: from Arrow to Rawls' *Philosophy and Public Affairs* **5**, 241-273.

K.W. Tracy`and K.W. Bouthoorn (1997), *Object-Oriented – Artificial Intelligence Using C++*, Basingstoke, England: Freeman.

E. Turban (1993), *Decision Support and Expert Systems: Management Support Systems*, New York: McMillan.

E. Turban and J. Aronson (2001) *Decision Support Systems and Intelligent Systems*. 6th Ed. Prentice-Hall, New Jersey.

A.M. Turing (1950), "Computing Machinery and Intelligence", *Mind*, Vol. 59, No. 236, pp. 433-460.

A. Tversky (1969) 'Intransitivity of preferences'. *Psych. Rev.* **76**, 31-48.

A. Tversky and D. Kahneman (1981) 'The framing of decisions and the psychology of choice' *Science* **211**, 453-458.

J. Von Neumann and O. Morgenstern (1947) *Theory of Games and Economic Behaviour.* 2nd Edn. Princeton University Press.

B.Warboys, P. Kawalek, I. Robertson and M. Greenwood (1999) Mc-Graw-Hill, London.

H.J. Watson, G. Houdeshel and R.K. Rainer Jr. (1997) *Building Executive Information Systems and other Decision Support Applications*. John Wiley and Sons, New York.

D.A. Waterman, (1986) *A Guide to Expert Systems*. Addison-Wesley.

M. West and P.J. Harrison (1989) *Bayesian Forecasting and Dynamic Models*. Springer Verlag, New York.

D.J. White (1976) *Fundamentals of Decision Theory*. North Holland.

D.J. White (1985) *Operational Research*. John Wiley and Sons, Chichester.

L.P. Willcocks (1999) Evaluating the outcomes of information systems plans. In Galliers *et al* (1999) 270-290.

L.P. Willcocks and S. Lester (1999) Information technology and organisational performance: beyond the IT paradox. In Galliers *et al* (1999) 551-572.

R.B. Wilson (1972) 'Social choice theory without the Pareto Principle' *J. Econ. Theory*. **5**, 14-20.

M.P. Wiper and S. French (1995) "Combining Experts' Opinions using a Normal-Wishart Model.". *Journal of Forecasting*, **14**, 25-34.

D. von Winterfeldt and W. Edwards (1986) *Decision Analysis and Behavioural Research*. Cambridge University Press.

G. Wright and P. Ayton, Eds. (1994) *Subjective Probability*. John Wiley and Sons, Chichester.

G. Wright and P. Goodwin (1999) 'Future-focused thinking: combining scenario planning with decision analysis'. *Journal of Multi-Criteria Decision Analysis*. **8**, 311-321.

D. Yeates and James Cadle (1996) *Project Management for Information Systems*. Financial Times Pitman Publishing, London.

Appendix: Basic Probability Concepts

The Laws of Probability

Mathematically, the laws of probability are simple and well known:

- probabilities are non-negative.
- the probability of an impossible event is zero.
- probabilities of disjoint events add up: *viz.*

$$P(A \cup B) = P(A) + P(B) \quad \text{if } A \cap B = \emptyset$$

Usually this works over a countably infinite[15] set of events and probability is said to be countably additive. The total probabilities over all possibilities is 1, i.e. the probability of the certain event is 1.

$$\sum_{\text{All possible events}} P(\text{event}) = 1$$

Conditional probabilities and independence

If we have a set of probabilities representing some uncertainties and then we learn that something has happened (or not happened), then we need to revise our probabilities to reflect our changed knowledge. The *conditional probability* of event *A* given that event *B* is known (or assumed) to have happened is defined to be:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \qquad (2)$$

i.e. the ratio of the probabilities of *A* and *B* happening together to that of *B* happening with

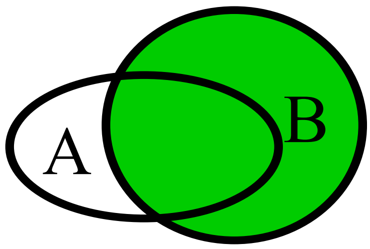


Figure 50: Illustration of the definition of conditional probability

[15] If words such as "countably infinite" mean little or nothing to you, pass them by. The aim of this section is to get the basic ideas over. The trouble is that

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

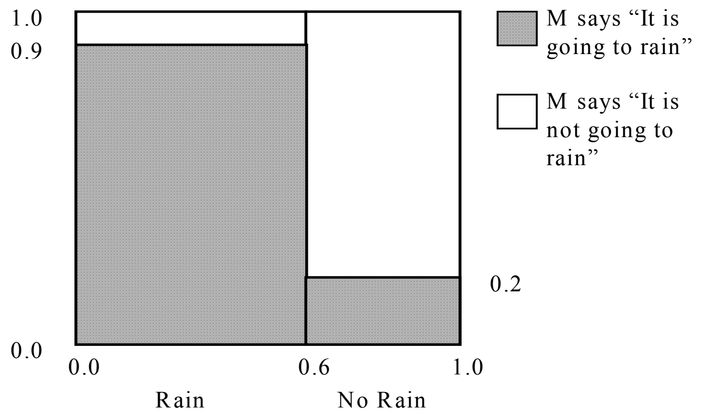


Figure 51: The rain example

or without A: see Figure 50. Intuitively this is a very reasonable definition and this intuition can be made more rigorous: see DeGroot (1970), French (1986) or French and Rios Insua (2000) for a discussion.

Note $P(B|B) = 1$, i.e. a certainty always has the probability 1.

A and B are (*probabilistically*) *independent*[16] if learning B tells us nothing about A: *viz.*

$$P(A|B) = P(A) \qquad (3)$$

Equivalently: A and B are *independent* if $P(A \cap B) = P(A) \times P(B)$, as may be deduced on substituting (2) into (3). This equivalent representation is, in fact, the most common definition of independence. However, it is not the most intuitive for modelling.

Independence is important for building models that are computationally tractable, but dependence is much more interesting in terms of the model itself. Without it we cannot learn. Only if the potentially observable data is dependent upon parameters of interest in the model – and thus on facets of interest in our perception of the world – can we assimilate and learn from data. Otherwise data tell us nothing. Thus there is a balance to be drawn which we shall return to at several points over the course.

Bayes Theorem

We remarked above that *dependence* (i.e. the absence of independence) gives the opportunity of learning. Consider a simple example.

Example A1

By consulting your seaweed, which is damp but not sopping wet and looking at the sky, you come to the conclusion that the chances of rain tomorrow are about 60:40. You then ask a friendly meteorologist and he says that it is going to rain tomorrow. Call him 'M'. Now you have asked M for his forecasts on many occasions in the past and have noticed that he either says "it is going to rain" or "it is not going to rain". Meteorologists clearly have a limited vocabulary. Moreover, you have noticed that M is not always right – this is a realistic example! On the occasions that it has rained, he has predicted rain about 90% of the times, but 10% of the times he hasn't. And on those occasions that it has not rained, he has predicted no rain 80% of the times, but rain 20% of the times.

We can illustrate these numbers on a diagram: see Figure 51. We have drawn a unit square, which we shall use as a Venn diagram, and remembering that probabilities sum to 1, we can make areas in this square correspond to probabilities. Your probability of 0.6 of rain is represented by dividing the square vertically in the ratio 0.6:0.4. Next we divide each vertical column according to the proportion of times that M says it will rain. The column which corresponds to your likelihood of rain will, therefore, be divided so that 90% of it corresponds to the times M says it will rain and the remaining 10% to the times that he says it

what is basic in terms of probability for some is far from basic to others.

16 We shall usually drop the qualification 'probabilistically'. However, when we discuss multi-criteria decision making, other concepts of independence will enter the discussion and bring with them the potential for confusion. You have been warned.

will not. The column corresponding to your likelihood of no rain is divided similarly, but in the ratio 0.2:0.8. Now he has said it will rain. So you attention is confined to the shaded area, the two regions corresponding to M's statement "it is going to rain". The larger area to the left corresponds to your likelihood of rain; that to the right to your likelihood of no rain. Thus your probability of rain *after* hearing the M's view is:

$$\begin{aligned} P(\text{rain} \mid \text{M says "rain"}) &= 0.9\times 0.6/(0.9\times 0.6 + 0.2\times 0.4) \\ &= 54/62 = 0.87 \end{aligned}$$

Your probability of rain has increased from 60% to 87% on hearing and assimilating M's statement. Translating the numbers above into probabilities,

$$\begin{aligned} P(\text{rain} \mid \text{M says "rain"}) &= P(\text{M says "rain"} \mid \text{rain}) \\ &\quad \times P(\text{rain})/P(\text{M says "rain"}) \end{aligned}$$

This is the simplest example of the application of Bayes' theorem:

Bayes Theorem[17]

Informally:

$$\begin{aligned} P(B|A) &= \frac{P(A\cap B)}{P(A)} \\ &= \frac{P(A|B)\times P(B)}{P(A)} \\ &\propto P(A|B)\times P(B) \end{aligned}$$

The probability *a posteriori* of B after you have learnt A is the probability of B *a priori* multiplied by the conditional probability of A if you assume B, known as the *likelihood*[18].

posterior probability ∝ likelihood × prior probability

More formally, Bayes Theorem states: if B_1, B_2, …, B_N forms a partition of the certain event, *viz*.

$$P\left(\bigcup_{i=1}^{N} B_i\right) = 1;\ B_i \cap B_j = \varnothing \text{ for } i \neq j$$

then

$$\begin{aligned} P(B|A) &= \frac{P(A\cap B)}{P(\bigcup_{i=1}^{N} A\cap B_i)} \\ &= \frac{P(A|B)P(B)}{\sum_{i=1}^{N} P(A|B_i)P(B_i)} \end{aligned}. \quad (4)$$

Bayes theorem underpins Bayesian approaches to statistics and decision making in the light of evidence. Note that in the example there is dependence: the probability that M says "It is going to rain" varies with the occurrence of rain. It is this dependence that allows you to update your belief in rain. Another example is relevant since it demonstrates that the result of a test should change ones opinion rather less than one might at first think.

Example A2

Suppose that a health authority screens a population for a particular type of cancer. It is known that 0.05% of the population has cancer. The screening test that is used will detect the cancer, if a person has the cancer, 98% of the time and will produce a spurious result, i.e. 'detect' the cancer when the person is free of cancer, 2% of the time. What is the probability of a person who gives a positive test result actually having cancer?

Define the events:

A_1 – the person has the cancer;

A_2 – the person does not have the cancer

B – the person gives a positive test result.

We require $P(A_1 \mid B)$. The information on proportions gives:

$P(A_1) = 0.0005$, $P(A_2) = 0.9995$,

$P(B \mid A_1) = 0.98$, $P(B \mid A_2) = 0.02$.

Whence, applying Bayes' theorem:

$$\begin{aligned} P(A_1|B) &= \frac{(0.98\times 0.0005)}{(0.98\times 0.0005+0.02\times 0.9995)} \\ &= 0.024. \end{aligned}$$

This means that only about 2 or 3% of the cancer suspects produced by the screening will, in fact, have cancer. This may seem counter-intuitive, but a little reflection supports its good sense. Most people on encountering the data, assume that a positive result in the test indicates a very high chance of cancer: after all, it is 98% accurate. But in making that assumption, they are forgetting that the

17 In Example 1 B corresponds to 'rain' and A to M's statement "It is going to rain".

18 In statistics, the likelihood function is the probability of the data actually observed given the unknown parameters *considered as a function of those parameters*.

incidence of cancer is very low[19]. Although the test detects 98% of the people who actually have cancer, it also throws up a spurious positive result for 2% of the healthy population. Since there are 2000 times as many healthy people as those who have cancer, the spurious results swamp the true ones.

Discrete and continuous distributions

All the above were stated in terms of discrete events, but the theory of probability applies with continuous variables or quantities. Probabilities are given by integrals of density functions; and generally in the laws of probability integrals replace sums. Thus Figure 52 illustrates a density function and a distribution function $P_X(\cdot)$ for a continuous variable X. The density function $p_X(\cdot)$ describes the probability distribution of X by[20]:

$$P_X(a < x < b) = \int_a^b p_X(x)dx$$

Similarly the distribution function describes the uncertainty about X by:

$$P_X(x < a) = \int_{-\infty}^a p_X(x)dx$$

Ideas of *conditionality* carry through (with suitable care about dividing by zero). Independence means that the joint distribution is the product of the *marginals*. To be a little more precise, for continuous variables X and Y:

$$P_{X,Y}(X \le x, Y \le y) = \int_{-\infty}^x \int_{-\infty}^y p_{X,Y}(x', y')dy'dx'$$

$$P_X(X \le x) = \int_{-\infty}^x p_X(x')dx'$$

$$p_X(x) = \int_{-\infty}^\infty p_{X,Y}(x, y')dy'$$

$$p_X(x|Y = y) = \frac{p_{X,Y}(x, y)}{\int_{-\infty}^\infty p_{X,Y}(x', y)dx'}$$

where:

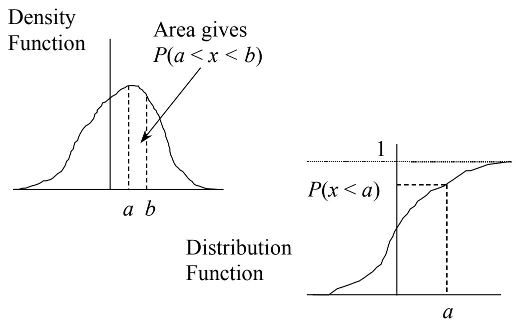


Figure 52: Continuous probability distributions

1. $P(.)$ represents a *distribution function*, that is the probability that uncertain quantities are less than or equal to given values;
2. we have used subscripts to indicate the uncertain quantities concerned;
3. the conditional density of X given that $Y = y$ is given by dividing the joint density of X, Y at (x, y) by the marginal density of Y at $Y = y$, assuming that this does not involve a division by zero.

As before, the definition of a conditional density leads to the definition that two quantities are (*probabilistically*) *independent* if learning anything about one provides no information about the other: or, equivalently, two quantities are probabilistically independent if their joint density is the product of their marginal densities:

$$p_{X,Y}(x, y) = p_X(x) \times p_Y(y)$$

For density functions of continuous quantities, Bayes' Theorem becomes:

$$p_X(x \mid Y = y) \propto_y p_Y(y \mid x) \times p_X(x)$$

where $\propto_y$ means is proportional to as a function of y. As before we may write:

posterior density $\propto$ *likelihood* $\times$ *prior density*

Note that the use of the word 'likelihood' here corresponds precisely to the likelihood function used in much non-Bayesian statistics: see e.g. Migon and Gamerman(1999).

19 Actually there is a subtlety that we should admit in real life. The example stated "suppose a health authority screens a population …' and the example goes on to state that the only 0.05% of the population has the cancer. The implication is that the screening test is applied to the entire population. In practice, few tests are applied so widely. Patients are selected for screening on the basis of some symptoms or because they belong to a sub-population at a statistically higher risk than the general population. However, this does not alter the behavioural point: most people grossly overestimate the posterior probability of having cancer given a positive test result.

20 We ignore subtleties relating to strict or weak inequalities and the existence of atoms of probability at certain points: i.e. to the use of mixed continuous and discrete distributions

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB

Index

Manchester Business School
University of Manchester
Booth Street West
Manchester, M15 6PB