Report 84-31
Stanford -- KSL

Scientific DataLink

Implementing Probabilistic Reasoning.
Matthew L. Ginsberg,
Jun 1984

card 1 of 1

Implementing probabilistic reasoning

M.L. Ginsberg
Department of Computer Science
Stanford University
Stanford, California 94305

1. Introduction

I have argued elsewhere [Ginsberg 1984] that it is possible (advisable, even) to reason non-monotonically be assigning to statements such as "Tweety can fly" not specific truth values, but ranges of possible probabilities. My intention here is not so much to repeat the arguments as to describe an implementation of the ideas they entail. Before doing so, let me repeat some of the results of the earlier paper.

We can generalize the usual truth values T and NIL by thinking of them as probabilities of 1 and 0 respectively. The probabilities can be generalized further to *ranges* of probabilities, where we think of a specific probability p as the singleton interval $[p, p]$.

We will denote a probability interval $[a, b]$ (where $0 \leq a \leq b \leq 1$) by the pair $(a \ \bar{b})$, where we have written $\bar{b}$ for $1 - b$, and $a + \bar{b} \leq 1$. Intuitively, a measures the extent to which some given statement is *confirmed* by the available evidence, and $\bar{b}$ the extent to which it is *disconfirmed*. We will call a the *activation* of $(a \ \bar{b})$, and $\bar{b}$ the *deactivation*.

As an example, a true statement will have truth value $(1 \ \ 0)$, corresponding to the probability interval $[1, 1] = \{1\}$. A false statement will have truth value $(0 \ \ 1)$.

Incomplete information can be handled similarly. "Tweety can fly" might have truth value $(.9 \ \ .02)$, for example; the .02 represents the fact that some birds are known *not* to fly. As an example of extremely incomplete information, "Tweety lives in California" may well have truth value $(0 \ \ 0)$. In general, the completeness of the knowledge represented by the truth value $(a \ \ d)$ corresponds to the sum $a + d$; this is called the *mass* of the truth value.

The main attraction of this approach is that it is possible to combine conflicting truth values for the same statement; this gives the method a non-monotonic nature. A formulation due to Dempster [Dempster 1968 or Shafer 1976] gives us, after some work,

$$(a \ \ b) + (c \ \ d) = \left(1 - \frac{\bar{a}\bar{c}}{1 - (ad + bc)} \quad 1 - \frac{\bar{b}\bar{d}}{1 - (ad + bc)}\right). \tag{1}$$

This rule of combination allows us to weigh conflicting evidence with regard to any specific statement. It is commutative and associative, so that conclusions drawn using it

1

will be independent of the order in which the inferences are drawn. It is also invertible, the inverse being given by

$$(a \ b) - (c \ d) = \left(\frac{\bar{c}(a\bar{d} - \bar{b}c)}{\bar{c}\bar{d} - \bar{b}c\bar{c} - \bar{a}d\bar{d}} \quad \frac{\bar{d}(b\bar{c} - \bar{a}d)}{\bar{c}\bar{d} - \bar{b}c\bar{c} - \bar{a}d\bar{d}} \right). \tag{2}$$

This allows us to retract specific inferences without affecting others and provides many of the same features as a truth maintenance system [Doyle 1979]. Finally, since $(0 \ 0) + (c \ d) = (c \ d)$, we can construct our entire theory incrementally, each statement having initial truth value $(0 \ 0)$ (corresponding to complete ignorance).

2. Probabilistic rules.

Suppose, then, that we want to find an incremental truth value for fly(Tweety) from the rule

$$bird(x) \rightarrow fly(x), \tag{3}$$

where the rule itself has truth value (.9 .02). If Tweety is known to be a bird, so that $bird$(Tweety) has truth value (1 0), then the increment to the truth value of fly(Tweety) is clearly (.9 .02). More generally, if $bird$(Tweety) has truth value $(a\ d)$, the increment to fly(Tweety) will be (.9a .02a). Note that the deactivation d does not contribute to this expression, since the rule (3) cannot be used to draw any conclusion fom the fact that Tweety is *not* a bird.

We will think of a rule such as (3) as consisting of a premise and a conclusion ($bird(x)$ and $fly(x)$ respectively in the above example). The truth value of the premise will be called the *application level* of the rule. In (3), the effect of the rule is to increment the truth value of the conclusion by the product of the activation of the application level and the truth value of the rule itself.

Now consider the following:

$$\neg bird(x) \rightarrow \neg fly(x) \qquad (.8\ .1). \tag{4}$$

It would be nice if we could manage to capture the relationship between this rule and the previous one. As a start, we can rewrite (4) in the equivalent form

$$\neg bird(x) \rightarrow fly(x) \qquad (.1\ .8). \tag{5}$$

Recall now that the application level of the rule (3), being the truth value of $bird(x)$, also contains information regarding the premise in (5). This allows us to combine (3) and (5) to obtain the single rule:

RULE1(x)

    premise: $bird(x)$

    conclusion: $fly(x)$ t: (.9 .02)

                                 nil: (.1 .8),   (6)

3

where we have included in the conclusion an indication of which entry in the application level of the rule itself should be used as a trigger.

As a more complicated example, consider the fact that "*is-a*" is transitive:

If all x's are y's, and all y's are z's, then all x's are z's.

If all x's are y's, and no y's are z's, then no x's are z's.

This becomes

$$is\text{-}a(x,y) \wedge is\text{-}a(y,z) \rightarrow is\text{-}a(x,z) \qquad (1\ \ 0)$$
$$is\text{-}a(x,y) \wedge \neg is\text{-}a(y,z) \rightarrow \neg is\text{-}a(x,z) \qquad (0\ \ 1)$$

$$(7)$$

This will resemble (6) if we think of $is\text{-}a(x,y)$ not as part of the rule's premise, but as a *precondition* for it. Thus the transitvity of *is-a* is replaced by:

$$RULE2(x,y,z)$$

precondition: $is\text{-}a(x,y)$

premise: $is\text{-}a(y,z)$

conclusion: $is\text{-}a(x,z)$ t: $(1\ \ 0)$

                                       nil: $(0\ \ 1)$ (8)

The application level of the rule is the product of the activation(s) of the precondition(s) and the truth value of the premise.

If the value of $is\text{-}a(x, \text{birds})$ is $(a\ \ d)$ and that of $is\text{-}a(x, \text{flyers})$ is $(.9\ \ .02)$, we can apply (8) with $y = $ birds and $z = $ flyers; the application level of the rule is $a(.9\ \ .02) = (.9a\ \ .02a)$, and we should therefore increment the truth value of $is\text{-}a(x, \text{flyers})$ by $(.9a\ \ 0) + (0\ \ .2a) = (.9a\ \ .02a)$. Note that the sum used in the intermediate calculation is a simple vector addition, rather than Dempster's expression (1). This is because the different entries in a single conclusion are assumed to be dependent; Dempster's rule assumes that the summands are the consequences of maximally independent information.

4

3. Tags and metarules.

The t and nil appearing in (8) are used to convert the application level of rule 2 to the single value needed to apply the rule. If $(a\ d)$ is the application level of the rule (or any other truth value), we have

$$t : (a\ d) \to a \qquad \text{and}$$
$$\text{nil} : (a\ d) \to d. \tag{9}$$

Any such mapping will be called a *tag*; in addition to the two above, the following are also of interest:

$$m : (a\ d) \to a + d$$
$$u : (a\ d) \to 1 - a - d$$
$$\text{poss} : (a\ d) \to \bar{d} \tag{10}$$
$$\text{poss-not} : (a\ d) \to \bar{a}.$$

The Dempster-Shafer mass is given by m; $u = \bar{m}$ corresponds to the uncertainty inherent in a given truth value. The extents to which the conclusion is possibly true or false are given by poss and poss-not respectively. Note that the six tags consist of the three basic ones t, nil and m, along with their complements.

By including a tag in each precondition to a rule, we can rewrite rule 2 in the slightly more general form

$$\text{RULE2}(x, y, z)$$

precondition: $t(\textit{is-a}(x, y))$

premise: $\textit{is-a}(y, z)$

conclusion: $\textit{is-a}(x, z)$ t: $(1\ 0)$

nil: $(0\ 1)$ (11)

A more valuable use for tags is suggested by the following: Although it is never possible to prove in advance that a rule will be applicable in any given situation, it may

5

be possible to show that it will *not* be applicable. This suggests thut we rewrite rule 1 as:

RULE1(x)

        precondition: poss(RULE1(x))

        premise: $bird(x)$

        conclusion: $fly(x)$ t: (.9 .02)

                                                  nil: (.1 .8),   (12)

The point of this modification is that we can now introduce the following:

RULE3(x)

        premise: $ostriches(x)$

        conclusion: $fly(x)$ t: (0 1)

                         RULE1(x) t: (0 1)   (13)

The first conclusion corresponds to the fact that ostriches can't fly; the second indicates explicitly that rule 1 should not be applied to ostriches.

It is important to realize that the viability of this scheme is a direct consequence of the invertibility of Dempster's rule. It is quite possible that we apply RULE1(x) before RULE3(x). (Indeed, the potential need *not* to apply a rule of questionable validity is a classic problem in the application of any non-monotonic reasoning system.) Having done so, however, we can use (2) if necessary to reverse any conclusions which were inaccurately drawn.

A similar modification can be made to rule 2. The idea is that if Tweety is an ostrich, we should use RULE2(Tweety, ostriches, flyers) in preference to the alternative RULE2(Tweety, birds, flyers). Note that this is not only because ostriches are a subset of birds, but also because rule 2, when applied to (Tweety, ostriches, birds), leads to strong conclusions. The Dempster-Shafer mass is well suited to this, and we have:

RULE4(x, y_1, y_2, z)

        precondition: t(is-$a(y_1, y_2)$)

        premise: RULE2(x, y_1, z)

        conclusion: RULE2(x, y_2, z) m: (0 1)   (14)

6

The application level of this rule is proportional to that of rule 2; the effect is therefore to deactivate $RULE2(x, y_2, z)$ to the extent that $RULE2(x, y_1, z)$ leads to substantive conclusions.

4. Summary and conclusions

At this point, the formalism we have presented is fairly self-contained, and can be used to replace the kernel of an existing derivation system such as MRS [Genesereth, 1984]. The ideas themselves have been implemented on Stanford's SAIL system. (SAIL seems a logical choice because it is virtually the only Stanford system which does not have MRS on it.)

When implementing probabilistic reasoning, it seems attractive to replace the usual truth values T and NIL with ranges of acceptable probabilities. Since Dempster's rule allows us to combine these ranges we can deal with apparent inconsistencies between various (non-monotonic) conclusions.

It seems natural in this framework to combine the premises of a rule of inference to obtain a probability range as well. In order to do this, we have split the premises of a rule into a single preferred one (which we continue to refer to as "the" premise) and a collection of preconditions. The preconditions contain tags which enable us to combine their truth values with that of the premise to obtain an application level for the rule, which is used to evaluate the rule's consequences.

It is also possible to include one or more rules among the preconditions, premise or conclusions of another rule. This extension relies upon the invertibility of Dempster's rule, which allows us to retract conclusions drawn earlier if necessary.

References

Dempster, A.P., "A generalization of Bayesian inference," J. Roy. Stat. Soc. *B*, **30** (1968), 205-247

Doyle, J., "A truth maintenance system," Artificial Intelligence, **12** (1979), 231-272

Genesereth, M., MRS reference ?

Ginsberg, M.L., "Non-monotonic reasoning using Dempster's rule," AAAI-84 (?)

Shafer, G., *A Mathematical Theory of Evidence*. Princeton: Princeton University Press (1976)