A methodology for constructing expert systems

F. T. Attarwala and A. Basden

I.C.I. plc, Decision Support Systems Group, Runcorn, Cheshire, WA7 4QE, U.K.

Abstract

Construction of Expert Systems has so far been seen as a craft or an art, not a science. This paper attempts to improve this situation by deducing some aspects of a methodology for constructing Expert Systems from a model of an expert's expertise. It suggests that the traditional methods, based on extracting problem-solving rules from an expert and encoding them directly into a suitable knowledge representation have certain disadvantages, while attempting to extract a causal model from the expert overcomes many of these. The simple Plausible Inference Expert System shell, of which several are now commercially available, is often very suited to the construction of causal models, and some practical issues concerning the construction of causal models are discussed.

The suggested methodology has been found to work in several Expert System projects in I.C.I., though it is still being developed in the light of experience. In particular, the importance of the top-down design of an Expert System is stressed, though it is not yet based on theoretical considerations.

A lot has been mentioned recently in the media about application of Expert Systems (ES). There is an air of excitement and an expectation of interesting and rewarding applications but so far there are very few ESs in productive use.

Traditionally, ESs have been constructed by eliciting problem-solving rules from an expert and encoding them in a suitable knowledge representation language. Experience in ICI suggests that the use of problem-solving rules is not always the best approach and this paper discusses an alternative approach. The experience has been gained during the construction of several ESs: two which help ICI personnel to make a complex choice, two which advise customers or technical service staff and one which predicts risks of faults.

We present an initial theoretical underpinning of our approach, based on a model of the nature of an expert's expertise. However, as yet, both the experience in ICI and the theoretical underpinning of our approach are limited in scope, and the comments in this paper are offered not as a final solution, but partly to stimulate discussion in these areas.

THE IDEA OF AN EXPERT

The process of building an ES is generally considered to involve taking some of an expert's knowledge and encoding it in a computer for use by others, and, as many bemoan, the process is still only an art or craft, not yet a science. However some progress can be made towards a more substantial foundation by considering the nature of an expert's expertise and its use in consultant or advisory roles.

By an expert we mean someone who has a deep and proven knowledge in a particular domain, and who can be trusted to give advice to solve real problems in that domain. For the sake of brevity we have adopted a simplified but requisite model of an expert's expertise, depicted in Figure 1. This paper does not attempt to justify the model but to explore the implications of it.

We will discuss each of the elements of the model and their relationships to each other as we progress. As shown in Figure 1 the relationship between experience and understanding is essentially a circular one irrespective of the point of start. Many experts develop their real understanding of the domain only after they have had a period of experience; whereas, a fresh graduate (who in years to come may become an expert) entering into industry has some basic

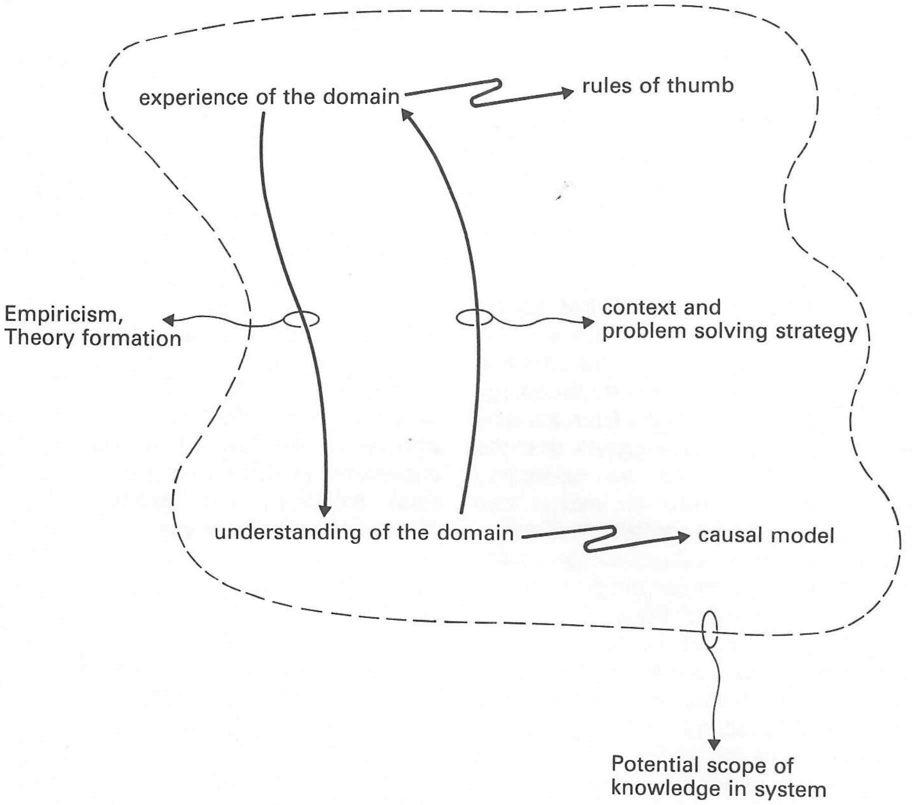

Figure 1 A model of an expert's expertise

understanding of the domain and acquires experience whilst being on the job. In each of these two cases, experience helps to improve understanding, and understanding helps to give a framework for experience.

(What is said about domains of expertise, about experience and understanding, could of course be applied to the domain this paper is discussing, namely the expertise involved in constructing ESs. However in this paper this is not attempted in any formal way.)

Let us now examine how these two components of an expert's knowledge are used.

In practice, experience is largely used to solve real world problems and is based on a rich experience of the real world so that the proffered solution is not only correct but workable. That is precisely the reason why we approach experts and find them valuable. However, in some instances, we seek an explanation of the advice. To be credible, an expert must be able to give a plausible explanation, not only telling us 'what to do' and 'how to do it' but also 'why', and it is largely the understanding of a domain that an expert uses to do this. Of course, the depth of explanation required depends on a number of factors such as the depth of understanding of the advice seeker, the nature of the domain and the problem being solved. We do not give here any hard and fast rules for what level of explanation will suffice but merely say some explanation is required.

PROBLEM SOLVING AND RULES OF THUMB

Now let us consider how an expert solves a problem. An expert will often apply what we call a rule-of-thumb method of problem solving. The rule-of-thumb consists of two main components:

(a) An 'if situation(s) then action/conclusion(s)' pair

(b) together with the context, only some of which is relevant, i.e. the expert can ignore irrelevant facts without necessarily being aware of doing so.

Rules-of-thumb increase an expert's efficiency in coping with a high variety of situations. Often they are a method of approximation, which states that similar actions will achieve similar desired effects (i.e. solve the problem) in not too dissimilar situations. Rules-of-thumb are normally constructed from observations, analogy or with the aid of the expert's understanding of the domain. The last element is discussed below.

UNDERSTANDING OF DOMAIN AND CAUSAL MODEL

When experts are asked to explain the reasons for their advice, they will often do so in terms of cause and effect. This seems to be as true for political or economic domains as for scientific ones. In fact, it seems (ref 1, 2) that causal reasoning is one of the more fundamental human modes of reasoning, and that what we call an understanding of a domain of expertise is largely (though not entirely) composed of a knowledge of causes and effects — a causal model.

Of course, there are major differences between scientific and non-scientific domains. In the domains of physical science the causality has been discovered and verified by rigorous experimentation, and it is thus accepted widely, and it is also of a precise — often mathematical — form. But in the domains of the behavioural sciences this is not so. One's belief that there is a causal link between two items is largely based on a correlation between them and the causality often involves many smaller causal links which have still to be discovered. Thus causality outside the domain of the physical sciences can be seen as a summary of the detailed causality which is actually in operation but perhaps yet to be discovered. This is useful when constructing ESs, and is expanded later.

It should be noted that we use the term 'causal' rather loosely in this paper — much as the man-in-the-street would use it — so that the concept of causality includes that of implication.

PROBLEM SOLVING STRATEGY

What is the nature of the links between understanding and experience in Figure 1? An expert required to solve a problem for which there is no rule-of-thumb will work one out from an understanding of the domain. The nature of this process is not well understood, but we can say it normally requires a problem-solving strategy and a knowledge of the context.

Problem-solving strategies vary among experts since they are based on intuition, past experience, preference and different problem solving skills and they involve a variety of techniques such as analogy, lateral thinking, etc. Contexts also vary, since each expert has, in general, a different territory of problems. Because of this it is not surprising that two experts in a domain often have different sets of rules-of-thumb, leading to the consensus problem mentioned by Barr (Forthcoming).

This suggests that the normal method of encoding the problem-solving rules-of-thumb of an expert can make the construction of ESs more difficult than it needs to be. Fig 1 shows that the potential scope of knowledge that can be encoded in an ES includes more than the experience expressed as rules-of-thumb — it also includes understanding expressed as a causal model, the context and the problem-solving strategy. While the latter, and the nature of their link to rules-of-thumb are not yet well understood, in many cases the causality of a domain is understood to some degree and, as has been found in ICI, useful ESs can be constructed by concentrating on the causal model rather than the rules-of-thumb.

CAUSAL MODELS

We are not saying that ESs should consist only of a causal model, since the ES itself will usually be built to aid problem-solving in a particular context (this is discussed later). But we do believe that Expert Systems can often be more successfully constructed if they are based on a causal understanding of the domain rather than problem-solving rules-of-thumb. We have identified several reasons for this. a) Agreement amongst experts about the causality of a domain is more likely than about rules-of-thumb, especially in the (semi-) scientific domains of interest to ICI. b) As mentioned earlier, explanations tend to be both given and expected in terms of causality as causal-based explanations are of more value to the user than mere reiteration of knowledge base rules, which is an acknowledged problem with ESs based on rules-of-thumb.c) Building a causal model can be easier than building one based on rules-of-thumb because text-books can be used. d) Exceptions can be highlighted and the completeness of the knowledge base can more easily be checked.

It is unfortunate that most current ESs are based on rules-of-thumb rather than causality; CASNET (1978) is one exception and Davis (1983) presents a useful discussion.

What do we mean by a causal model? In a conventional program, the causality is expressed in terms of algorithms, and their repeated application over a grid of space, time, etc. In an ES the algorithm might still be there, but the whole model is expressed in a way that is usually more understandable than in conventional programs — in terms of items and relationships.

Each item represents a fact or numerical value that is relevant to the domain of knowledge, and a causal relationship between two items means that some change in one item causes (or implies) a change in the other. Thus there is a relationship between acidity and corrosion rate which shows that an increase in acidity causes an increase in corrosion rate. Or there is a causal relationship between the assertions 'there are crevices in the metal' and 'stress-corrosion-cracking is likely to occur' which shows that, because crevices facilitate some of the mechanisms involved in stress-corrosion-cracking, an increase in the risk of crevices leads to an increased risk of stress-corrosion-cracking.

As mentioned earlier, causal links do not have to represent the detailed physics and chemistry, but can be a summary of several more-detailed causal links. This is particularly true of economics where the detailed causality is not known, all we have is in the form of observed correlations in behaviour (wherever X happens we normally find that Y will happen) for which there is good reason to believe that there is a causal link. But, even when the detailed causality is known, it may be irrelevant to the objectives of this particular ES and the link we use can be a summary of several known or unknown causal mechanisms. Thus:

$$X \dashrightarrow Y$$

is a summary of

$$X \dashrightarrow A \dashrightarrow B \dashrightarrow C \dashrightarrow Y$$

and

$$\begin{tikzcd} & Y1 \\ X \arrow[ur, dashed] \arrow[dr, dashed] & \\ & Y2 \end{tikzcd}$$

is a summary of

$$\begin{tikzcd} & & Y1 \\ X \arrow[r, dashed] & A \arrow[ur, dashed] \arrow[dr, dashed] & \\ & & Y2 \end{tikzcd}$$

Representing as it does the causality between two items, each link must carry with it some transfer function defining the relationship between the two items. Where the items are numeric values, such as acidity and corrosion rate, the function will be a mathematical expression derived either from experiment or scientific theory. For non-numeric facts the function will either be a logical expression or a mathematical expression of the probabilities of the facts. For a summary link, the function can be worked out if the detailed functions are known, or it can be derived from observation or experiment if not. Among facts Bayesian techniques are appropriate for the latter type of summary link since they are based on probabilities, derived from observations or experiments on populations.

This suggests, and experience in ICI supports it, that the Plausible Inference Expert System (PIES) shells do have an important place in the application of ES technology. This is not only because they offer Bayesian accumulation of evidence, but also because the type of relationship they employ ('evidence for') can be used directly for causality or implication.

Some might object that causal models can best be constructed using conventional computer techniques. This may be true where our understanding of the causality is complete, accurate and unlikely to change, but ES techniques (and especially PIES) offer the various advantages of a) being able to support summarised causality; b) being more easily modified (particularly useful where the causal model may have to be changed); c) being able to easily display the chains of reasoning in the causality.

THE USE OF TEXTBOOKS

Aiming for a causal model opens up an alternative source of expertise that is not normally available when using rules-of-thumb, namely textbooks. Textbooks (of the scientific kind) are written with the purpose of imparting understanding. Thus they can provide ready material for the builder of the ES to gain an initial acquaintance with the domain of expertise.

Textbooks are thus likely to be of most use in the earlier stages of construction. Indeed they can sometimes be a better source than experts in the early stages because many experts operate in a problem-solving manner, seldom having to teach or communicate the more detailed understanding of their expertise. So textbooks can sometimes offer an initial framework so that the expert's problem-solving skill and deeper understanding can be unravelled.

Although textbooks may thus offer definite advantages over some experts in the early stages the knowledge derived from a textbook should always, and particularly in the later stages, be subject to that derived from experts. This is because experts can supply practical knowledge of the situation and this often modifies what the textbooks say. Generally, textbooks give general rules or information but do not give very much context data, while experts not only know such rules but also have at least a feeling for the general shape of the data the ES can expect to receive in the application context, such as the frequency/rarity with which various factors occur, or the average values of numeric items.

It is this general shape of the data which defines the prior probabilities and weights in the Bayesian links found in PIES.

GOAL DIRECTED CAUSAL MODEL

SOPHIE (Brown et al, 1982) employs a causal model, with which it attempts to perform some general reasoning, much as an expert described earlier in the paper. This causal model is devoid of problem-solving strategy because it could be called upon for any purpose. Our experience has not been with such general purpose models but with what we call goal-directed causal models, which are adapted to a particular purpose and context and embody a given problem-solving strategy. In such causal models certain items are identified as those whose value or probability we particularly wish to evaluate; and, when using PIES, they are the goal hypotheses. Any part of a general causal model that is not relevant to these goals is ignored, and in some parts a knowledge of purpose and context can be used to increase computational efficiency.

CASNET's causal model is of this form (Weiss et al, 1978).

Thus, in a goal-directed causal model some items are goals and others are evidence. Ultimate evidence is manifested as questions to be put to the user or some other method of obtaining information from outside the system. Having a clearly defined set of goals aids the process of building the knowledge base.

The purpose and context should be identified early on in construction of the ES by a Top-Down design process that determines who the likely users will be, what their requirements and expectations are and in what situations they will use the ES. Indeed it can be argued that the lack of such Top-Down Design is a major factor in explaining why so few ESs are yet in regular use.

CONSTRUCTION OF THE KNOWLEDGE BASE

Having found our source of specialist expertise, how should we build the knowledge base? We have found the following method to be generally successful when using PIES and, although it may be known to many, it may be of interest to those just starting with the technology.

Many ESs have several stages of evaluation, such as assessments of the risks of various diseases followed by the choice of treatment. For each such stage we usually find there are three steps in construction, which are not necessarily of equal length:

1 Decide the goal hypotheses for the stage. These will be determined by the purpose of the ES determined in the design phase.
2 Construct an evidence net (as a causal model) for these goals. See below.
3 Attend to those parts that will affect how the ES will appear to the User, such as the structure that controls which parts of the net are activated in which order, the style and wording of the questions to be put to the User, and the facilities for explaining the reasoning to the User. These are not discussed in this paper.

The evidence net is a network of items and relationships showing the causal mechanisms that contribute to the given goals. At the right-hand side (when drawn on a piece of paper) are the goal items and on the left, items which are put to the user as questions. In the middle are all the intermediate items, each normally representing one meaningful causal factor of the domain.

A method that we have found to be successful for constructing the evidence net is to list all the causal factors that might be relevant to the given list of goal items; for a corrosion goal these might be 'pH changes locally', 'ions become concentrated', etc. A brainstorming process or a keyword search in a textbook can be useful here. Then group them together and start drawing arrows among them and from them to the goals which show causality. Then add the questions that would be put to the user to establish these intermediate items.

At some stage we find some causal links that are too detailed and some that are not detailed enough. The former can be summarised as explained earlier but the latter need expanding. A useful method of expanding a link is to ask the expert 'Why does X lead to Y: what mechanisms are involved?'. Because it is natural for an expert to give causal explanations, causal models do not seem very onerous to construct in this way.

SOME ISSUES ON CONSTRUCTING CAUSAL MODELS

There are a number of issues that we have found to be of significance when attempting to construct a causal model using PIES:

(a) Level of detail

Above, it was shown that we can choose the level of detail when building a model, by using causal links that summarise a collection of more detailed ones. But to how much detail should we go? At the one extreme is the detail of the physics and chemistry — or one might even go down to the nuclear physics — but such detail may not be appropriate when trying to get some feel for the risk of corrosion. At the other extreme there is no detail at all, but the questions and the goals are linked directly.

The level of detail should be dictated

by the requirements of the users, as identified during a Top-Down Design process. But, in general we have found that, for an ES which will be used interactively in lieu of experts an appropriate level of detail is one level of causality deeper than most users know. This level seems to give the most useful explanations. Too shallow a model gives no new information and it is thereby less useful, only being usable as a calculator. Too detailed a model does not provide the cognitive hooks needed for knowledge retention, and can irritate the users by the amount of detail, especially when it has to be expressed in scientific jargon that most users will not understand. It may also run more slowly. Perhaps more seriously, a very detailed model can be inaccurate while giving a misleading impression of accuracy because, on the one hand, in many domains not all the causal mechanisms that come into play are exhaustively known, and on the other, detailed models very often require more information than the user can accurately supply. Probabilistic reasoning, based on observed correlations, may be more accurate because it covers all the causal links even though we might not know what they are. Paradoxically, therefore, a rather less-detailed model may be more accurate.

In addition, a constraint is sometimes imposed on the level of detail by what data are available for calculating prior probabilities and weights (see below) or by the presence of feedback loops in the actual causality which cannot be handled by most PIES. (For instance, a transistor with a feedback resistor can be summarised as an amplifier stage.)

(b) Macroscopic vs. microscopic

In some problems the causal mechanisms are microscopic in scope while the effects the users are interested in, and thus the meaning of many of the items in the knowledge base, are macroscopic in scope. For instance, a farmer's crop consists of many individual plants. Generally, in such models, most of the evidence information, such as temperature, will also be macroscopic. How do we link the two types of item?

Usually a macroscopic item can be used as evidence for a microscopic item as long as there are no local variations. If there are, and the local variation is approximately the same for all the individuals, then it can be calculated and superimposed on the macroscopic item before being used in evidence. But if the local variation is different for each individual then it is difficult to see how this can be efficiently represented by current PIES as, strictly, there should be a separate piece of the knowledge base for every individual.

A microscopic item can be used as evidence for a macroscopic item as long as the macroscopic effect is merely the average of all the individual effects. This is usually true, but may not always be so.

(c) Non-numeric facts vs. numeric values

Most causal models have a mixture of numeric values and non-numeric facts. How do we use each as evidence for the other? Fuzzy functions (Zadeh, 1965) are appropriate here. It is also often possible to side-step the problem by redefining numeric items as facts or vice-versa. For instance, the numeric item, 'severity of disease' might be replaced by the probabilistic fact, 'disease will cause a problem'. Whether doing this will distort the knowledge base in a detrimental way depends on the application and the needs of the Users (as determined by the design activity). Much work is still needed in this area.

(d) Time

Especially in predictive models, time is an important factor. How should we deal with it? In conventional causal models a common method is to set up a generalised causal model as a set of algorithms that apply for one time-slice and then apply it repeatedly over a number of time-slices, with input information for each slice coming from both external sources and the results of earlier slices. At the other extreme is the kind of reasoning that a farmer might use: 'We've just had a good spell of rain; that'll put the crop in a good position for a high yield at harvest', which is not only highly summarised but also spans a large slice of time. Often a mixture of the two will be most appropriate.

But what size of time-slice should we use? Using too small a time-slice gives problems similar to using too-detailed a level of causal reasoning. Too small a time-slice may be inappropriate to the use of ESs, future external input cannot always be predicted accurately; and repetitive application means repetitive cumulation of errors. In models where the macroscopic effect is the average of the microscopic effects, each individual may be at a different stage; for instance, in a crop the individual plants will be at different growth stages. So the time-slice must be large enough to span such a range of growth stages. It is of course permissible to have time slices of varying lengths, for instance to reflect the lower predictive accuracy of external input factors further into the future. Also, the causal model may be slightly different for different time-slices.

(e) Bayesian weights and prior probabilities

The numeric factors in Bayesian accumulation of evidence, the prior probabilities of facts and the weights on the evidence links, should strictly be derived from historical observations. That is, because the prior probability is the probability that a fact is true, given no evidence at all about it, it should be calculated from the number of times that fact has been found to be true in the past in a number of situations that are similar to those that will be encountered when the ES is being used. Similarly the weights on the evidence links too should be calculated from historical data, because the weight on the evidence link from X to Y reflects the answer to the question, 'What does the probability of Y become if we know for certain that X is true (false)?'

Strictly speaking, this means that the priors and the weights in the model are only valid for the population of usages expected when the ES was built, and if the pattern of use changes these factors ought to be re-calculated.

However, such historical data are sometimes not available. Very often all that is available is the heavily-summarised information about (say) the severity of disease against temperature, and the internal items of the causal evidence net have often not been the subjects of surveys or experimental observation. In such cases it is necessary to use estimates made by the experts in lieu of historical data. These might be either in the form of probabilities or expressed as subjective statements like 'X has a strong positive effect on Y'.

(f) Rules-of-Thumb

Whilst this paper has deliberately emphasised the advantages of aiming for an ES based on a causal understanding of the domain rather than on problem-solving rules-of-thumb, this should not be taken to the extreme. There will always be places where it is more appropriate to the purpose of the ES to include some rules-of-thumb in the Knowledge Base, perhaps because none of the expected users will be interested in detailed causal explanations, or because the causality is not known. Additionally the task of choosing between a number of items is often better implemented as rules-of-thumb, even though the basis for the choice lies in some underlying (perhaps unknown) causality, viz. the effects of making such a choice. This is because the reasoning in this part of such an ES is in a direction opposite to the causality. However, the underlying causality should still be sought as it will normally be used in justifying the choice. These issues should be considered in the Top-Down Design activity.

CONCLUSIONS

This paper has presented an approach to building Expert Systems based not only on experience but also argued on a basis of a model of the structure of knowledge of an expert. This has been done in an attempt to start moving the process of building ESs from being an art or craft to being a science. The proposed approach and the model has arisen from experience of a number of ESs developed within ICI.

When building ESs there are advantages in not following the normal approach of trying to capture an expert's problem-solving rules-of-thumb but in constructing a causal model. Explanations are of a higher quality, there is less risk of experts disagreeing, and it can be easier to construct such an ES. Building the causal model can benefit from the use of textbooks in many instances.

However the causal model will often not be general-purpose, but goal-directed, adapted to a particular purpose and users. It is important that such issues are considered carefully as a top-down design activity. In our experience we have been able to develop goal-directed causal models using Plausible Inference Expert System shell software fairly easily; though the current software needs to be improved.

ACKNOWLEDGMENTS

The authors wish to thank the Directors of ICI plc. for permission to publish this paper, and M A M Rogers for many constructive comments.

REFERENCES

Barr, A. (to be published). 'Consensus'. See also Hunter, P. (1984). 'When experts beg to differ . . .', *Computer Weekly*. 10th May, p. 16.

Basden, A. (1983). 'On the application of expert systems, *International Journal of Man-Machine Studies*. Vol. 19, pp. 461—477.

Brown, J. S., Burton, R. R., deKleer, J., 'Pedagogical, natural language and knowledge engineering techniques in SOPHIE I, II and III', in Sleeman, D., Brown, J. S. (eds), (1982). 'Intelligent Tutoring Systems' Academic Press. pp. 227—82.

Davis, R. (1983). 'Diagnosis via causal reasoning: paths of interaction and the locality principle', *Proc. AAAI conference. Washington DC. USA*. pp. 88—94.

Eden, C., Jones, S., Sims, D. (1983). 'Messing about in problems', Pergamon Press.

Nakamura, K., Iwai, S., Sawarasi, T. (1982). 'Decision support using causation knowledge base', *IEEE Transactions*. Vol. SMC—12, No. 6, pp. 765—777.

Weiss, S. M., Kulikouski, C. A., Amarel, S., Safir, A. (1978). 'A model-based method for computer-aided medical decision-making, *Artificial Intelligence*. Vol. 11, pp. 145—172.

Zadeh, L. (1965). 'Fuzzy sets', *Information and Control*. Vol. 8.