REASONING WITH UNCERTAINTY FOR EXPERT SYSTEMS

Ronald :. Yager
Machine Intelli lence Institute
Iona C llege
New Rochelle, lew York 10801

ABSTRACT

We discuss a methodology for handling uncertain information in expert and other intelligent systems. This approach combines the theories of approximate reasoning and Dempster—Shafer.

I INTRODUCTION

The construction of expert and other intelligent computer systems requires sophisticated mechanism for representing and reasoning with uncertain information. At least three forms of uncertainty can be identified as playing a significant role in these types of systems. The first of these possibilistic uncertainty appears in situations where the value of a variable can only be narrowed down to a set of values one of which is the actual value of the variable. This is manifested by a situation in which we know that a person's age is between 20 and 30. The second kind of uncertainty is related to situations in which there exists uncertainty as to satisfaction of a predicate by an element. This is manifested by concepts which have imprecise or gray boundaries. A very powerful tool for handling this type of uncertainty which also handles the first type of uncertainty is the fuzzy set. The third type of uncertainty is related to situations in which the value of a variable assumes can be modeled by the performance of a random experiment.

The theory of approximate reasoning [1] which is based upon the theory of fuzzy subsets and possibility theory and the Dempster-Shafer mathematical theory of evidence C23 are two important attempts at providing a framework for the representation and manipulation of uncertain information. In this paper we briefly describe an approach to reasoning with uncertain information which is an amalgam of this two approaches.[3]

II CANONICAL REPRESENTATION OF DATA

Assume V is a variable or attribute which takes values in the set X. A canonical statement is a datum of the form V is A in which A is the value of the variable expressible in terms of the base set X. An example of this is the statement John is young. In a situation like this we can represent the value young as a fuzzy subset A of X such that for each $x \in X$, $A(x)$ indicates the degree of membership of x in the set young. The effect of such a canonical statement is to induce a possibility distribution π_V on X such that

$$\pi_V(x) = A(x),$$

in which $\pi_V(x)$ indicates the possibility that V assumes the value x given the knowledge that V is A.

The use of the above type of formalism allows us to represent uncertain knowledge having a possibilistic as well as a fuzzy component. We should note that the certain knowledge that John is seventeen can be represented in this situation by simply making $A=\{17\}$.

If A and B are two values expressible as fuzzy subsets of X and one is asked if V is B conditioned on the fact that V is A one must use the measures of possibility and certainty to answer this question where

$$\text{Poss}[V \text{ is } B/V \text{ is } A] = \pi_{B/A} = \text{Max}_x[A(x) \wedge B(x)]$$
$$\text{Cert}[V \text{ is } B/V \text{ is } A] = 1 - \pi_{B^-/A}$$

In situations in which the knowledge about the value of a variable V contains probabilistic as well as possibilistic and fuzzy uncertainties we must use a more general type of canonical statement which we call a D—S granule.

Assume m is a mapping from the set of fuzzy subsets of X into the unit interval $m: I^X \rightarrow [0,1]$. We call the fuzzy subsets of X, A_i, $i=1...p$, for which $m(A_i) \neq 0$ the focal elements of m. If the following two conditions are satisfied

1. $\sum m(A_i) = 1$ and 2. $m(\emptyset) = 0$

we call m a basic probability assignment function (bpa). A canonical statement of the form

V is M

is called a D-S granule.
In the face of a D-S granule the measures of possibility and certainty become their expected values which we denote as the plausibility and belief measure which are

$Pl(B) = \Sigma_i \Pi_{B/A_i} * m(A_i)$

$Bel(B) = \Sigma_i Cert[A/B_i] * m(A_i)$

More complicated structures can be obtained from these canonical forms. Consider that V and U are two variables taking their values in the sets X and Y respectively. Consider the statement

"if V is A then U is m."

In the above A is a fuzzy subset of the base set X and m is a bpa on Y with focal elements B_i, i=1,2,...p. The above statement induces a D-S granule

U/V is m^*

in which m^* is a bpa on X x Y such that the focal elements of m^* are the fuzzy sets E_i on X x Y in which for each (x,y), $E_i(x,y) = Min[1, 1-A(x)+B_i(y)]$ and $m^*(E_i) = m(A_i)$.

III REASONING PROCEDURE

In this section we shall describe the basic reasoning mechanism used in this approach. We first introduce some useful rules of reasoning

1.Conjunction- Assume m_1 and m_2 are two bpa's on the set X with focal elements $\{A_i\}$ and $\{B_k\}$ respectively. The conjunction of of the two D-S granules

V is m_1 and V is m_2

is a D-S granule

V is m

such that m is a bpa on X in which for each $A \subset X$, $m(A) = \Sigma\, m_1(A_i)*m_2(B_k)$, where the sum is taken over all i,k such that $A_i \cap B_k = A$.

2.Cylindrical Extension- Assume V is m is a D-S granule where m is a bpa on X. Let U be another variable taking values in the set Y. The cylindrical extension of V is m to X x Y is the joint D-S granule

V,U is m^*

where m^* is a bpa on X x Y in which if A_i, i=1,....p, are the focal elements of m then the focal elements of m^* are B_i, i=1,...p, where

$B_i(x,y) = A_i(x)$

and $m^*(B_i) = m(A_i)$

3.Projection- Assume V,U is m is a joint D-S granule on X x Y. The projection of this on X is a D-S granule, V is $m^\wedge$, where $m^\wedge$ is a bpa on X such that if A_i, i=1,...p, are the focal elements of m then the focal elements of $m^\wedge$ are B_i, i=1,...p, where $B_i(x) = Max_y [A_i(x,y)]$ and $m^\wedge(B_i) = m(A_i)$.

The basic procedure for reasoning can be described as follows.

(1). Represent each of data as a D-S granule.

(2). Cylindrically extend each granule so that they are all on the same space.

(3). Conjunct all the individual pieces of data.

(4). Project onto the variable of interest.

A simple example will clarify the procedure.

Example Assume V is a variable which can take its value in the set

X = {1,2,3,4}

and U is a variable which can take its value in the set

Y = { Bob,Jim,Sue,Mary}={B,J,S,M}.

Assume we have the knowledge that " if V is small then there is at least a ninty percent chance that U is a women." Futhermore, let us assume that there exists a probability distribution on X such that p_1=.5, p_2=.2, p_3=.2 and p_4=.1. We are interested in using this data to find the value of U. Our first piece of data can be represented formally as

"if V is small the U is m"

in which m_1 is a bpa such that

$m_1(W) = .9$ and $m_2(Y) = .1$,

where W = {S,M}. Small is a fuzzy subset of X which can be

small = {1/1, 1/2, .5/3, 0/4}.

The above piece of data is representable as conditional D-S granule

U/V is m_2.

m_2 is a bpa on X x Y where $m_2(A_1) = .9$ and $m_2(A_2) = .1$ in which

$A_1(x,y) = Min[1, 1-small(x) + W(y)]$.

A_1 is representable as the matrix

		B	J	S	M
A_1 =	1	0	0	1	1
	2	0	0	1	1
	3	.5	.5	1	1
	4	1	1	1	1

$A_2(x,y) = Min[1, 1-small(x)+Y(y)]$,
however since $Y(y) = 1$ then $A_2(x,y) = 1$.

The second piece of data can be represented as a D-S granule V is m_3 in which $m_3(1)$=.5, $m_3(2)=m_3(3)$=.2 and $m_3(4)$=.1. The cylindrical extension of m_3 is m_4 whose focal elements are C_1, C_2, C_3 and C_4 where the membership function of C_i is

$C_i(x,y) = 1$ for $x = i$
$= 0$ for $x \neq i$.

In addition $m_4(C_i) = p_i$.

The conjunction of these two pieces of data is the D-S granule

V,U is m

with focal elements

$D_i = C_i$ i = 1,2,3,4

D_5 = {1/(1,S), 1/(1,M)}

D_6 = {1/(2,S), 1/(2,M)}

D_7 = {.5/(3,B),.5/(3,J),1/(3,S),
1/(3,M)}
D_8 = {1/(4,B),1/(4,J),1/(4,S),1/(4,M)}
The weights associated with this focal elements are
$m(D_1)=.05$, $m(D_2)=.02$, $m(D_3)=.02$,
$m(D_4)=.01$, $m(D_5)=.45$, $m(D_6)=.18$,
$m(D_7)=.18$, $m(D_8)=.09$
The projection on U is the D-S granule

U is m^*

in which the focal elements are
E_1 = Y = {B, J, S, M}, E_2 = {S, M}
E_3 = {.5/B, .5/J, 1/S, 1/M} and
$m^*(E_1)=$
$m(D_1)+m(D_2)+m(D_3)+m(D_4)+m(D_8)=.19$
$m^*(E_2)= m(D_5)+m(D_6)=.63$
$m^*(D_3)= m(D_7)=.18$

IV ENTAILMENT

A very useful principle in reasoning is called the entailment principle. It allows us to infer that John is tall from the knowledge that he is six feet six. In [4] Yager has introduced an entailment principle for D-S granules.

Def: Assume m_1 is a bpa on X with focal elements A_1,.....A_p with weights $m_1(A_i)$. Let m_2 be another bpa on X with focal elements $B_{11}, B_{12}, \ldots, B_{1,n(1)}, B_{21}, \ldots, B_{2n(2)}, \ldots, B_{p1}, \ldots, B_{pn(p)}$ such that for each i

$A_i \subset B_{ij}$ for all j

and $\sum_j m_2(B_{ij}) = m_1(A_i)$

then we say $m_1 \subset m_2$.

Entailment Principle: Assume $m_1 \subset m_2$ then from knowlege that V is m_1 we can infer the D-S granule V is m_2.

The entailment principle is use for as aid in making inferences and answering questions. The entailment principle can also be used to simplify bpa's to forms that are more comprehendible. This simplification can be either for the purpose of presenting the results of a reasoning process more succintly or for easing the process of evaluating the weights in providing a bpa.

For example, in the case of the problem we just worked out since $E_2 \subset E_1$ = Y we can infer that the result of our reasoning is that the probability that U is a female is at least .63.

A situation in which one could use a D-S granule would be the following. Assume we are interested in John's age, which we will denote as V. Therefore V is a variable which takes its value in the set of integers less then 120, denote this X. We are given the information that John graduated from high school this year. We know that people usually graduate from his at "about seventeen years of age." However there are some people who for various reasons don't graduate at that age. We can use a D-S granule to represent this information. In particular we can say that

V is m,

where m is the bpa with focal elements
A = "about seventeen"
B = "young graduate"
C = "old graduate"
However since $B \subset X$ and $C \subset X$ we can simply use the focal elements A and X and let $m(A) = \alpha$ and $m(X) = 1-\alpha$. In this case α is the value such that "at least α portion of the people graduate at about seventeen."

With the aid of the entailment principle one can get a better understanding of the Dempster rule used by Shafer in combining bpa's.

Assume m_1 and m_2 are two bpa's on X. We previously defined the conjunction of two as $m = m_1 \cap m_2$, in which for each $A \subset X$ $m(A) = \sum m_1(A_i)*m_2(B_j)$, with the sum taken over all A_i and B_k such that $A_i \cap B_k = A$.

Shafer's combination of these two bpa's is $m^* = m_1 \oplus m_2$ in which
$m^*(\emptyset) = 0$
$m^*(A) = m(A)/(1-m(\emptyset))$ for all other A.
We note if $m(\emptyset)=0$ then $m=m^*$.
In the above m^* is obtained by proportionally allocatting the weight in the null set to the other focal elements of m. Since $\emptyset \subset A_i$ the Dempster rule can be seen as simply a conjunction followed by a special application of the entailment principle.

REFERENCES

[1] Zadeh, L.A., "A Theory of Approximate Reasoning." In Machine Intelligence. Vol 9, Hayes, J., Michie, D. & Mikulich, L.I. (Eds), John Wiley and Sons: New York, (1979) 149-194.

[2] Shafer, G., A Mathematical Theory of Evidence. Princeton University Press: Princeton, 1976.

[3]Yager,R. R.,"Toward a General Theory of Reasoning with Uncertainty," Tech. Report# MII-509 and 510, Machine Intelligence Institute, Iona College, 1985.

[4]Yager, R.R., "The Entailment Principle for Dempster-Shafer Granules," Tech Report# MII-512, Machine Intelligence Institute, Iona College, 1985.