ARTIFICIAL INTELLIGENCE AND HUMAN DECISION MAKING

Jean-Charles POMEROL

1 - Introduction

It's quite difficult to define precisely the word decision, but each of us generally agrees that he has already experienced the concept. Every human being, right or wrong, think that in many occasions he has made a choice between different alternatives. Whether he exercises his free will or bends to some kind of causal necessity, is another (philosophical) question we do not enter into. The intuitive notion of human free will in choosing between various alternatives will suffice for our talk.

On the other hand, it is important to specify what we mean when using the expression "Artificial Intelligence" (AI). There are at least two different views about AI. The first one equates AI to "Sciences of the Artificial" (Simon, 1969), or the science of designing and building computer-based artifacts performing various human tasks. Adopting this view has the advantage of throwing out most of the philosophical discussions about the nature of intelligence and the feasability of the AI project. This view of AI has relatively few links with decision to the extent that an artifact cannot properly be said to make a decision. The decision, if any, has of course previously been made by the designer of the system (at least if he is able to trace, for any input set of data, the instructions trigerred). In other words, the concept of "decision" is antinomic to the idea of program. When a task is programmed, the decision no longer exists since the actions are determined according to each possible situation that may occur (see Pomerol, 1992a, 1992b and Lévine-Pomerol 1995, for a discussion and the consequences for Decision Support Systems (DSS)).

But even if an artifact does not make any decision, its designer has previously modelled a decision process embedded in the system. And this is a first question for us : how to model and program decision processes in the artifacts ? The most natural answer to this question is that "it suffices" to observe how people make the decision in the task at hand and to reproduce the process into the machine. So, even if we adopt a view of AI not referring to "human intelligence", we have to deal with human reasoning.

Ce papier doit être présenté comme tutorial à EURO XIV, à Jérusalem, en Juillet 1995.

We thus reach the second definition of AI, relating to its cognitive side. We know that AI is often regarded as the science of knowledge representation and reasoning (Newell and Simon, 1972). If we therefore think about AI as a science aimed at mimicking human beings, then it obviously has a non void intersection with decision. The difficulty is that each human being may have his/her own way of reasoning and deciding, at least at the preference level. In this case AI becomes a science of the persons (subjects), a subjective science. This point of view has already been advocated by some authors (e.g. Dubois in Courbon et al., 1994). Following this idea, AI is the science of the design and development of systems mimicking not mankind (genericity), but a given human being (subjectivity). Let us note that, on the contrary, up to now, AI has considered that it makes sense to look for generic properties and representations rather than developing specific and subjective skills. It is only recently that AI has been sensitised to subjectivity, and consequently to decision since it is generally acknowledged that decisions are personal. Since the introduction of utility functions by the economists (see e.g. von Neumann and Morgenstern, 1944), it is actually accepted that two "rational" decision makers, confronting the same situation, may make two different decisions depending on their **subjective probabilities**.

This debate about genericity vs. subjectivity has existed since the origins of decision theory. Some researchers defend the idea that everybody decides or should decide in the same (rational) manner ; they represent the normative current in decision theory. On the other hand, some people try to understand how a given subject decides. They find several typical behaviors, with a more or less large variabilities ; they are representative of the descriptive or behavioral school in decision theory.

So it is difficult to envisage the relationships between AI and decision making without beforehand establishing the amount of subjectiveness we concede to AI. In what follows I will share the idea that *AI encompasses a large subjective component*. In other words without denying the social component of intelligence (see Sternberg, 1990 for a survey and Latour, 1992), we will concede a strong personal component underlying every example of intelligent behavior.

This first statement having been made, our paper is organized as follows. In section 2, we posit the different aspects of any decision, then in the rest of the section we examine the links between each aspect of decision and AI. Section 3 is devoted to diagnosis and pattern matching. In the next section, we examine the usual views of AI about goals and plans. Finally, in the last section we introduce some new trends of decision science (multicriteria analysis, decision support systems), techniques which become difficult to discriminate from the AI engineering of "decision maker actifacts".

2 - Reasoning to decide

We can admit that any decision has its origin in a dissatisfaction. *Let us agree to call this dissatisfaction a decision problem.* The dissatisfaction arises from the difference between the current state of affairs and another, not yet existing, more desirable one.

current state of the world $\neq$ more desirable (future) state

$\downarrow$

decision
problem
or
"what to do to reduce the difference ?"

Figure 1 - Decision Problem

Note that the personal dimension appears right from the start, because what is desirable for one person may not be wished by another. Note also that, in a context where things are known with certainty, the current state is perceived as being unique for a sane person ; on the other hand, it may happen that there exist many different attributes in the desirable state which are not fully compatible (i.e. get a better wage and more free time).

Before making a decision, the subject recognizes the current state. Let us agree that the current state contains information about the past and the present. In other words, the subject knows a part of what happened before the present and has his own perception of what occurs now. Then, keeping in mind his perception of the current state, the subject tries to identify it by reference to his experience. *This means that we can assume that the subject has recorded many situations or states he has already met or has learnt about. Let us call these states "recorded states".*

*The first phase of decision is then to find one or several recorded states close to the perceived current state. According to the context and complexity, this operation can be denoted "pattern matching" or "**diagnosis**".* In AI we generally reserve the expression "pattern matching" for the recognition of a simple pattern, whereas "diagnosis" is used for the identification of more complex states ; in control, we use the term state-estimation problem or simply observation. Finally, in some cases, it is very difficult to know what happened and what is the exact status of the situation we are dealing with ; thus, we cannot dismiss the possibility of a subject identifying various possible current states with estimated probabilities. The diagnosis is uncertain.

Sometimes it is possible to draw a kind of table, or a one to one function (injection), from the set of recorded states to the action set. Let us denote this injection by φ ; $\varphi(S_i) = A_j$ means that if the reference state is S_i , then the best (or usual) decision is the action A_j.

Up to now we have not yet considered the future. But it is clear that any decision involves the future. People with no future do not worry with decision making ! The decision necessarily depends on the future (desirable) state one wishes to reach. In decision theory these **preferences** are expressed by the **utility function** of each subject. We have therefore to add one attribute to φ, assuming that P denotes the preferences about the future states, the chosen action is $A_j = \varphi_P(S_i)$. This view of an action as a reaction to a given state is reminiscent from control theory (see Dean and Wellman, 1991 for a presentation in a AI perspective). In a control problem, let S_t be the state at the instant t, the state at t+1 is given by the state equation

$$S_{t+1} = f(S_t, U_t)$$

where U_t (input) is the action carried out at the instant t. When U_t is directly deduced from the observation of the state of the system by a regulation function r we can assimilate r to φ (Dean and Wellman, 1991 figure 1.11), being reminded that a system has no preferences.

With regards to the simple picture we have just sketched before, we have to introduce two unavoidable concepts difficult to handle. The first one is the inherent *uncertainty of future* and the second one is the *granularity of the decision.*

Let us begin with uncertainty. The future states are obviously not known with certainty. Making the decision A_j may result in the future state F_1 in some cases and in the future state F_2 in some other circumstances. Suppose that F_1 is desirable and that you dislike F_2 very much; then the decision becomes difficult. In some cases the probability of occurence of F_1 vs. F_2 is known because the events are frequent (e.g. the probability for a train of a given railway network to be delayed may be known since there are many trains each day), but in many cases it is very difficult to obtain shared (objective ?) probabilities, each subject having their own ; this is the reason why they are called subjective probabilities. A lot of researchers have also advocated that there is no probability at all and various models have been proposed to take into account the representation of uncertainty in the mind.

In any case, *let us denote E (for **Expectations**) a representation of the future events uncontrolled and uninfluenced by the subject (states of nature with probabilities in decision theory)* and let A denote the set of all possible actions. A multifunction $\psi(S_i, A, E)$ defines the set of all *the future states attainable from the current state* S_i and $\psi(S_i, A_j, E)$ is the future state attainable from S_i when the subject chooses the action A_j, E being their anticipations about the states of nature. In an uncertain setting, it may happen that many states are attainable with different probabilities. The decision consists of choosing an action $A_j = \varphi_P [\psi(S_i, A, E)]$. We can summarize the situation in a diagram (fig. 2).

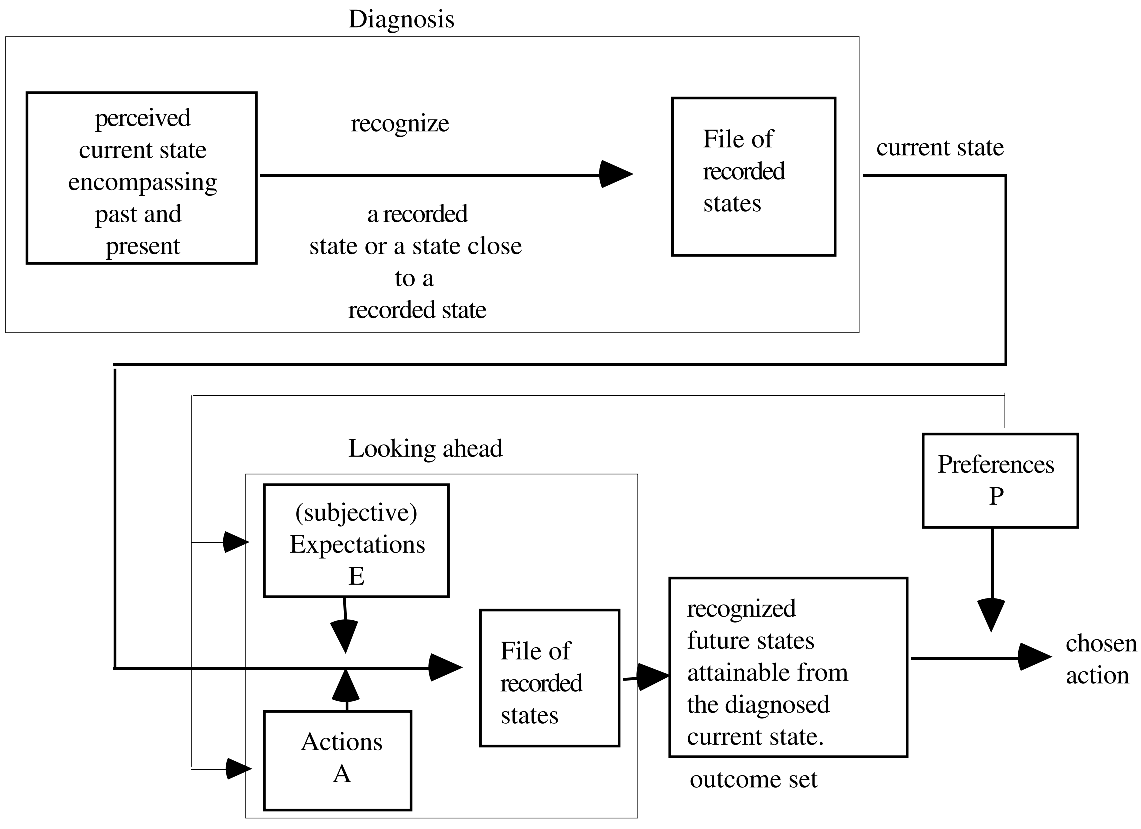

Figure 2 - Decision process

Figure 2 deserves some further comments. We have drawn an arc (thin line) from the preference box to the actions, because a lot a people think that it is possible, to some extent, to define the actions according to the preferences. First define what you want, then define the actions to reach it ! This is expressed by the research current mainly originated by Keeney (1988, 1992) about value-driven thinking. These works draw the attention on the fact that the action (or alternative) set is not given and can be changed during the decision process reasoning. It has been often observed that many real decision makers are over-constrained in their perception of the alternative set and study just a small part of the possible alternatives.

In some cases, it is also defendable to draw an arc from the preferences to the expectation box (thin line). This may be regarded as a psychological bias, because it means that the future is seen according to the preference. This is probably a frequent situation that must be avoided in rational decision making.

The set of the "*recognized future states attainable from the diagnosed current state*" can be viewed as the **outcome set** in decision theory. For this reason, we will denote it OC. The preferences of the subject apply to this set and, by a regressive reasoning, the preferred outcome defines the action to be chosen (decision). We may think of the elements of OC as the possible **goals** of the subject. At least, it seems that this is the usual meaning of the word "goal" in AI (see § 4.1). Thus, in our scheme the goals are defined among the attainable states, the difficulty being that it is not easy to determine what is attainable or not, and this depends on the actions and the expectations. In a word, a goal is the result of a complex alchemy combining possibles actions, recorded experiences, expectations and preferences.

We have also adopted some simplifications in Figure 2. The first is that we have assumed that the subject makes an unambivalent diagnosis and consequently that the current state is known with certainty and is unique. This is obviously false because in many cases we have not enough information to be certain of the current state. In most cases we think that we are in a given current state with some uncertainty (probability) and possibly in another with some probability.

Also, the preferences may influence the diagnosis process and the file of the recorded state (memory). Numerous psychological biases are observed in this domain (see von Winterfeldt and Edwards, 1986 and Bell et al., 1988). Another simplification on Figure 2 is that the decision process may appear as being "linear", this is not the case and many backtracks can occur, especially when the subject becomes aware that the attainable future states are not satisfactory. Moreover, in many practical organizational settings, due to feedback phenomena, it is not obvious to distinguish an event from an outcome. For example, in an oligopolistic market, does price rising (or decreasing) is an event (uncontrolled and uninfluenced) or an outcome ? In many cases the decision makers and the modelers do not know where to set up the limit of the model, because depending on the level of analysis, any significant decision may have far away consequences, and the frontier between events and results becomes rather vague.

In Figure 2, we have sketched what may be regarded as a realistic decision process tracking the main ingredients of decision reasoning. For simplicity's sake we have divided the process into two main parts, diagnosis and look-ahead. It is of course not always easy to separate these two reasonings but, from an engineer's point of view, this facilitates the design of the system.

Let us return to what I consider as the second (with uncertainty) major difficulty in decision modelling especially in management (with contrast to control), it is the decision and time granularities and the related problems. In the framework of process control, each action (input) is almost immediately (depending on the process, from less a second to few minutes) followed by a modification of the state and a new decision. We are faced to very reactive systems. In such systems, the role of the environment is very low because it has not enough time to lead to large effects during the lapse of time between two decisions. For example, the external temperature does not have a big influence upon the control of a cement kiln, just as moderate wind upon car driving. On the other hand, if we consider the decision of doing an investment, then a change in the market or in the rate of the currencies may dramatically affect the result.

The difference comes from that we do not consider, depending on the context, the same "quantity of decision" as a decision chunk. We can admit that almost any decision process is theoretically continuous : i.e. at each instant the decision maker is making a decision. (In this setting "doing nothing", is also a decision). *The continuity assumption makes sense for an agent driving a bicycle, a car or an industrial process.* At each instant, he moves the handlebar or the wheel. But it is very different for a CEO making the decision to purchase a new facility, he gets the results some months later after the environment (competition, market, etc...) has produced its effects. Another difference is that, for the CEO, the object of the decision is also not continuous. Turning a wheel may be regarded, within some limits, as a continuous process, whereas purchasing a significative facility may certainly not. In some sense, the main difference between this two types of decision that we will denote hereafter **"almost continuous decision"** and **"discrete decision"** is a difference of granularity. The difficulty being that in management people generally consider large grain decisions, i.e. discrete decisions.

This question of the granularity of the decision has repercussions on the actions. The significance of an outcome for a subject frequently involves many sub-actions to reach this outcome. For example, the outcome "spend a week of agreable holidays by the sea", is possibly obtained, not only by deciding to go to the Bahamas, but also by many sub-decisions such as choice of the period, the tour operator, the boy (girl)-friend, and so on. It is the completely successful chaining of all these sub-choices which implies the fulfilment of the goal. From a theoretical point of view, it does not matter whether you reason from one "atomic chunk" of action or a sequence (or a vector) of sub-actions. A sequence of sub-actions is also called a **policy** in various contexts. But, from a practical point of view, it does matter ! Let us stress just two points : it may occur that some events happen between the implementation of two different sub-actions. This is the situation envisaged in decision trees. A sub-action a_1 is followed by an event and according to the event, for example, in case of an event with two outcomes (e_1 and e_2), the subject makes the sub-decision a_2 or a_3 according to e_1 or e_2.

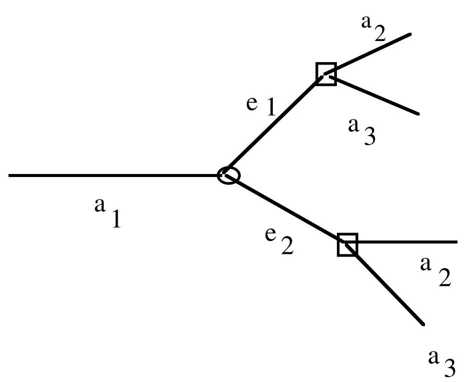

Figure 3 - An elementary chance node

So the two decisions to be considered are $a=(a_1,a_2)$ and $a'=(a_1,a_3)$ and the decision tree allows each action, a and a', to be evaluated according to the value of the final outcomes (terminal nodes) and their probabilities.

To make this clear, suppose that our traveler has to choose between Israel and Egypt to spend his holidays and then to announce whether he will stay or leave before the end of the holidays in case of troubles in the chosen country (e.g. terrorist bombing). Let us assume that the occurrence of troubles is independent in Egypt and Israel (otherwise we should consider the couples of events "troubles in Israel" (TI) or no troubles (NTI) and "troubles in Egypt" (TE) or no troubles (NTE) and the four events (TI, TE), (TI, NTE), (NTI, TE) and (NTI, NTE)). In case of independence, the decision tree is drawn in Figure 4 on which the figures on the event arcs indicate the probabilities of the events.

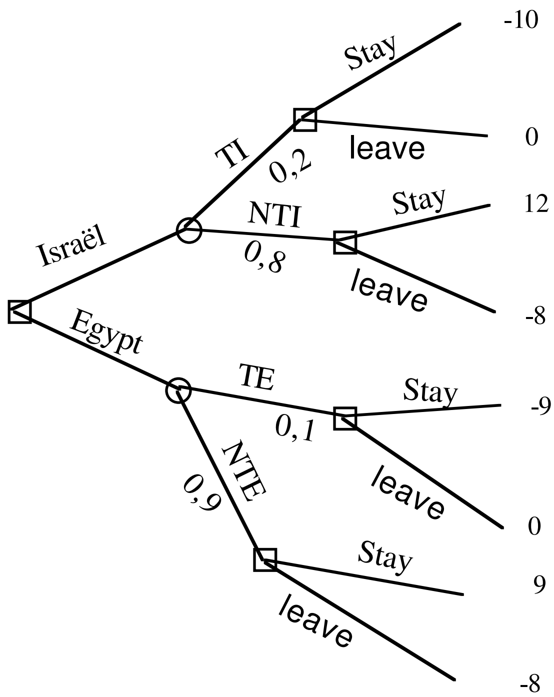

Figure 4 - A decision tree

We have indicated the value of the outcome for each terminal node. A path from the root of the tree to a terminal node is a **scenario**. For example the first scenario is : our traveler goes to Israel, troubles occur and he stays. The possible decisions are a_1=(Israel, leave), a_2=(Israel, stay), a_3=(Egypt, leave) and a_4=(Egypt, stay) with expected value $v(a_1)= -6.4$, $v(a_2)=7.6$, $v(a_3)=-7.2$ and $v(a_4)=7.2$. The expected values and the choice of the decision is recursively computed from the terminal nodes (see e.g. Dean and Wellman, 1991). In our example, the best decision is to go in Israel and decide to stay even in case of troubles. It is noteworthy that this decision results in the worst outcome in case of troubles in Israel. The reader is referred to Dean and Wellman (1991, Ch. 7) or Shenoy (1994) for a simple presentation of decision trees and probabilistic networks.

A *sequence of sub-actions intertwined with events is what we have called a scenario*. From another more theoretical point of view, a sequence of sub-actions is also an action. Let us observe that when we say that $a=(a_1, a_2)$ is an action (i.e. the couple (Israel, stay in case of trouble)) the decision maker is committed to this action or policy before the chance draw ; he cannot in this scheme change his mind after the occurence of the troubles. This would be another decision. With AI vocabulary we would say that (a_1, a_2) is a **plan**, an **irrevocable plan** in this case (see § 4.1).

Another point is that, it is possible, in decision reasoning, to reason at different hierarchical levels. For example, you may first choose the place for your holidays and the style of your stay, then make a decision about the details (hotel, car renting, etc...). Hierarchical reasoning appears to be very frequent in decision making. This must not be confused with reasoning on a sequence of sub-actions ; in this latter case the action is just a short word (abbreviation) for the sequence, this is different from reasoning at different hierarchical levels.

According to Savage, classical decision theory (see Shafer, 1988) does not distinguish between the current states and future states. Any state of the world encompasses indistinctly past and future. Given an action, all the possible outcomes for each state of the world are known and valued by a utility function. Under some "rational" axioms (Shafer, 1988) the choice must lead to maximization of the expected utility. Nor does the decision theory refer to the file of the recorded states, because it does not consider the questions relative to both the construction of the action and of the "state of the world" sets, nor does it consider the question of problem emergence (why a person intends to act).

We must mention that we are of course not the first author to draw such a (more or less simplified) picture of decision as Figure 2. In the framework of the supervision of industrial processes, Rasmussen (1980, 1986) has already proposed a "decision ladder" which describes the main sub-tasks occuring in control. This model has been improved by many authors (e.g. Millot, 1988, Boy, 1991). In an industrial setting, the main difference with Figure 2 is that, due to the almost continuity of the decision, people generally do not need to consider the uncertainty of the environment, and the importance of the look ahead phase is therefore reduced, moreover the controller is supposed to have no preference, just a simple "optimal" behavior.

The only originality of our decision scheme (if any) is the emphasis on the look-ahead reasoning. A lot of neuro-physiologist (e.g. Calvin, 1991, 1994) think that one of the attribute of intelligence that emerges during evolution is the forecasting ability which is related to the capability of chaining language sequences including future (therefore virtual) events and acts. As many practioners, I am convinced that moreover this very common look-ahead thinking, or reasoning about scenarios, is generally performed both backward from the goals and forward from the current state (see Rollier and Turner, 1994, for an experiment about the respective merits of both modes of reasoning). Another interesting point is that there are many evidences that *scenario thinking is important in learning* (e.g. see De Geus, 1988). Let us add that this very human look-ahead capacity is the conditions for being able to make tradeoffs between short term and long term outcomes.

In this brief sketch about decision reasoning, we have not tried to develop the informational aspects underlined by Simon (1977) and his famous four phases of decision (see Lewis, 1991, about the universality of this model). Also, we have said nothing about "intelligence" in Simon's sense and we do not examine the origin of subject information. The "design" of the actions is partly covered by the thin line between "Preferences" and "Actions". But the expectations and the file of recorded states probably also play their role during the design of the actions. The "choice" phase deserves no further comments, being merely the application of preferences to the outcome set OC.

And finally, nothing is said about the "review" process because we do not cover the learning and recording process occurring after the decision.

Before turning back to goals and plans according to AI views, we will provide, in the next section, some brief considerations about AI and diagnosis. According to our second definition, AI should be concerned with the modelisation of each operation involved in the decision process. This is surprisingly not the case and, while there has been much work on diagnosis, very little has been done on the look-ahead process ; one of main the purposes, of this paper is to review the research about the latter.

3 - Pattern matching and diagnosis

3.1 • Classification and decision

The first step during the decision reasoning process is a diagnosis phase. The problem is to recognize, as accurately as possible, the current state (including the past). This state may be recognized with certainty (e.g. the present President of France is a socialist) or with uncertainty (e.g. Earth has entered an era of increasingly high average temperature). Contrary to decision theory, it is of practical importance to distinguish between the current state and the future ones because the former is generally known with much less uncertainty than the latter.

The main point is that the current state may be subject to evaluation and comparison. For example, in process control, one can measure the present temperature, stress, and so on, so that, the current state is characterized by a vector of many different measurements. Assume that the number of actions is finite, then, from this valuation of the current state according to many attributes, the decision maker has to choose an action $\varphi_P(S_i)=A_j$. In other words, he has to determine whether $S_i \in \varphi_P^{-1}(A_j)$ or not. Assuming that $\varphi_P^{-1}(A_j)$ for j=1,...., n realizes a partition of S , the decision problem is then solved. Things are often not so simple, but in many cases, the simple preceeding scheme works and the decision problem amounts to affecting a vector of R^n to a class. Decision is then a mere classification problem.

Thus, classification is a first approach to decision. Many mathematical solutions to this problem, belonging to the field of data analysis, have existed for a long time. These methods primarily depend on the fact that the classes are given or are to be built (see Dubuisson, 1990 or Milgram, 1993, where some of these methods are developed for diagnosis). They are generally denoted "pattern recognition methods" and use various statistical and similarity distance ideas.

The situation is more difficult when the number of possible actions is very large, for example increasing the intensity of an electric motor, by steps of 0,01 A, from 0 to 20 Ampere. It is generally more convenient, in this case, to design a regulation function φ which gives directly the intensity according to the state S_i (a fuzzy controller for example). In the next sub-sections we will introduce some significant methods introduced by AI researchers to deal with the preceeding questions.

3.2 • Expert systems

Basically an expert system is nothing more than a diagnosis machine. Starting from a list of facts in whatever syntax you want, it produces an output which is another list of facts. We do not intend to describe the way an expert system performs its tasks of producing new facts from input facts. The important point concerns the semantic of the transformation.

The input facts describe a given situation (current state) and the output is either a diagnosis (e.g. the patient is suffering from this illness (MYCIN, Shortliffe, 1976), the situation of the the client is quite good (Risk manager), the trouble with the device is..., etc...), or a recommendation or an action (e.g. increase the flow of oil, decrease the temperature of the kiln, accept the credit, propose this contract to the underwriter, etc...). The situation is particulary clear in on-line control processes, a domain where expert systems have met large success ; in this domain, the expert systems perform an analysis of the current parameters of the process, deduce the current status of the process and determine the action to undertake. In this case, the expert system exactly performs the function φ. Returning to Figure 2, we can simplify it (Figure 5) to show the position of the function. This function φ shunts the "look-ahead" step. It is therefore not surprising that this type of system is often seen as very naïve in terms of organizational decisions. We also observe that φ does not depend on the preference P. Most expert systems are designed to provide a universal (objective ?) mode of reasoning independent of the user. This is the reason why so much time is devoted to the development of sound, validated knowledge bases shared by everybody, allowing the determination of the unique action to be undertaken. In some less frequent cases, expert systems are used in a decision framework to express the preferences in place of a human agent. A typical example is given by the system VOTE of Slade (1990). In this case the system reasons on the outcome set. This kind of device may be useful either to mimic a virtual agent (for example in negotiation training) or to anticipate the decision of a lacking agent (e.g. a competitor).

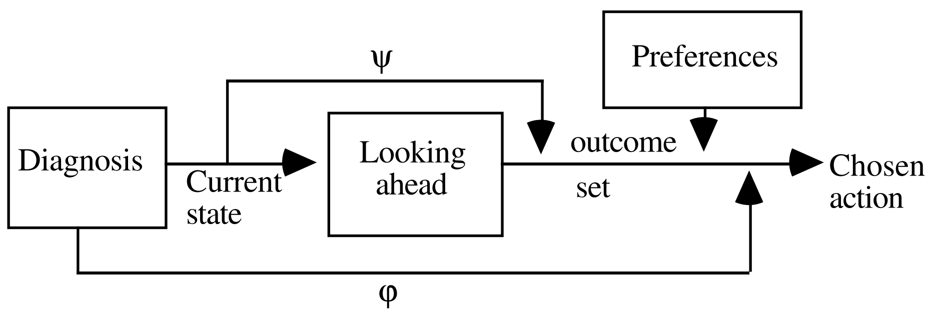

Figure 5 - Simplified decision process

Our interpretation of expert systems as a function φ is supported by the already-known interpretation of each rule as a discrete function between symbolic attributes (Lévine and Pomerol, 1989b), or a real function in the case of fuzzy expert systems. With this interpretation, the inference engine is nothing more than a system composing discrete functions (or real in the case of fuzzyness), a system that we hope to be transitive.

It is also worth stressing that, in most expert systems, the facts are assumed to be known with certainty. We will see methods adapted to the case of uncertainty in paragraph 3.4.

To conclude this sub-section, let us repeat that expert systems involving only a small part of the decision process and relying on a fundamental short cut (Figure 5), can only hope to be used in simple decision processes where : **1)** ***the current data about the situation are known with certainty*** **; 2)** ***subject preferences does not matter*** **; 3)** ***the experience of human performers has led to a direct correspondence between the various possible current states and the decisions at hand*** **; 4)** ***decision is almost continuous*** (see Section 2). This is the case for the control of many industrial processes or machines. This would also be the case for many low level clerical tasks if we did not stumble against the data acquisition obstacle (mainly manuscript reading and vision, see Pomerol and Retour, 1990). Our conclusion is backed by the fact that many "automatic decision process" systems exist in the cases where the four above prerequisites hold (process control, airplane pilot, underground train driving, etc...). As has been observed by Boy (1991), in case of incident, even for human systems, the most common failure is the faulty recognition of the current state. In fact, the simple scheme obeying the four above assumptions supposes that φ^{-1} defines a finite partition of the current state set, and trouble may occur either when the actual current state is wrongly affected to $\varphi^{-1}(A_j)$ when it really belongs to $\varphi^{-1}(A_k)$ $(k \neq j)$, or when it is not comprised in the current state set known by the system. In these two cases, the system is endangered unless rescued by external intervention.

Finally, note that the above four conditions are not satisfied by expert systems addressing discrete decisions, as the ones generally made by live experts. *Expert systems do paradoxically not suit for simulating expert's judgments*. One question is : what is the value of these judgments and the interest of the derived systems ? Let us give just two references which give a rather pessimistic answer about the value of expert's judgments : Hammond (1987) and Mumpower (1987).

3.3 • Numerical methods for classification and diagnosis

As we have seen in sub-section 3.1, classification is diagnosis when the current states can be enumerated. This classification must be made according many attributes or parameters. We do not intend to develop the numerous methods arising from data analysis (see Saporta, 1990, Auray et al., 1990, Diday, 1991, Barthélemy and Guénoche, 1991, Zighed et al., 1992) but rather to emphasize new methods promoted in the AI community.

Among the most promising new methods, is the rough set theory (Pawlak, 1991 and Pawlak and Slowinski, 1994a and b). In few words, starting from a decision table of elementary decisions (each alternative is described by a row of a matrix giving the evaluations of the alternatives according to fixed attributes and the last component is the decision), rough set method provides the upper and lower approximation sets of alternatives corresponding to a decision. One of the interesting features of rough set theory is that it authorizes inexact membership of a class and gives a kind of measure of accuracy of the classification ; it also allows the elimination of not useful attributes. Another appealing feature of rough sets is that they can be used to deduce sorting rules from decision tables. Rough set theory has been used sucessfully for diagnosis : medical (Slowinsky and Stefanovsky, 1994), company failure clues (Slowinsky and Zapoundinis, 1994), decision (Pawlak and Slowinsky, 1994b), credit scoring (Pawlak and Slowinski, 1994a) and many other domains (see Novicki et al., 1992 and Slowinski, 1992).

Neural networks are also a very popular classification method. Briefly, a neural network is an algorithm which affects a given vector V to a class. This is done by comparing the score of V according to a non-analytical function f. It means that V is classified according to its value f(V) (a real number), for an introduction see Hertz et al., 1991. In fact, a neural network can approximate, almost any continuous functional (see e.g. Chen and Chen, 1993). In multicriteria analysis (Pomerol and Barba-Romero, 1993) the function f would be regarded as an aggregation function. Here, the function is built by layers of partial sums of sub-functions. The weights involved in these sums are adjusted by the algorithm in order to fit with a given classification (learning sample). Then, once the weights are fixed, any new input vector V is affected to the closest class, depending on its global score f(V). Many applications have been developed in the decision field using neural networks (see for example, Olmeda and Barba-Romero, 1993 and the references therein). Overall, a neural network is a classification machine just as our expert systems are. There are two differences : in an expert systems the input vector (facts) generally has a very large number of components and is rather symbolic as opposed to numeric, and the adjusment of the intermediary sub-functions is made by a tedious process of knowledge acquisition and validation, whereas in neural network, it is made by a "learning" algorithm. Keeping in mind the deep analogy between expert systems and neural networks, there is nothing to add about the use of neural networks for decision making ; they obey the same four restrictions as expert systems (see sub-section 3.2).

3.4 • Uncertainty and inaccuracy in classification and diagnosis

We have already mentioned that it is not always possible to determine with certainty or accuracy the current state of the system. For example, an oil explorer does not know with certainty whether the rock he is walking on contains oil or not. However, he does not drill at random. Observing some symptoms a doctor is not always sure of the illness of the patient, and so on. AI has a long tradition of dealing with such problems. First Shortliffe (1976), in MYCIN experience tried to introduce uncertainty into expert systems. He was followed by Duda et al. (1978) with PROSPECTOR. But these two attempts where impeded by the lack of a strong coherent model underpinning the calculations performed by the systems (Jaffray, 1985 and Henrion, 1987).

Some more interesting approaches has been proposed. Among the best known is the "possibility theory" introduced by Zadeh in 1978 (see Dubois and Prade, 1985). This theory has many links with previous works of Dempster (1967) and Shafer (1976). These works, essentially relax the additivity assumption of probabilities. Up to the mid-eighties, the two main obstacles in using probabilities in diagnosis were : 1) the amount of data to collect to cope with any realistic situation (probabilities of each event and every combination of events in case of non-independency), 2) the huge amount of calculation.

These two obstacles have been swept away by Pearl (1988) and some others (see Lauritzen and Spiegelhalter, 1988, Neapolitan, 1990, Oliver and Smith, 1990). They have developed algorithms able to propagate probabilities in the case of (restricted) dependency (a node is supposed to be independent of its ancestors conditionally to its parents). This field is now known as bayesian or probabilistic networks (for an introduction see Dean and Wellman, 1991 ch. 7 or Shenoy, 1994). In a graph of dependent facts it is possible now to isolate independent subsets and to propagate the probabilities accordingly. The technique of bayesian networks allows uncertain diagnosis, relying on probability theory, to be made.

More involved situations may require some inaccuracy. Fuzzy set theory or the derived fuzzy logic aim to address the problem of lack of accuracy. The reader is referred to the many books on the topic (e.g. Dubois and Prade (1980), Bouchon-Meunier (1993)). A critical survey may be found in IEEE-EXPERT (1994) and new developments in Dubois et al. (1995). In fact, fuzzy sets in AI seem especially to be used to avoid the gaps involved in any discretization process. Fuzzy expert systems or controllers are mainly interesting because they carry out and produce continuous real functions. When the decision at hand is also continuous (for example, decide the amount of increase in the current of an electric motor), fuzzy decision devices are well adapted. However, they do not escape the restrictions made for expert systems, see 3.2.

For the same reason of avoiding discretization gaps in preferences, the fuzzy set theory is now being extended to preference theory (Perny and Roy (1992), Fodor and Roubens (1994)).

3.5 • Case-based reasoning

With respect to decision, one of the most interesting reasoning models emerged recently in AI. This is the case-based reasoning model (Schank, 1982, Kolodner et al., 1985, Kolodner and Simpson, 1989). Here, it is acknowledged that when people reason, they do so on cases as a whole, and not on facts isolated as in expert systems. A case is very similar to an object in programming : it describes, with many attributes, a significant chunk of knowledge. Basically, the first step of case-based reasoning consists of searching among the recorded cases to see whether there exists a case close to the case at hand. This is a kind of (large) pattern matching. If a case is sufficiently close to the case in hand, then the same actions as in this closest case are carried out. If no similarity is found, then the case is subject either to an ad hoc new reasoning or to modification of the recorded near cases (repairing). Moreover, this type of system is often equipped with a learning module. Using this module, when a new case is recognized, this case and its consequences, as well as the results of the subsequent decision are recorded, thereby enriching the case memory (Kolodner, 1983). This kind of reasoning "by similarity" has been recognized for a long time, Gilboa and Schmeidler (1992) quote Hume (1748) and Keynes (1921). It is clear that people draw, by analogy (real or alleged), many conclusions from their experience . *I think that subjective probabilities are also very sensitive to analogical reasoning and could be deduced from cases.*

Up to now, case-based reasoning did not worry too much about future and preferences and has appeared to be appropriate as a substitute for function φ (Section 2). But recently some researchers have tried to introduced preferences in case-based reasoning (Gilboa and Schmeidler, 1992, 1993). Let us briefly outline their ideas. They try to reconcile case-based reasoning and decision theory. In their framework, a case is a triplet (q, a,r) where q is a problem, a is an action and r a result (outcome). Given a similarity real function s(p,q) on the couples (p, q) of problems (more or less, respectively, our current and desirable states), and a utility function U on the outcomes, they suggest that the chosen action be the argument of Max U(a), where U(a), for the problem p at hand and the memory of recorded case M, is

$$\sum_{(q,a,r)\in M} s(p,q) \;\; U(r) \quad .$$

From our point of view these ideas are very appealing because, they introduce preferences and a base of recorded cases, which we do believe are indispensable for practical decision modelling. Gilboa and Schmeidler give axioms ensuring the consistency of their model and the existence of a similarity function. However their model rises many questions which are discussed in their two papers. From my own point of view, one of the weakness of the model is that it does not include explicit reasoning about the future (the look-ahead phase), implicitly the look-ahead reasoning is contained in the similarity function. Nevertheless we think that the propositions of Gilboa and Schmeidler are very interesting to modelize the function φ (Section 2).

4 - Goals and plans

4.1 • What is a goal ? and a plan ?

In Section 2, we defined a goal as an element (possibly a subset) of the outcome set. *The goal is the outcome that the subject wants to obtain*. Assuming that either there is no uncertainty between the decision and the outcomes or that the decision is almost continuous, then there exists, according to our definitions, an action (or sequence of sub-actions) linking the current state to the goal.

The question envisaged by AI, under the name of **planning**, is the following : *given a goal and a current state, find the sequence of actions (or sub-actions) which leads from the current state to the goal. This sequence of actions is called a plan*. Much effort has been made for almost twenty years to provide algorithms performing the planning task thus defined.

A lot of ideas have been put forward in cognitive science to enrich the notion of plan. A goal reduced to a selected outcome, does not involve other important human dimensions, mainly intention and commitment. For Cohen and Levesque (1990) intention is choice with commitment. It is obvious that when we consider that a goal has been selected, as a result of the subject preferences, we cannot precise how much he is self-committed to attain the goal. I incline towards thinking that this strength is already contained in the preferences and/or the utility function if we accept its *cardinality* (which rises more problems than brings solutions, see Pomerol and Barba-Romero, 1993, for various aspects of the question and references). Another modelling of intention, suggested by Doyle (1987), consists of translating the strength of intention into priorities. It is patent that priority reasoning has given good results in many "intelligent" systems, but the distinction between priorities and preferences is not clear. This is similar to the discussion about the respective roles of weights and values in multicriteria analysis. The concept of intention seems to be more a matter of psychology or inter-personal relations, than of decision. In inter-personal decisions, not only intention, but also confidence and the value of the arguments exchanged are also worthwhile to consider. *We restrict ourself in this paper to personal, isolated decision*. Anyway, goals in AI are unaware of the richness and relevance in that matter of multiattribute preference theory.

Actually, the idea of Doyle (1987) is that we should diagnose two states before a decision : the state of the environment and, among all the possible mental attitudes of the agent, the agent's state which is a subset of his mental attitudes. A mental attitude comprises : beliefs, desires and intentions. Beliefs are expressed by likelihoods (possibly probabilities), desires by preferences (possibly utilities) and intention by priorities. With these ingredients Doyle builds a formal model of reasoning and rationality. If we assimilate the usual decision-theoretic views about utility as a synthesis of subject's preferences and priorities, and if we include subject's beliefs into Expectations, Doyle's model is then not too far, but more complicated, from decision-theoretic ideas recalled in Figure 2. Wellman and Doyle (1991) have also tried to enrich the goal concept. They propose (see Dean an Wellman, 1991, Ch. 5) to consider an outcome as a multi-dimensional element just as in multicriteria analysis (Pomerol and Barba-Romero, 1993). A goal is then an outcome which is preferred according to some attributes, the preference being independent of the value of the other attributes. In term of multicriteria analysis a goal is a maximizer for a criterion, whatever the values of the other criteria are. This is a very poor modelling and we wonder what is, in this case, the usefulness of the other criteria.

A goal being the result of the subject's preference, we must also mention that the subject generally does not feel absolutely free. For example, a CEO stating a goal of a market share of 10 %, probably thinks that 20 % would be more desirable but is not possible. He feels himself to be bound by some constraints (his production capabilities, competition, etc...). The exact role of what is generally felt as constraints in expressing both preferences and the set of possible actions is not clear, see Smith and Browne (1993) for a brief survey. It is well known, in multicriteria analysis (for example in goal programming) that people express sometimes their goals under the form of constraints ; consider for example the goal : keep the pollution level under x. In this case, the decision maker will often put forward another attribute as a goal, for example : the goal is to maximize the production while keeping the pollution level under x, the latter is then expressed as a constraint (see Pomerol and Barba-Romero, 1993). Let us finally observe that most often, the constraints bear on the actions, underlining that actually goals and actions are separated for simplicity's sake but have in reality some links. In human reasoning, it seems that people does not separate goals from possible actions. Goals are never stated without more or less explicitly referring to the possible actions. The most useful concept seems to be that of *attainable outcome* which, once more, echos some *scenario thinking*.

A last point, already raised from a practical point of view (Thompson and Tuden, 1959) is the uncertainty about the goals or about the preferences. The point seems to be relevant in many organizational settings, but I am unaware of any formalized model, I wonder if a probability of changing the subject preferences would answer the question.

4.2 • Planning with certainty

Numerous planning algorithms are based on the idea of a regressive search (in term of decision : backward reasoning). A seminal example is STRIPS (see Nilsson, 1980) which starts from the final goal. A stack is then created and, under the goal, are set the sub-goals featuring in the precondition of one of the rules leading to the goal, and so on. Let us give a simple example. We consider the two situations displayed in Figure 6 in the (small) world of the blocks (see Nilsson, 1980).

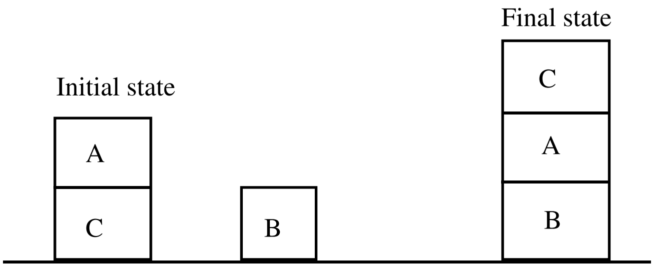

Figure 6 - Block world

Let us assume that the final situation is described by the conjunction of predicates (Nilsson's notation) ON(C, A) $\wedge$ ON(A, B) $\wedge$ ONTABLE(B) ; then STRIPS creates a stack of goals ON(C, A), ON(A, B), ONTABLE(B). To realize ON(C, A) it has to trigger the rule *stack(C, A)* with the precondition and delete lists : HOLDING (C) $\wedge$ CLEAR(A). These are two new goals and the stack is now HOLDING(C), CLEAR(A), ON(C, A), ON(A, B) ONTABLE(B), STRIPS continues until the preconditions of a rule match the current state (here the initial state). Then the rule is triggered, a new state is reached and the new head goal of the stack is considered.

We will not enter any further into the details of STRIPS. We hope that it suffices to understand, *1) that a goal is exactly an outcome to be reached (assuming that every block construction is attainable) and 2) that a regressive search consists of managing a goal stack which is initiated by the final goal, and fulfilling the sub-goals one after the* *other*. Of course, this type of reasoning has some shortcomings which are well-known in AI and have been widely discussed. The main problem is with conjunctive goals, i.e. when the fulfilment of a sub-goal impedes the realization of another sub-goal.

From the decision point of view it is clear that STRIPS functions in an ideal small world of certainty and objectivity, because no event can disturb the robot and moreover it has no preference ; the goal is set. In a decision framework, on the other hand, the choice of outcome is not separable from the choice of actions. Nevertheless, regressive analysis has its virtues and is probably a basic mechanism in decision theory. Value-driven, or more exactly *outcome-driven thinking seems to be very human*. If we not care how the chosen outcome is brought about, it makes sense to start from the goals. "I want to take fifty per cent of the market". "I would like an income of 100 000 $", etc. One of the disavantages of this way of thinking is that it becomes purely foolish when the goal is not attainable. On the other hand, if it is attainable, it suffices to find the path to it (plan) and, if possible, a certain path (against every defense) ; otherwise the plan having the largest probability of success.

So when the goal is attainable, planning is not so unhelpful at may seem at a first glance. But it has another weakness : it assumes that nothing changes during the execution of the plan and this raises the problem of planning horizon (a well-known problem in economics). To what extent does it make sense to plan in an uncertain world ? Where is the equilibrium between reactiveness and planning ? To answer these two questions AI has developed two responses : the first being planning under uncertainty or decision-theoretic planning and the second, reactive systems. In what follows, we will only consider the former aspect and we will not try to enter into the large world of reactive systems and distributed AI (DAI). But remember that when going through a crowded room, you are able to combine a plan to move towards the exit and reactive behavior to avoid collisions. In fact, some results obtained in DAI are somewhat puzzling because a meta-behavior may emerge from purely reactive agents (as was shown by simulation of artificial ants (Drogoul, 1993)). From the outside of the system this meta-behavior appears as a plan. The same thing occurs in economics where purely individual reactions result in what might be taken for a purposeful global behavior.

Before leaving this paragraph, let us mention that a process of *"universal sub-goaling and chunking"* was proposed by Newell (see Laird et al., 1986 and Newell, 1990) as a universal system of cognition. To demonstrate his idea, Newell designs a system, called SOAR, intended to show intelligent behavior in a very large variety of tasks. We refer to this system because, in a sense, SOAR generalizes STRIPS and is one of the first attempts to introduce preferences and rationality into planning.

SOAR principles are "universal subgoaling" and "chunking". The first one is a generalization of regressive thinking involved in STRIPS : each goal generates sub-goals which in turn become new goals the system tries to achieve. The second idea "chunking" is inspired by the observation that the mind learns until it reachs a steady state of performance (Seibel law of practice, see Newell, 1990). In order to learn the system records fixed sequences of rules (chunk of knowledge) it has discorvered to be successful while solving a given problem. We recognize here the same problematic we have put forward for decision. A problem is a difference between two states. When SOAR meets an impasse, it has to decide in which problem space it has a chance to solve it. It then institutes the difference it has to reduce as a goal, and it performs a heuristic search in the chosen problem space (Figure 7).

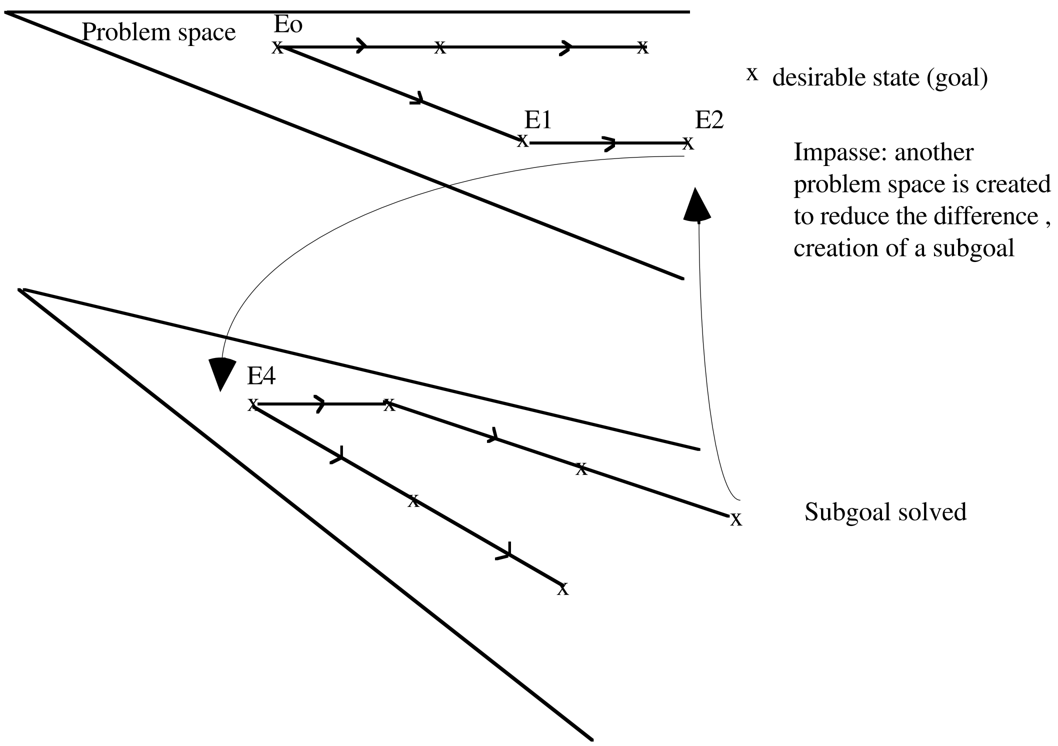

Figure 7 - Problem solving with SOAR

Once it has found a solution, it records the path (chunk) and it returns to its original problem space. In SOAR any current situation is modelled by : 1) a goal, 2) a problem space, 3) a state, 4) an operator. There are, accordingly, four levels of decision : choose the goal, the problem space, the state, the operator. A chunk is characterized by a current state, a problem space and a goal. When the system meets a current state and a goal already recorded (into right problem space) it triggers the corresponding chunk simulating a short cut in reasoning resulting in a saving of time.

Strictly speaking, there is no preferences in SOAR, but there are many choices (goal, problem space, state, operator). The rationality in SOAR is expressed by production rules which produce, during the quiescence phase, *"augmentations"* to the goals, states and operators. These augmentations indicate the relevance of the objects for solving the task at hand, it is a kind of operational preference ; the values of the possible operators are therefore evaluated before firing. Afterwards (decision phase), the most preferred action (i.e. the most effective for the problem at hand) is chosen and executed.

Although, Newell (1982) advocates that the knowledge level is, among other thing, a level where the rationality of the subject is carried out, there is no subjective preferences in SOAR. The only preferences in SOAR are objective and operative, for example, choosing the most promising operator. SOAR is substantively in the line of using a heuristic evaluation function h, such as in A* type algorithms, where the function h indicates the most promising path.

Once more, as a decision system, SOAR is very primitive but it puts forwards and demonstrates the operativeness of regressive reasoning tied to learning. A chunk being an acquired short cut in the reasoning, one may think that many such short cuts exists also into decision making ; a subject is then able to "automatically" make a decision when he recognizes a current state, a difference and a problem space. Thus, the so-called context in SOAR (goal, problem space, current state, operator) is very similar to a case in case-based reasoning (see sub-section 3.5).

4.3 • Planning under Uncertainty

Planing under uncertainty amounts to chose a policy subject to modifications of the environment (events). Formally the situation is the same as the one described in decision trees (Section 2). A plan reachs its target (goal) with a probability depending on the policy chosen. The planner tries to make the decision which maximize this probability. Thus, it is not surprising that decision-theoretic planning heavily relies on decision theory (see Dean and Wellman, 1991). Actually, there are two views about planning under uncertainty the first one relies on various nonmonotonic, especially temporal, logics and the second on the theory of decision. Despite sustained efforts of the researchers, nonmonotonic logics have not proven their ability to handle other knowledge than those contained into toy problems, we shall therefore not emphasize this point. We think that "logic is not to be confused with human thinking" (Simon, 1991). Moreover, we share the view of Simon (1991, p. 192) that formal logic, even with default or temporal reasoning, is not an appropriate language for AI programs, especially when some decision making is involved. We think that is an awfull mistake to pursue the search for logic languages as a base of practical human reasoning ; but it does not exclude the possibility to use formal language as a normative base or to computerize, in a concise and effective way, what has been learnt.

One can find in Langlotz and Shortliffe (1989) a pertinent discussion about the respective advantages and disavantages of nonmonotonic (temporal, with default) logics and of the theory of decision. It is argued that if, in some cases, nonmonotonic logics allow some measure of likelihood on the inferences (which may be interpreted as a plausibility in an uncertain framework), they generally contain no reference to preferences. On the other hand, let us mention another interpretation of default logic. Doyle and Wellman (1989) have noticed that expressing a default rule in Reiter style such that P : Q/R (read : if P is believed and (not Q) is not believed, then conclude R), may be interpreted as expressing preferences between propositions (when P is believed the subject prefers to believe R than believing (not Q) which itself is prefered to .(not Q).and R).Using Arrow's theorem, Doyle and Wellman (1989) have concluded that "there is probably no universally acceptable method for rationally resolving conflict in default reasoning" (Doyle, 1991).Thus, finding a logic for planning under uncertainty with preferences seems, at least, very difficult. One of the interesting features of decision theory is that it merges uncertainty and preference. Referring to human reasoning this seems to be very appealing. On the other side, it is clear that default reasoning cannot be seriously promoted for practical use if there is no link between the defaults and their utility for the decision maker. In many circumstances defaults are more important than rules for the user. Thus, as concluded by Langlotz and Shortliffe (1989, p. 94) : "nonmonotonic logics are most appropriate when the utilities are relatively unimportant (that is low-stakes situations when the decision maker is relatively indifferent to the consequences of action)". This are rather rare situations in human decision.

For an introduction to decision-theoretic planning, the reader is referred to Dean and Wellman (1991, Ch. 7 and the references therein). They extend the framework of probabilistic networks (see sub-section 3.4) to planning and present an application to robot navigation. Briefly, probabilistic netwok approach merges action nodes, chance nodes and utility of the terminal nodes as in decision trees (see Section 2). Accordingly the calculation on these networks follows the rules used on decision trees with, moreover, evidence propagation. Evidence propagation is the basic procedure in probabilistic networks to propagate, along the network, a new information relative to an event. Nevertheless, the burden of calculation remains heavy for coping with real robot planning and probably more efforts are necessary to arrive to a fully satisfactory approach, including temporal reasoning. Already some Markov temporal probabilistic networks (Dean and Wellman, Ch. 7.9) seems to be a very promising direction.

In all these models the utility function is considered as being given encompassing the tradeoffs made by the decision maker between his multiple desires. It is probably one of the weakness of these approaches, as far as it is clear that a skilled decision maker try to escape to early choices in order to keep the possibility to arbitrate at the latest moment between different plans able to result in different multi-attribute "satisficing". As far as I know, no model incorporates this basic feature of human decision : to decide at the latest moment to retain as many possibility of change as possible.

5 - Looking ahead and decision machines

5.1 • Decision machines

A decision machine is an automaton performing the one to one correspondance between the diagnosed current state and an action. As we have recalled the word "decision" is, in this case, improper because the decision has been made by the dessigner. But when people are unaware of the program or when it is so sophisticated that it is almost impossible to foresee its issue, one can talk of these machines as decision machine. A decision machine may therefore be regarded as a program performing the function φ of Figure 5. Numerous such machines already exist in the context of almost continuous decision (control of industrial processes, underground or train driving, and so on). We have already seen that, in this context, the look-ahead phase is very reduced.

But, even with a programmed decision machine relating the current state to an action, one do not capture all the complexity of human decision and such artifacts may have some undesirable effects. In human decision, it may happen that nobody wishes, for various organizational reasons, to materialize into a program the decision process. This is a first obstacle on the decision machine road. But I think that the main difficulty is that in many concrete situations the set of all the possible current states cannot be described neither extensionally nor intensionally. Thus, the human decision maker is always indispensable, working in an interactive way with the machine, mainly because not expected (not programmed) states might occur. Many studied emergency or accident cases were provoked by a bad recognition of the current state (diagnosis) by the subject (Boy, 1991). The designers of decision support systems are therefore confronted with the paradoxical question of developing systems able to help people in situations that neither the user nor the program expect. This is probably, with uncertainty, one of the most difficult challenge for the development of decision support systems.

5.2 • "What if" analysis

Although various previous attempts, have been made to cope with uncertainty, many decision makers are not satisfied with the proposed methods. In most cases the dissatisfaction stems from what we have identified as **look-ahead reasoning** (Figure 2). The main cause of dissatisfaction is that, in real situations, the events are either very interdependent and/or the probabilities remain unknown (e.g. what is the probability that the price of oil will be higher or lower than today in three months?). Another difficulty is forecasting or even identifying all the possible reactions of the other agents, especially in a competitive setting. We have already mentioned that the consciousness of the future seems to be a phylogenetic acquisition. The capacity for anticipation seems to exist only in most advanced animals. Even for human beings, the capacity for anticipation and the ability to decide against immediate optimization to allow future gains seems to be, at least, one of the components of intelligent behavior. Does this component is correlated to calculation skills and/or language is, I think, an open question.

The basis of this anticipating behavior seems to be a capacity of the human mind to situate itself in a non-existing (future or past) situation and perform a "what if" analysis (a cliché of DSS literature, see Lévine and Pomerol, 1995 for an AI perspective). Thus it is not surprising that we arrive at what was regarded, for a long time, as the main function of decision support systems (DSS). *I claim that the basis of human ability to perform look-ahead reasoning is "what if" analysis*. The same idea may be viewed from the point of view (see section 2 for a definition) of scenario scrolling. This is the popular reasoning of the type : "If I do that, they will react like this, and I will do this if the price (or whatever else) has increased or that if it has decreased", and so on. What is important in scenario reasoning is to be able to develop many scenarios and assess, at least roughly, their probabilities (which is an impossible task in the mind). Thus, *supporting people in decision making amounts to mainly help them to foresee the consequences of a choice* (Hatchuel, 1994).

In any case, "what if" analysis or more accurately "scenario reasoning" should produce two outputs : all the possible outcomes at a given horizon and the probability or plausibility of each outcome. The decision maker exercising his preferences on the probabilistic outcomes (preferably multi-attribute), then makes his decision and triggers his actions in accordance with the chosen scenario. Let us notice that dealing with probabilistic outcomes suppose that the subject is able to make tradeoffs between the value of a given outcome and its risk (this tradeoff is generally hidden in the utility function). Unfortunately for non-aided decision makers, scenario reasoning is very combinatory and it is almost impossible to handle long, precise and diverse scenarios. This is the reason why machines are necessary.

5.3 • Look-ahead machines

Two capabilities appear to be necessary in a look-ahead machine : *1) the ability to combine many actions and events (with probability or whatever similar) ; and 2) the ability to imagine the possible actions and all the possible reactions of the other agents and of nature*. In the scheme of Figure 2, this "imagination" ability is simply provided by the file of recorded states. One may estimate that there is nothing new under the sun, and that for a given subject, all possible events and reactions of the other agents are drawn from a mind-recorded stock. In any case, it has been claimed that forecasts never foresee what is really new (see Hogarth and Makridrakis, 1981, Makridakis, 1985 and Mumpower, 1987), so we cannot hope that our look-ahead machine might escape this weakness. It would be nice if it could foresee what has already happened at least once somewhere ! Another reason for using recorded states is that human forecasts are often too optimistic because human beings keep in mind more easily success than failures (Kahneman and Lovallo, 1993).

With this restriction that the basis for scenario building is recorded actions, events and situations (let us say states), what candidates are there for look-ahead machines ? At first glance they are of two types : simulation machines and DSSs.

A simulation machine (simulator) is a machine in which a real industrial or social process has been modelled on a reduced scale (concepts or devices as nodes, links, flows as arcs between the nodes, treatments tied to the nodes). Then, some initial data and parameters are fed into the simulator and the user observes the evolution of the variables he is interested in. One of the most interesting features of this technology is that some variables characterizing the uncertain events are randomly fed into the system according to a given probability law. The simulator then produces random variables which can be observed ; for example mean and standard deviation.

Simulation seems to be the only way of looking ahead when it is impraticable to completely model the process with hard modelling (equations, rules, etc...). This impossibility generally stems from the entanglement of causes and consequences. It is very difficult to arrive at a satisfying model of a process possessing many intertwined feedback loops. The problem of feedback loops in look-ahead reasoning deserves some brief comments. It has been observed many time that subjects are insensitive to the implication of feedback when medium or long delays occur between a decision and its effects (see Kleinmuntz, 1985 ans 1993, and Sterman, 1989 and the references therein). Decision makers generally fail to see the ramifications of a decision as soon as a delayed feedback occurs. One of the only way to sensitize the decision makers to feedback is to use system dynamics initiated by J. Forrester (see for example Wolstenholme, 1990). Studying the dynamics of the variables of the model via simulation, allows a progressive understanding of how a modification of the input variables leads to very counter-intuitive effects, due to not obvious feedbacks. In case of delayed or involved feedbacks, assuming that no causal model exists, nobody can really anticipate the future of the system without simulation. But when it is possible to develop at least a partial model of the given process, it is possible to exploit it by realizing a DSS.

After many years of experience in the DSS field, I am quite convinced that *DSSs are look-ahead machines*. As has already been stated, a DSS is a multi-model, interactive system used by a decision maker to perform an exploration (Bonczek et al., 1981 and Lévine - Pomerol, 1989, 1995). This exploration, or heuristic search, occurs at several levels. At the input data level, it is called "what if" or sensitivity analysis. Roy (in Courbon et al., 1994) has also advocated another "what if" analysis called *robustness analysis*. In robustness analysis the question raised by the decision maker is "What input modifications can a decision endure before really becoming bad ?" It differs from sensitivity analysis, which looks for the maximal modification of the input which leaves a decision unchanged. For an example of robustness analysis see Pomerol et al. (1995).

At the model level, the heuristic search allows the decision maker to explore various variants or diverse models to look ahead among the many possible situations that may occur. In a DSS the model as a whole (or metamodel) is never complete, it is completed by the decision maker interaction (Pomerol, 1992a and 1992b). In a DSS, one operation is especially left to the decision maker, this is to perform the numerous evaluations that occur along the exploration process. In evaluating the plausibility of a model and expressing preferences about different outcomes, the decision maker expresses his preferences and he is the only person able to do that. In some sense a DSS designer acknowledges that the preferences involved in the decision maker's mind are certainly multi-attribute (which excludes any simple utility function) and personal. Many multicriteria DSSs (for a survey see Pomerol, 1993) have been designed which do not incorporate this idea and tried to model and impose an aggregation function in order to make the decision. Very few of them are in use. Rather than focusing on the choice, designers would do better to make richer scenarios by being able to produce and to handle complex actions and situations (see Pomerol (1993) for a discussion and Pomerol et al. (1995) for an example).

Let us illustrate the above ideas by a final example. Assume that you are the CEO of a small company trying to make a decision about a big investment. To do that, you use a spreadsheet (spreadsheets are prototypal and most use DSSs). The first task of our CEO is to make a model of his business, he then explores various sets of data and parameters and evaluates the result according to several attributes (gains, productivity, payback time, etc...). He evaluates each situation and will make his decision (for example adopt the investment budget) according to the data which produces the most "satisficing" outcome, having made a tradeoff between gains and risks (or probability of success). In the case where no set of data is satisficing, the decision maker may continue his reasoning by trying other different models. We have not made mention of the fact that the CEO has to evaluate the quality of the model (this evaluation may depends on his preferences and experience).

This small example shows the things that are modelled in a typical DSS (business, process, etc...) and the parts which are left to the decision maker (mainly evaluation or, in some sense, the expression of his preferences). The contribution of AI to DSS is its abiliity to put forward better and more sophisticated representations allowing more complex states and reasonings to be handled. AI also contributes to the modelling process of unstructured tasks, for example via expert system technique. It has given rise to many "intelligent" DSSs involving knowledge bases (see Turban, 1988, Lévine and Pomerol, 1989, Klein and Methlie, 1990).

AI has yet probably more to offer at the junction between the set of recorded states and the generation of the scenarios. Notice also that this link necessarily encompasses some learning, because it is clear that a review process after decision should lead to an updating of a recorded state bases. Such a learning process plays a significant role in the human mind.

6 - Conclusion and perspective

After this tour around decision and human reasoning, we can draw some conclusions. The first is that decision theory and AI are only at the beginning of their mutual cross-fertilization. The emphasis in decision theory was on making the best decision, either in a certain or uncertain world, with alleged preferences. The look-ahead phase is the one which is mainly considered, but no one is interested in building of the scenarios relating to the subject's experience and memory.

On the other hand, AI has focused much attention on diagnosis and on human knowledge representing and recording. It is only a few years since it first embraced uncertainty, and has pratically never dealt with preferences. In some sense, the planning process, as viewed by AI, starts after the decision is made, since the result of the decision is the goal. It is certainly possible to ignore the preferences and to help a user to find a good path from the current state to his goal, but AI cannot carry out that plan without paying attention to uncertainty. This is the main reason why AI planning programs have been of no use up to now. But, as we have seen, uncertainty planning is now taking off.

Another weakness of AI is its disregard to multi-attribute preferences. It is rather surprising because H. Simon was one of the first, with his bounded rationality model, to aknowledge using by any of us of several criteria at different instants of the decision process and subsequently a kind of tradeoff reasoning to arrive at a satisficing issue. It is obvious than one cannot simulate human reasoning without taking into account tradeoffs between different, not always compatible, aspirations. In the design of realistic interactive system, it becomes more and more frequent that developers introduce multicriteria objectives.

As such, AI and decision theory appear to be mainly complementary : diagnosis representation and handling of the recorded states for AI, look-ahead, uncertainty and (multi-attribute) preferences for decision theory. However, decision theory remains very normative and offers few models which are relative to human reasoning, with the recent exceptions of some non-classical decision models not relying on expected utility (see Jaffray, 1988 and references therein for these models).

But, as proven by DSS experiments, whereas it is important to deal with uncertainty, it is not necessary to have a sophisticated theory of preferences, as long as the system is interactive and that the evaluation is left to the decision maker.The DSS field probably offers the best framework for merging AI and decision theory tools. There is no doubt that diagnosis plus look-ahead machines have a brillant future, if not to mimic human reasoning, at least to support human decision.

References

Auray J.-P., Duru G. and Zighed A., 1990, Analyse des données multidimensionnelles, Editions Alexandre Lacassagne, Lyon.

Barthélemy J.-P. and Guénoche A., 1991, Trees and proximity representations, John Wiley.

Bell D., Raiffa H. and Tversky A. (Eds.), 1988, Decision making, descriptive, normative and prescriptive interactions, Cambridge University Press.

Bonczek R.-H., Holsapple C.-W. and Whinston A.-B., 1981, Foundations of decision support systems, Academic Press.

Bouchon-Meunier B., 1993, la logique floue, Que Sais-Je, Presses Universitaires de France.

Boy G., 1991, Intelligent assistant systems, Academic Press.

Calvin W.-H., 1991, The ascent of mind : ice age, climate and the evolution of intelligence, Bantam Books.

Calvin W.-H, 1994, "La naissance de l'intelligence", Pour la Science n°206, pp.110-117.

Chen T. and Chen H., 1993, "Approximations of continuous functionals by neural networks with application to dynamic systems", IEEE Transactions on Neural Networks 4, pp.910-918.

Cohen P.-R. and Levesque H.-J., 1990, "Intention is choice with commitment", Artificial Intelligence 42, pp.213-261.

Courbon J.-Cl., Dubois D. and Roy B., 1994, "Autour de l'aide à la décision et de l'intelligence artificielle", Rapport LAFORIA/IBP 94/01.

De Geus A.-P., 1988, "Planning as learning", Harvard Business Review, March-April 1988, pp.70-74.

Dean T.-L. and Wellman M.-P., 1991, Planning and Control, Morgan Kaufmann, Los Altos.

Dempster A.-P., 1967, "Upper and lower probabilities induced by a multivalued mapping", Annals of Mathematical Statistics 38, pp.325-339.

Diday E., 1991, "Des objets de l'analyse des données à ceux de l'analyse des connaissances", in Induction Symbolique et Numérique à partir de données, Kodratoff Y. et Diday E. Eds., CEPADUES.

Doyle J., 1987, "Artificial intelligence and Rational Self-government", working paper, Pittsburgh Carnegie Mellon University.

Doyle J. , 1991, "Rational control of Reasoning in Artificial Intelligence", in the logic of theory change, Fuhrman A. and Morreau M. (Eds), lecture Notes in AI, Springer, pp.19-48.

Doyle J. and Wellman M. P., 1989, "Impediments to Universal Preference-Based Default theories", in Proceedings of the first conference on principles of Representation and Reasoning, Brachman R. J., Levesque H. J. and Reiter R., (Eds), Morgan Kaufman, pp. 94-102.

Drogoul A., 1993, De la simulation Multi-Agents à la résolution collective de problèmes, Thèse de l'Université Pierre et Marie Curie, Novembre 1993, Paris.

Dubois D., Fargier H. and Prade H., 1995, "Refinements of the maximin approach to decision making in fuzzy environment", working paper, université Paul Sabatier.

Dubois D. and Prade H., 1980, Fuzzy sets and systems : theory and applications, Academic Press.

Dubois D. and Prade H., 1985, Théorie des possibilités, Masson.Translated in English Possibility theory, An approach to computerized Processing of Uncertainty, Plenum Press, 1988.

Dubuisson B., 1990, Diagnostic et reconnaissance des formes, Hermès.

Duda R.-O, Hart P.-E. and Nilsson N.-J., 1978, "Development of the PROSPECTOR consultation system for mineral exploration", SRI Report.

Fodor J. and Roubens M., 1994, Fuzzy preference modelling and multicriteria decision support, Kluwer.

Gilboa I. and Schmeidler D., 1992, "Case-based Decision theory", Working Paper, Northwestern University.

Gilboa I. and Schmeidler D., 1993, "Case-based knowledge representation : an expository note", Working Paper, Northwestern University.

Hammond K. R., 1987, "Toward a Unified approach to the study of Expert Judgment", in Expert Judgment and Expert Systems, Mumpower J.L., Phillips L.D., Renn O. and Uppuluri V.R.R. Eds., Springer, pp.1-16.

Hatchuel A., 1994, "Apprentissages collectifs et activités de conception", Revue Française de Gestion, Juin-Juillet 1994, pp.109-120.

Henrion M., 1987, "Uncertainty in Artificial Intelligence : is Probability epistemologically and heuristically adequate ?", in Expert Judgment and Expert Systems, Mumpower J.L., Phillips L.D., Renn O. and Uppuluri V.R.R. Eds., Springer, pp.105-129.

Hertz J., Krogh A. and Palmer R.-G., 1991, Introduction to the theory of neural computing, Addison Wesley.

Hogarth R.-M. and Makridakis S., 1981, "Forecasting and planing : an evaluation", Management Science 27, pp.115-138.

IEEE EXPERT, 1994, "A Fuzzy logic Symposium", IEEE EXPERT 9 n°4, pp.3-49.

Jaffray J.-Y., 1985, Cours de DEA, Université P. et M. Curie, Paris, Mimeographed.

Jaffray J.-Y., 1988, "Choice under risk and security factor : an axiomatic model", Theory and Decision 24, pp.169-200.

Kahneman D. and Lovallo D., 1993, "Timid choice and bold forecast : A cognitive perspective on risk taking", Management Science 39, pp.17-31.

Keeney R.-L., 1988, "Value-Driven expert systems for decision support", Decision support systems 4, pp.405-412.

Keeney R.-L., 1992, Value-Focused thinking, Harvard University Press.

Klein M. and Methlie L. B., 1990, Expert Systems, A decision support Approach, Addison-Wesley.

Kleinmuntz D.-N., 1985, "Cognitive heuristics and feedback in a dynamic decision environment", Management Science 31, pp.680-702.

Kleinmuntz D.-N., 1993, "Information processing and misperceptions of the implications of feedback in dynamic decision making", System Dynamics Review 9, pp.223-237.

Kolodner J.-L., 1983, "Maintaining organization in a dynamic long term memory", Cognitive Science 7, pp.243-279.

Kolodner J.-L. and Simpson R.-L., 1989, "The MEDIATOR : analysis of an early case-based problem solver", Cognitive Science 13, pp.507-549.

Kolodner J.-L., Simpson R.-L. and Sycara-Cyranski K., 1985, "A process model of case-based reasoning in problem solving", Proceedings of IJCAI-85, pp.284-290.

Laird J., Rosenbloom P. and Newell A., 1986, Universal subgoaling and chunking, Kluwer.

Langlotz C.-P. and Shortliffe E.-H., 1989, "Logical and Decision-theoretic methods for planning under uncertainty", AI Magazine 10 n°1, pp.39-47.

Latour B., 1992, "Des quelques services rendus par l'IA à la philosophie des sciences", in l'Intelligence Artificielle, une discipline et un carrefour interdisciplinaire, Sayettat C. Ed., Compiègne, pp.55-65.

Lauritzen S.-L. and Spiegelhalter D.-J., 1988, "Local computations with probabilities on graphical structures and their application to expert systems", Journal of the Royal Statistical Society 50, pp.157-194.

Lévine P. and Pomerol J.-Ch., 1989a, Systèmes interactifs d'aide à la décision et systèmes experts, Hermès, Paris.

Lévine P. and Pomerol J.-Ch., 1989b, "Knowledge representations by Schemata in Financial Expert Systems, Theory and Decision 27, pp.147-161.

Lévine P. and Pomerol J.-Ch., 1995, "The Role of the Decision Maker in DSSs and Representation levels", in proceedings of the 26 Th Hawaï International conference on System Sciences, volume 3, Numaker Jr. J. F. and Sprague R. H. (Eds), IEEE, pp. 42-51.

Lewis P.-J., 1991, "The decision making basis for information systems : contribution of Vickers concept of appreciation perspective", European Journal of Information Systems 1, pp.33-43.

Makridakis S., 1985, "The art and science of forecasting : an assessment and future directions", in SIAD 85, Agence de l'Informatique, Paris.

Milgram M., 1993, Reconnaissance des formes, méthodes numériques et connexionnistes, Armand Colin.

Millot P., 1988, Supervision des procédés automatisés et ergonomie, Hermès.

Mumpower J.L., 1987, "Very simple expert systems : an application of judgment analysis to Political Risk Analysis", in Expert Judgment and Expert Systems, Mumpower J.L., Phillips L.D., Renn O. and Uppuluri V.R.R. Eds., Springer, pp.217-240.

Neapolitan R.-E., 1990, Probabilistic reasoning in expert systems : theory and applications, John Wiley and Sons.

von Neumann J. and Morgenstern O., 1944, Theory of games and economic behavior, Princeton University Press, Princeton.

Newell A., 1982, "The knowledge level", Artificial Intelligence, 18, pp. 87-127.

Newell A., 1990, Unified theories of cognition, Harvard University Press.

Newell A. et Simon H.A., 1972, Human Problem Solving, Prentice-Hall, Englewood Cliffs.

Nilsson J.-J., 1980, Principles of artificial intelligence, TIOGA Publishing Company.

Nowicki R., Slowinski R. and Stefanowski J., 1992, "Evaluation of vibroacoustic diagnostic symptoms by means of the rough sets theory", Computers in Industry 20, pp.141-152.

Oliver R.-M. and Smith J.-Q. (Eds.), 1990, Influence diagrams, Belief nets and Decision Analysis, John Wiley and Sons.

Olmeda I. and Barba-Romero S. (Eds), 1993, Redes neuronales articiales, fundamentatos y aplicaciones, University of Alcala Press.

Pawlak Z., 1991, Rough sets, Theoretical Aspects of Reasoning about Data, Kluwer.

Pawlak Z. and Slowinski R., 1994a, "Rough set approach to multi-attribute decision analysis", European Journal of Operational Research 72, pp.443-459.

Pawlak Z. and Slowinski R., 1994b, "Decision Analysis using rough sets", Transactions of Operational Research 1, pp.107-114.

Pearl J., Probabilistic reasoning in intelligent systems : networks of plausible inference, Morgan Kaufmann.

Perny P. and Roy B., 1992, "The use of fuzzy outranking relations in preference modelling", Fuzzy Sets and Systems 49, pp.33-53.

Pomerol J.-Ch., 1992a, "Autour du concept de décision assistée par ordinateur", Revue des systèmes de décision 1, pp.11-13.

Pomerol J.-Ch., 1992b, "Aide à la décision et Intelligence Artificielle", in l'intelligence artificielle, une discipline et un carrefour inter-disciplinaire, AFIA-92, C. Sayettat Ed., pp.147-149.

Pomerol J.-Ch., 1993a, "From aggregating by rules to the integration of expert systems in multicriteria DSSs", IEEE SMC conference Vol.1, Le Touquet, pp.483-488.

Pomerol J.-Ch., 1993b, "Multicriteria DSSs : State of the art and problems", Central European Journal for Operations Research and Economics 2 n°3, pp.197-212.

Pomerol J.-Ch., 1994, Des préférences au choix, des mathématiques à l'intelligence artificielle : le monde de l'aide à la décision, Rapport LAFORIA n°94/20.

Pomerol J.-Ch. and Retour D., 1990, L'introduction et l'utilisation des systèmes experts dans le tertiaire, Etude ADITECH 141, Paris.

Pomerol J.-Ch. et Barba-Romero S., 1993, Choix multicritère dans l'entreprise, Hermès, Paris.

Pomerol J.-Ch., Roy B., Rosenthal-Sabroux C. and Saad A., 1995, "An intelligent DSS for the multicriteria evaluation of railway timetables", Foundations of computer and decision science 20 n°3, to appear.

Rasmussen J., 1980, "The human as a system component", in Human Interaction with Computer, H.T. Smith and T.R.G. Green Eds, Academic Press.

Rasmussen J., 1983, "Skills, rules and knowledge, signals, signs and symbols and other distinctions in human performance models", IEEE Transactions SMC 13, pp.257-266.

Rollier B. and Turner J.-A., 1994, "Planning forward by looking backward : retrospective thinking in strategic decision making", Decision Sciences 25, pp.169-188.

Saporta G., 1990, Probabilités, analyse des données et statistiques, Editions TECHNIP.

Schank R.-C., 1982, Dynamic memory, Cambridge, University Press.

Shafer G., 1976, A mathematical theory of evidence, Princeton University Press.

Shafer G., 1988, "Savage revisited", in Bell et al. Eds., 1988.

Shenoy P.-P., 1994, "A comparison of graphical techniques for decision analysis", European Journal of Operational Research 78, pp.1-21.

Shortliffe E.-H., 1976, Computer-based medical consultations : MYCIN, American Elsevier.

Simon H.-A., 1969, The sciences of the artificial, 2nd Edition 1981, The MIT Press, Cambridge.

Simon H.-A., 1977, The new science of management decision, revised edition, (first edition 1963), Prentice-Hall, Englewood Cliffs.

Simon H.-A., 1983, Reason in Human Affairs, Basil Blackwell, Oxford.

Simon H.-A., 1991, Models of my life, Basic Books.

Slade S., 1991a, "Qualitative Decision theory", Proceedings of Case-Based Reasoning Workshop, Washington, 1991, Morgan Kaufmann, pp.339-350.

Slade S., 1991b, "Case-based reasoning : a research paradign", AI Magazine 12, n°1, pp.43-55.

Slowinski R. (Ed.), 1992, Intelligent decision support. Handbook of applications and advances of the rough sets theory, Kluwer.

Slowinski R., 1993, "Rough set learning of preferential attitude in multi-criteria decision making", in Methodologies for Intelligent Systems, J. Komorowski and Z.W. Ras Eds, Lecture notes in Artificial Intelligence 689, pp.642-651.

Slowinski R. and Stefanowski J., 1994, "Rough classification with valued closeness relation", in new Approaches in classification and Data Analysis, E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, B. Burtochy Eds, Springer, pp.482-489.

Slowinski R. and Zopounidis C., 1994, "Rough-set sorting of firms according to bankruptcy risk" , in Applying Multiple Criteria Aid for Decision to Environmental Management, M. Paruccini Ed., pp.339-357.

Smith G.-F. and Browne G.J., 1993, "Conceptual Foundations of Design Problem Solving", IEEE Transactions SMC 23, pp.1209-1219.

Sterman J.-D., 1989, "Modelling managerial behavior : misperceptions of feedback in a dynamic Decision-Making Experiment", Management Science 35, pp.321-339.

Sternberg R.-J., 1990, Metaphors of mind, conceptions of the nature of intelligence, Cambridge University Press.

Thompson J.-O. and Tuden A., 1959, "Strategies, structures and processes of organisational decision", in Comparative Studies in Administration, J.D. Thompson, P.B. Hammond, R.W. Hawkes, B.H. Junker and A.Tuden Eds, University of Pittsburgh Press.

Turban E., 1988, Decision support and expert systems, MacMillan.

Wellman M.-P. and Doyle J., 1991, "Preferential semantics for goals", in Proceedings AAAI-91, pp.698-703.

von Winterfeldt D. and Edwards W., 1986, Decision analysis and behavioral research, Cambridge University Press.

Wolstenholme E.-F., 1990, System enquiry - A system dynamics approach, John Wiley.

Zadeh L.A., 1978, "Fuzzy sets as a basis for a theory of possibility", Fuzzy Sets and Systems 1, pp.3-28.

Zighed A., Auray J.-P. and Duru G., 1992, SIPINA, méthode et logiciel, Editions Alexandre Lacassagne, Lyon.