Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER 137050
UNCLASSIFIED	

TITLE

EXPERT SYSTEMS FOR SHIP NOISE INTERPRETATION

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to: JR

EXPERT SYSTEMS FOR SHIP NOISE INTERPRETATION

Joseph N. Maksym

Defence Research Establishment Atlantic
Box 1012, Dartmouth
Nova Scotia, Canada. B2Y 3Z7

Abstract. This tutorial looks at topics in the design of expert systems for interpreting acoustical signals. It is an elementary introduction to issues in programming methodology, knowledge representation, and reasoning with uncertain evidence and knowledge.

1 INTRODUCTION

One of the most significant developments arising from artificial intelligence research is the knowledge-based, or expert, system. In essence this is a computer program designed to store and use specific factual and problem-solving knowledge acquired from human experts. Early systems of this kind such as **MYCIN** [1], a diagnostic expert system in medicine, and **PROSPECTOR** [2], an advisory program for mineral exploration, introduced many of the basic properties of knowledge-based computer programs, namely:

1) Inference rules obtained from experts in the problem domain.

2) Explanation capabilities in terms understandable to human users.

3) A user-extendable knowledge base.

4) Performance at human expert levels in the task domain.

Experimental knowledge-based systems have been constructed in the underwater acoustic signal analysis domain. These include **SU/X** [3], **SIAP** [4], and **INTERSENSOR** [5]. This tutorial examines the basic principles of knowledge-based systems, and looks at some current research on the inference techniques they use.

H. G. Urban (ed.), Adaptive Methods in Underwater Acoustics, 653 -670.

2 THE SIGNAL ANALYSIS PROBLEM

Figure 1 is a typical starting point for acoustic signal interpretation. It is a spectrogram displaying twenty minutes of acoustic signal from one beam of an experimental sensor array. One hundred sequential *looks*, each a short term average spectrum, are shown.

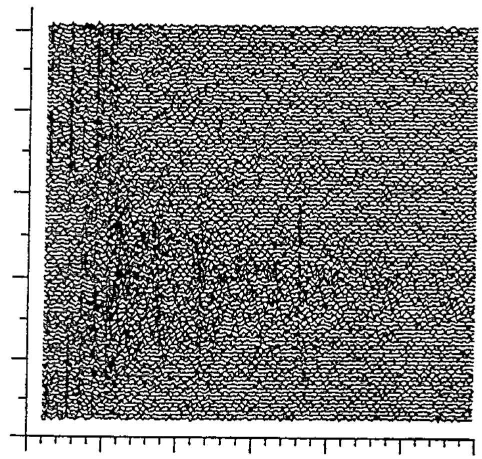

Figure 1. An acoustical spectrogram

In Figure 2 the analysis problem is illustrated by a network model in which causal events generate acoustic manifestations through a multi-level network. Figure 2 is very much simplified as it leaves out several important mechanisms such as the acoustic channel and sources of noise. These effects make the links in the network model uncertain rather than deterministic. The analysis problem is thus one of reasoning with uncertain evidence, and is often made more difficult by the presence of several platforms.

Specific techniques for constructing knowledge-based systems for the analysis problem will now be examined, beginning with the production system.

3 PRODUCTION SYSTEMS

A central concept in many expert systems is the programming strategy known as the production system. First proposed in 1943 by Post as a general computational formalism [6], the idea has seen much recent development as a programming methodology for expert systems. A production system consists of three basic components: 1) A global data base, 2) A set of production rules, and 3) A control strategy.

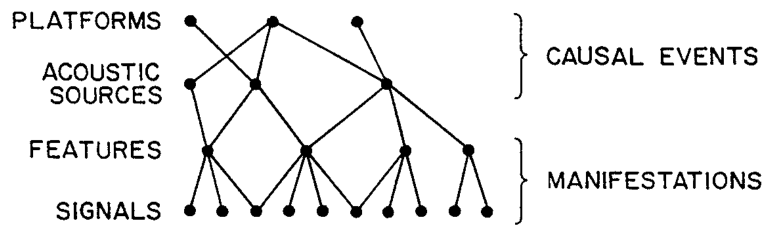

Figure 2. Multi-level data representation

The system operates by applying rules to the global data base in a sequence determined by the control strategy. The global data base may be considered as a kind of *blackboard* that records the state of the problem at any given time. Rules are typically of the form: If *condition* then *action*. If the condition part of a rule matches an assertion in the global data base the action is implemented, resulting in replacement, addition, or deletion of information stored in the global data base. An efficient control strategy is one that generates essentially only those states of the global data base on a direct path to a solution. Expert systems utilize task-domain knowledge to guide efficient application of rules.

Production, or rule-based, systems have several advantages over the procedural style of programming [7]. One is the separation of domain knowledge from the computer code that implements it. Another is the notion of rules as self-contained *chunks* of problem-solving activity. Together, these factors allow system users who are task-domain specialists, but not necessarily programming specialists, to add knowledge to the system and thereby increase its capabilities.

A potential disadvantage of production systems is a decreased visibility of program flow. Theoretically this could be constructed, if desired, by reference to the set of productions and the consequent states of the global data base.

4 KNOWLEDGE REPRESENTATION

4.1 RELATIONAL STATEMENTS

Expert systems require appropriate data structures for representing data and knowledge. A simple example of a data structure is a relational triplet such as: THE *attribute* OF *object* IS *value*. Many rule-based inference systems such as MYCIN represent domain knowledge in the form of facts, observations, hypotheses, and the condition and action parts of inference rules by means of relational triplets or simple extensions of them. To illustrate, consider a simple production system for testing hypotheses about acoustical signals. The top-level conclusions are represented by a list of triplets:

(THE *class* OF *platform* IS *freighter*)
(THE *class* OF *platform* IS *trawler*)
(THE *class* OF *platform* IS *destroyer*)

The global data base stores a list of triplets, which initially consists of observed data, and prior information. The operation of the production rules, as discussed above, updates the set of triplets in the global data base. At some point in the analysis it may contain statements like the following:

(THE *type* OF *engine* IS *diesel*)
(THE *#blades* OF *propeller* IS *four*)
(THE *fundamental* OF *harmonic-set-1* IS *12.8 Hz*)

Several different ways to control the application of rules are possible. One is the backward-chaining approach used by MYCIN. Backward chaining begins by selection of a test hypothesis. This forces selection of a rule whose action part contains that hypothesis. The rule's conditions are then matched against the contents of the global data base. Any condition not found there forces selection of a rule whose action part asserts it. The process unwinds from the hypothesis downward, generating a tree whose branches point to supporting data. Rules can also be applied from the data upward, adding assertions to the global data base until one or more of the top-level hypotheses appears there. In addition both the top-down and bottom-up approaches may be combined, meeting somewhere in the middle.

4.2 FRAME-BASED KNOWLEDGE REPRESENTATION

Simple data structures such as the object-attribute-value relations above have shortcomings as general datatypes for knowledge representation: 1) They lack organizational structure, and 2) They lack representational flexibility. These are provided by frame-based knowledge representations.

A number of knowledge-representation systems are based on frames. There is not space here to discuss the similarities and differences among them. Rather, the basic ideas will be illustrated with some simple examples from ATHENA, a frame-based knowledge representation system that we are using in our expert system implementations [8]. In ATHENA the data structures are called units. A unit has a set of slots for storing information about itself and about its links to other units. Figure 3 is an illustration of a unit that represents information about fishing trawlers. Some of the slots contain simple facts while others link to other units. For example, the slot named **BEARING**

```
Unit: TYPE-X-TRAWLER
  SIGNALS:          [A SIGNAL]
  BEARING:          [0 360]
  DRIVE-TRAINS:     [A DRIVE-TRAIN with
                      ENGINES: [An ENGINE with
                                 #-CYLINDERS: 8
                                 STROKE: 2]
                      PROPELLERS: [A PROPELLER with
                                    #-BLADES: 3]
                      MODE: [DD]
  RANGE:
  SPEED:            [0.0 15.0]
  LOAD:             [EMPTY FULL]
```

Figure 3. TYPE-X-TRAWLER unit

stores the information that bearing falls within the range [0, 360], while the slot called **DRIVE-TRAINS** links the unit **FISHING-TRAWLER** to a unit that describes a type of drive train. By storing information about objects, or concepts, together in this way storage and retrieval of information is facilitated.

Frame-based representation systems facilitate organization of objects into a generalization hierarchy as in Figure 4. Here each unit describes a more general concept than units directly below and to the right of it. Two kinds of hierarchical link are shown. The first kind expresses the fact that one object is a subclass, or specialization of another. The second specifies that an object is a specific instance of a class. For example, in Figure 4 the unit **TYPE-X-TRAWLER1**, embedded in asterisks, is an instance, or individual trawler. In such hierarchies descendents may inherit slots and slot values from ancestors. Frame-based systems usually offer one or more mechanisms to effect the automatic inheritance of slots, restrictions, and values. Such mechanisms simplify the construction of knowledge bases.

4.3 REPRESENTATION OF MULTIPLE PLATFORMS

Frame-like data structures such as UNITS lend themselves readily to interpretation of signals from multiple platforms. Figure 4 represents a typical situation, where several platforms generate an observed set of manifestations in a spectrogram. Given a specific class of platform, its possible manifestations can be described by a network of units associated with that platform. Conversely, a given manifestation can be linked to a set of platforms as possible causes.

An interpretation, at least for the case of manifestations observed with complete certainty, can be obtained via the covering principle in set theory

```
OBJECT
    MACHINE
        PLATFORM
            SURFACE-VESSEL
                MERCHANT
                    FISHING-TRAWLER
                        TYPE-X-TRAWLER
                            *TYPE-X-TRAWLER1*

    DRIVE-TRAIN
    SOURCE
        ENGINE
        PROPELLER
        SHAFT

SIGNAL
    HARMONIC SET
    LINE
    TRACK-SAMPLE
```

Figure 4. Generalization hierarchy

[9]. Having observed manifestations m_1, m_2, ... ,m_n, candidates for the explanation are sought among those sets of causes that cover the observed set of manifestations in the sense that the causes produce at least the observed manifestations. Reggia, Nau, and Wang [9] suggest that the best explanation is the minimal set of causes that covers the observed manifestations.

In practice the situation is rarely that straightforward. Rather than a small set of manifestations observed with complete certainty, there will usually be a much larger set of manifestations with varying degrees of equivocation. Finding the best explanation requires informed search as well as mechanisms for reasoning with uncertain evidence.

5 REASONING WITH UNCERTAIN EVIDENCE

Uncertainty in expert systems arises from noise and distortion, uncertainty in subjective inference rules obtained from human experts, and the use of qualitative or linguistic terms often used by human experts.

Ordinary (Bayesian) probability theory is a well-known tool for dealing with noise and uncertainty in many areas of research such as signal detection and pattern recognition. For a number of reasons, Bayesian statistics were rejected by the builders of early expert systems. For example, in MYCIN [1] certainty was modelled by numerical factors designed to approximate human subjective belief. More recent studies [10], [2], [11], [12], [13] have renewed interest in Bayesian techniques. To overcome some of the shortcomings of ordinary probability theory some newer theoretical frameworks have been proposed for use in expert system inference. Among them are the Dempster-Shafer theory [14], [15], [16], [17], possibility theory [18], and fuzzy set theory [19], [20]. Prade [21] gives a comprehensive, but terse, review of these techniques and their application to modelling of uncertainty in expert systems.

Several workers have criticized the numerical approach and have suggested some alternatives [22], [23], [24].

In the following subsections several different models for certainty, or belief, are described. To do this it is useful to look at a simple inference tree such as that in Figure 5. The nodes h, x, y, r, s, t, u, and v represent evidential statements that are coded in a form similar to: THE *attribute* OF *object* IS *value*. The tree structure in Figure 5 is implemented by the rules:

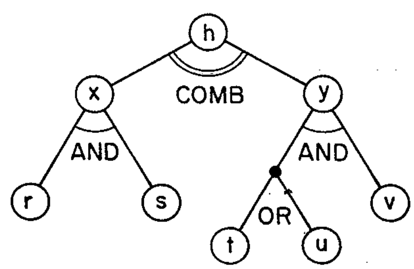

Figure 5. An inference network

R1: IF x THEN h with c_1

R2: IF y THEN h with c_2

R3: IF r AND s THEN x with c_3

R4: IF (t OR u) AND v THEN y with c_4

The symbols c_1, c_2, c_3, and c_4 are numerical certainty factors assigned by an expert. They measure the degree of belief that the conclusion of a rule follows if its conditions are satisfied.

5.1 Subjective Certainty Factors

In MYCIN the certainty factors are not modelled by conditional probabilities since subjective belief is not always consistent with the Bayesian requirement that $P(h|e) + P(\neg h|e) = 1$ be satisfied.[1] Human experts generally do not wish to say that the same evidence that tends to confirm a hypothesis also confirms its complement [25], [26]. MYCIN defines a measure of belief MB(h,e), and a measure of disbelief MD(h,e), which are contributed to by separate pieces of evidence. Certainty measure of a hypothesis is then the combined effect of confirming evidence and disconfirming evidence as follows:

$$c_{h,e} = MB(h,e) - MD(h,e) \ . \tag{1}$$

The intuitive relation of a certainty factor $c_{h,e}$ to prior probability of a hypothesis P(h) and the posterior probability P(h|e) is as follows:

$$c_{h,e} = \frac{P(h|e) - P(h)}{1 - P(h)} \quad \text{if } P(h|e) \geq P(h) \ . \tag{2}$$

$$c_{h,e} = \frac{P(h|e) - P(h)}{P(h)} \quad \text{if } P(h|e) < P(h) \ . \tag{3}$$

Further, when $P(h) = 1$ then $c_{h,e} = 1$, and when $P(h) = 0$ then $c_{h,e} = -1$.

Certainty factors associated with rules are part of the prior knowledge built into an expert system. When the system is used to interpret data, certainty values become associated with the network nodes. These dynamic certainty values are propagated by means of combining functions.

Figure 5 illustrates typical combinations of evidence that occur in an inference network. They are the logical AND, the logical OR, and the combination of independent pieces of evidence denoted by COMB. Assume that the certainty of nodes r, s, t, u, and v is known and given by c_r, c_s, c_t, c_u, and c_v. The certainty of the nodes x, y, and h is computed by propagation of the numerical values through the use of combining functions. The propagated value of c_x is computed by

$$c_x = c_3 \max[0, \min(c_r, c_s)] \ . \tag{4}$$

Similarly a propagated value for c_y is given by

$$c_y = c_4 \max\{0, \min[\max(c_t, c_u), c_v]\} \ . \tag{5}$$

1. The symbol $\neg$ stands for complementation.

The certainty factor of a rule is a multiplicative factor for the certainty of the evidence in the condition part of the rule. Propagation of certainty over logical AND and OR functions is handled in MYCIN by the min and max operators respectively. Note that R3 and R4 are confirming in their action. Consequently the values of c_x and c_y are not permitted to become negative in Equations 4 and 5.

The COMB symbol in the network of Figure 5 indicates the combination of independent pieces of evidence that relate to a single conclusion. To compute the resulting certainty requires a consistent combining function of the form:

$$c_{h,x,y} = f\ (\ c_{h,x}\ ,\ c_{h,y}\ )\ . \tag{6}$$

MYCIN uses an approximate combining function that tends to model human intuition [1]. The form of this combining function for the case where both $c_{h,x}$ and $c_{h,y}$ are positive and less than unity is given by

$$c_{h,x,y} = c_{h,x} + c_{h,y}\ (1 - c_{h,x})\ . \tag{7}$$

The numerical certainty measures and combining functions described above have been used successfully in expert systems patterned after the MYCIN model. They have the advantage of easy computation, and do not disrupt the modular flavour of inference rules. There are shortcomings, however. Errors can occur since $c_{h,x,y}$ computed by Equation 7 is not related to $P(h|x,y)$ in the way required by Equation 2 and Equation 3. An alternative combining rule, consistent with the Bayesian updating scheme in PROSPECTOR [2], has been developed by Ishizuka, Fu, and Yao [27]. The major objection, however, is to the the ad-hoc nature of MYCIN-like belief propagation methods, which limit fundamental understanding of the inference process, and may limit acceptance by intended users [12]. There is therefore ample justification in looking at formulations of inference that have firmer theoretical foundations.

5.2 Bayesian Techniques

Consider a set of mutually exclusive and exhaustive hypotheses h_1, h_2, ... , h_n and related events e_1, ... , e_m. The most likely hypothesis is the one that maximizes the conditional probability $P(h_i|e_1,\ldots, e_m)$. In theory, then, a general Bayesian implementation requires the computation of a set of conditional probabilities. This set of conditionals is large, of the order of 2^m for binary-valued events. In practical implementations of Bayesian inference, such as PROSPECTOR [2], several simplifying assumptions are made. First, a tree structure is assumed. The tree structure naturally separates events into local event groups of interacting events. A second simplifying assumption is conditional independence under a given hypothesis. Conditional independence allows the posterior probability of the root node h in Figure 5 to be expressed as

$$P(h|x,y)=\frac{P(h|x)P(h|y)}{P(h|x)P(h|y)+O(h)P(\neg h|x)P(\neg h|y)P(h)} \quad (8)$$

where, $O(h) = P(h)/P(\neg h)$, and $\neg$ represents complementation.

Equation 8 is the updating rule in the early version of PROSPECTOR [2]. As noted by Konolige in Appendix D of [2], the above result is a special case applying only for the situation with two hypothesis and two evidences. Confusion over this fact has resulted in a certain amount of discussion concerning the validity of Bayesian updating of mutually exclusive and exhaustive hypotheses under assumptions of conditional independence [28]. We can show, however, that the general form is

$$P(h_i|e_1, e_2,\ldots, e_m) = \frac{P(h_i)^{1-m} \cdot \prod_{j=1}^{m} P(h_i|e_j)}{\sum_{k=1}^{n} P(h_k)^{1-m} \cdot \prod_{j=1}^{m} P(h_k|e_j)} . \quad (9)$$

This can also be expressed in sequential form as

$$P(h_i|e_1, e_2,\ldots, e_m) = \alpha \cdot P(h_i|e_1,\ldots,e_{m-1}) \cdot \frac{P(h_i|e_m)}{P(h_i)} , \quad (10)$$

$$\text{where,} \quad \alpha = \sum_{k=1}^{n} P(h_k|e_1,\ldots,e_{m-1}) \cdot \frac{P(h_k|e_m)}{P(h_k)} .$$

The expressions above converge to unit posterior probability for the true hypothesis, and to zero for the rest, with the accumulation of independent pieces of evidence. Dependent evidence in general requires the joint probability, but can often be dealt with by combining dependent propositions through logical AND and OR, as is done in PROSPECTOR and MYCIN.

We summarize this section by noting a renewal of interest in probabilistic models for expert system inference. This has led to a number of research papers on Bayesian approaches, some of which are listed in the references. Some recent work by Pearl [12] and by Kim and Pearl [13] describes propagation of belief, as measured by posterior probability, in tree structures. This work brings up the possibility of a distributed computing approach, since the updating is effected by each node passing information to immediate neighbours up and down the tree. Evidence diffuses through the network in a single pass without infinite relaxation or instability.

5.3 Dempster-Shafer Theory of Belief

Several alternatives to the classical Bayesian approach to quantifying uncertainty have been proposed recently. One of these is the Dempster-Shafer belief theory [14], [15], [16], [17]. We will examine how it differs from the Bayesian approach and illustrate its use in an acoustic signal interpretation system.

Let $\Theta = \{ A_1, A_2, A_3 \}$ be the set of exhaustive and mutually exclusive elementary events, or outcomes, of an experiment. The set Θ is called the frame of reference. The definition of probability as a measure assigned to sets allows one to compute the probability of various compound events defined by subsets of Θ .

The difference between Bayesian and Dempster-Shafer theory shows up in the way evidence contributes probability mass to events. In the Bayesian theory all of the probability is focussed on the elementary events. The posterior probability must satisfy

$$P(A_1) + P(A_2) + P(A_3) = 1. \qquad (12)$$

The posterior probability of compound events such as $A_1 \cup A_2$, for example, is constrained by and computable from the posterior probabilities of the elementary events. Dempster-Shafer theory removes this constraint. A unit of basic probability mass can be divided among elementary and compound events. For example, the following assignment is possible:

$$m(A_1) + m(A_2) + m(A_3) + m(\Theta) = 1. \qquad (13)$$

In the extreme situation of total ignorance, where there is no evidence upon which to base estimates of prior probability, the whole of the unit probability mass can be assigned to Θ. This is a more natural expression of a true state of ignorance than an equal division of prior probability among the events A_1, A_2, and A_3.

Evidential support [2] for an event is defined as the sum of probability mass

$$s(A) = \sum_{B \subseteq A} m(B) . \tag{14}$$

In addition to support, another function called plausibility is defined by

$$p(A) = 1 - s(\neg A) . \tag{15}$$

The interval between support and plausibility is a measure of ignorance or uncertainty, usually referred to as u(A). These new measures carry more information concerning uncertainty throughout an inference network than is possible with a single Bayesian probability.

To illustrate further the Dempster-Shafer approach consider a decision problem in acoustic signal analysis. Figure 6 illustrates a set of events relevant to the analysis. The nodes A through H are elementary events comprising a frame of reference. Nodes to the right are compound events comprising unions of elementary events. Thus, {subsurface} is the union of {diesel} and {nuclear} while {platform} is the union of all the elementary events. Dempster-Shafer theory allows a knowledge source to allocate probability mass among all the nodes. Thus, inference rules do not have to be specific and can point to composite events such as {surface} as well as more refined ones such as {type-x-trawler}. Support at any node in the decision network is computed according to Equation 14 by adding the probability masses on the tree rooted at that node, including the mass at that node itself.

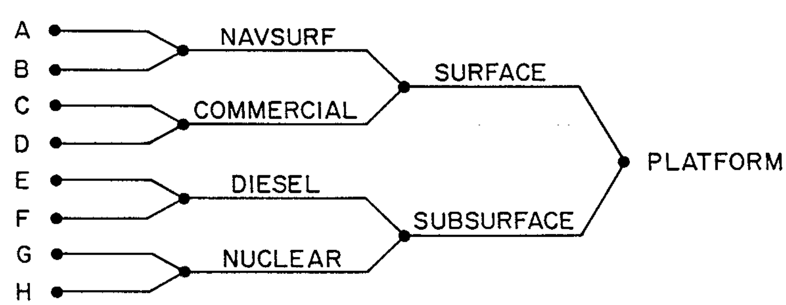

Figure 6. Acoustic decision network

2. Support is a specialized but very useful belief function [14].

To be really useful a theory of belief requires a means of integrating distinct bodies of evidence. This is accomplished by Dempster's rule of combination [14]. Suppose a knowledge source KS_1 contributes $m_1(X)$ to the nodes in Figure 6, where X is a variable representing the event associated with a node. Also let an independent source of knowledge KS_2 contribute $m_2(Y)$, where again Y represents an event associated with a node. The combined mass of an event, say Z, is computed by

$$m(Z) = K \cdot \sum_{X \cap Y = Z} m_1(X) \cdot m_2(Y) , \qquad (16)$$

where, $K^{-1} = 1 - \sum_{X \cap Y = \varnothing} m_1(X) \cdot m_2(Y)$.

We leave the reader to pursue further details in Shafer's text [14], or in [15].

For networks like that in Figure 6 the computational problems associated with integration of evidence from multiple knowledge sources can be simplified [29].

5.4 Fuzzy Sets

Expert systems are often required to process subjective, qualitative, or linguistic data or knowledge. A formal way to cope with these concepts is the theory of fuzzy sets [19]. According to the theory, if x is a support set of a universe of discourse (for example, the cruising speed of all ships), then the degree of membership in a fuzzy set F= *{fast ships}* is described by a set membership function $\mu_F(x)$ which can assume values in the range [0,1].

A calculus for logical combinations on fuzzy sets can be defined as follows [19], [20]:

$$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)] , \qquad (18)$$

$$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)] , \qquad (19)$$

$$\mu_{\neg A}(x) = 1 - \mu_A(x) . \qquad (20)$$

A computational scheme for inference rules with fuzzy sets can also be defined [20], [30]. Let x and y represent elements of support sets for fuzzy sets A and B respectively. Then the inference rule: $A \rightarrow B$ with certainty $c_{y,x}$ defines a fuzzy relation R given by

$$\mu_R(x,y) = \min[1, 1 - \mu_A(x) + c_{y,x} \cdot \mu_B(y)] . \qquad (21)$$

The fuzzy relation can be used to infer a fuzzy set B′ from a fuzzy set A′, not necessarily the same as A, by means of a composition rule $B' = A' \otimes R$ as follows:

$$\mu_{B'}(y) = \max_{x}\{\min[\mu_{A'}(x), \mu_R(x,y)]\} . \tag{22}$$

This account of fuzzy sets represents what might be termed the classical theory, largely developed by Zadeh. Alternative developments exist, especially in the definition of the implication relation. The reader is referred to the literature, for example Shefe [31].

5.5 Alternatives to Numerical Measures

Several authors have expressed reservations concerning purely numerical measures of confidence as computational models for inference. One of these is Doyle, who points out that numerical certainties can often be arbitrarily perturbed by as much as 30% without significant performance change. [23]. This would seem to indicate that expert systems derive their reasoning power from something other than precision of their numerical inference schemes. Doyle suggests, in fact, that the numerical value in a rule of the form: If A then B with certainty 0.8, obscures useful information about specific cases where the implication is not true. He suggests that in a really complete system these *defeasibility conditions* would need to be added anyway [23]. A more extensive discussion of Doyle's ideas is given in his paper concerning truth maintenance in inference systems [22].

Other authors, such as Cohen and Grinberg, also argue in favor of non-numerical approaches [24], pointing out that humans have specific knowledge-intensive techniques for dealing with uncertainty and minimizing its effects. To model this they advocate the use of qualifying factors that they call *endorsements*. As data filters through an inference net its endorsements would be carried with it. Knowledge sources would pass these on and add more to the derived inferences. A conclusion's endorsements would, in theory, be more descriptive of the credibility of the whole reasoning chain than is a single numerical belief value.

6 CONCLUSION

Knowledge-based signal analysis draws upon many diverse methodologies. This tutorial account has presented the reader with an introduction to basic techniques such as production rules and frame-based knowledge representation. The main emphasis, however, was on current research on inference with uncertain data and knowledge. Numerical certainty factors, Bayesian statistics, Dempster-Shafer belief functions, fuzzy sets, and non-numerical methods were discussed.

References

1. E. H. Shortliffe, MYCIN: Computer-Based Medical Consultations, American-Elsevier, New York, 1976.
2. R. O. Duda, P. E. Hart, K. Konolige, R. Reboh, *A Computer-Based Consultant for Mineral Exploration,* Final Report on SRI Project 6415, SRI International, Menlo Park, California, 1979.
3. H. P. Nii and E. A. Feigenbaum, *Rule-Based Understanding Of Signals,* Pattern-Directed Inference Systems (eds. D. A. Waterman and F. Hayes-Roth), pp. 483-501, Academic Press, New York, 1978.
4. R. J. Drazovich and S. Brooks, *Surveillance Integration Automation Project,* Proc. A.R.P.A. Distributed Sensor Net Symposium, pp. 119-123, 1978.
5. J. N. Maksym, A. J. Bonner, C. A. Dent, and G. L. Hemphill, *Machine Analysis of Acoustical Signals,* Pattern Recognition, Vol. 16, No.6, 1983.
6. E. Post, *Formal Reductions of the General Combinatorial Problem,* American Journal of Mathematics, Vol. 65, pp. 197-268, 1943
7. R. Davis, and J. King, *An Overview of Production Systems,* Machine Intelligence 8, 1976.
8. C. A. Dent and R. G. Smith, *A Guide To ATHENA,* DREA Technical Memorandum 80/L, 1980.
9. J. A. Reggia, D. S. Nau, and P. Y. Young, *A New Inference Method For Frame-Based Systems,* Proc. 3rd Annual Conf. on Artificial Intelligence: AAAI-83, Washington D. C. 1983.
10. R. O. Duda, P. E. Hart, and N. J. Nilsson, *Subjective Bayesian Methods for Rule-based Inference Systems,* Technical Note 124, Stanford Research Institute, 1976.
11. E. Charniak, *The Bayesian Basis of Common Sense Medical Diagnosis,* Proc. 3rd Annual Conf. on Artificial Intelligence: AAAI-83, Washington D. C. 1983.
12. J. Pearl, *Reverend Bayes on Inference Engines,* Proc. 2nd Annual Conf. on Artificial Intelligence, Pittsburgh, Pa. 1982.
13. J. H. Kim and J. Pearl, *A Computational Model for Causal and Diagnostic Reasoning in Inference Systems,* Proc. 8th Int.Joint Conf. on Artificial Intelligence, Karlsruhe, West Germany, 1983.
14. G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, Princeton, 1976.
15. J. A. Barnett, *Computational Methods for a Mathematical Theory of Evidence,* Proc. 7th Int.Joint Conf. on Artificial Intelligence, Vancouver, Canada, 1981.
16. T. D. Garvey, J. D. Lowrance, and M. A. Fishler, *An Inference Technique for Integrating Knowledge from Disparate Sources* Proc. 7th Int.Joint Conf. on Artificial Intelligence, Vancouver, Canada, 1981.
17. R. A. Dillard, *Computing Probability Masses in Rule-Based Systems,* NOSC/TD-545,Naval Ocean Systems Center, San Diego, California, 1982.
18. L. A. Zadeh, *A Theory of Approximate Reasoning,* Machine Intelligence 9, Horwood Press, Wiley, 1979.
19. L. A. Zadeh, *Fuzzy Sets,* Information and Control, Vol. 8, pp. 338-353, 1965.
20. E. H. Mamdani, *Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis,* IEEE Trans. Computer, vol. C-26 pp 1182-1191, December 1977.

DISCUSSION

Comment: Ph. L. Stocklin

What metric do these expert systems use in their inference procedures?

Reply: J. M. Maksym

Various evaluation functions have been used to make the search processes more efficient. The talk describes some measures that have been used to model subjective belief in the validity of data and conclusions made by these systems.

Comment: Ph. L. Stocklin

It has been claimed by psychologists that humans are almost never conscious of their reasoning and decision-making process and that some sort of Gestalt principle applies.

Reply: J. M. Maksym

That may be true. However it has been found that when human experts are forced to enter knowledge into an expert system they tend to re-evaluate their decision-making process. What usually results is in fact an orderly and structured set of rules.

Comment: G. C. Carter

As systems become more complex, conflict in knowledge bases will be a problem. Has there been much research on this?

Reply: J. M. Maksym

In early expert systems, ensuring consistency was up to the expert putting knowledge into the system. Typically one would have one expert act as a "knowledge Czar", thus ensuring consistency. Ensuring consistency of knowledge bases is a current research problem of considerable importance.

Comment: R. Griffiths

Would you like to comment on prolog?

Reply: J. M. Maksym

I have no personal experience with prolog.

Comment: L. Bjørnø

What is the limiting factor in the wider use of expert systems, in medical use for instance?

Reply: J. M. Maksym

There may be legal considerations that need to be resolved. Also, medical practitioners tend to be highly independent and may not like it to be known that a computer system helps them to make decisions.

Comment: N. Owsley

What is the current state-of-the-art in learning capability? Do we have systems that can expand their knowledge bases?

Reply: J. M. Maksym

The current emphasis appears to be one of getting knowledge adequately represented in order to achieve high performance. Most practical systems do not incorporate self-learning mechanisms.

Comment: Ph. L. Stocklin

The talk implies a decision-theoretic structure. Has there been any significant effort to look at expert systems from a decision theory point of view?

Reply: J. M. Maksym

I think work is now being done in this area by people with decision-theoretic backgrounds. There have been a number of good research papers on inference mechanisms presented at recent conferences on artificial intelligence.