

Possibility theory and the generalised Noisy OR model

Simon Parsons and John Bigham
Department of Electronic Engineering,
Queen Mary and Westfield College,
Mile End Road, London E1 4NS,
United Kingdom
S.Parsons@qmw.ac.uk, J.Bigham@qmw.ac.uk

Abstract

The probabilistic Noisy OR model has been widely used as a means of reducing the amount of probabilistic information that is needed to specify the interaction between groups of variables. Motivated by our work on the diagnosis of faults in satellite sub-systems, which required a similar synthesis of partial possibilistic information on interactions, we consider a number of ways of generalising the probabilistic Noisy OR model to possibility theory.

1 INTRODUCTION

One of the major problems in building a probabilistic network model of a domain is that of providing a full set of conditional probabilities $\Pr(x | \mathbf{u})$ relating each possible value x of each variable X to every set of possible values \mathbf{u} of the set of variables \mathbf{U} which have been identified as the direct causes of X . Thus for a variable D , with possible values $\{d_1, d_2, d_3\}$, which has direct causes A , B , and C , which have possible values $\{a_1, a_2, a_3\}$, $\{b_1, b_2, b_3\}$, and $\{c_1, c_2, c_3\}$ respectively, it is necessary to provide estimates of $\Pr(d_1 | a_1, b_1, c_1)$, $\Pr(d_1 | a_1, b_1, c_2)$, \dots , $\Pr(d_3 | a_3, b_3, c_3)$. This is no inconsiderable task, and, as Pearl [9] points out, it is complicated by the fact that the various causes often point to disparate frames of knowledge whose only connection is that they have a common consequence. As a result it is useful to study canonical models of interaction between multiple causes which provide a means of computing the necessary set of conditional probabilities from those in terms of a single cause. In other words these models provide a means of estimating $\Pr(d_1 | a_1, b_1, c_1)$, $\Pr(d_1 | a_1, b_1, c_2)$, \dots , $\Pr(d_3 | a_3, b_3, c_3)$ from the more easily accessible sets of values $\Pr(d_1 | a_1)$, $\Pr(d_1 | a_2)$, \dots , $\Pr(d_3 | a_3)$, $\Pr(d_1 | b_1)$, $\Pr(d_1 | b_2)$, \dots , $\Pr(d_3 | b_3)$, and $\Pr(d_1 | c_1)$, $\Pr(d_1 | c_2)$, \dots , $\Pr(d_3 | c_3)$.

The first such model to be suggested was the Noisy OR model (or Noisy OR-gate to give it the title first

conveyed upon it) [9] which took as its basis a form of disjunctive interaction in which any member of a set of causes makes a particular event likely to occur, and in which this likelihood is not diminished by the occurrence of several causes simultaneously. A suitable example of this kind of interaction is if the event in question is that of a patient suffering from a fever, and the set of causes is a set of fever-inducing diseases. If the patient suffers from one of the diseases then she is likely to have a fever, and if she is suffering from several diseases at the same time then she is yet more likely to have a fever. Pearl shows that with two simple assumptions (that the event does not occur unless there is a cause, and the mechanisms that prevent a cause making the event occur are independent) disjunctive interactions can be modelled quite elegantly. The simple model he provided, which only allows for binary valued causes and effect, has subsequently been adapted and generalised by a number of authors.

Henrion [7] extended the model to cover the case in which the event in question may occur even when all the direct causes identified by the model are absent by including a “leak” probability, showing how this model was useful in building Bayesian networks that solve real world problems. In addition, Heckerman [4] showed that models built using the leaky OR-gate could be solved tractably. In particular, he developed an algorithm that could solve such models and which had a time complexity that was exponential in the number of causes that actually occurred. In collaboration with Shwe [6], the same author has provided empirical justification for the validity of the Noisy OR model by comparing its performance with that of other canonical models of causal interaction.

There have also been further generalisations of the Noisy OR model beyond that of Henrion. Díez [1] generalises the model to take account of events which have more than two values, while Heckerman [3] discusses the use of a Noisy Adder—a model in which the possible values of the event variable are the sums of the possible (integer) values of the variables identified as its direct cause. Finally, Srinivas [10] completes the generalisation of the probabilistic Noisy OR, by allowing multi-valued causes, and permitting the com-

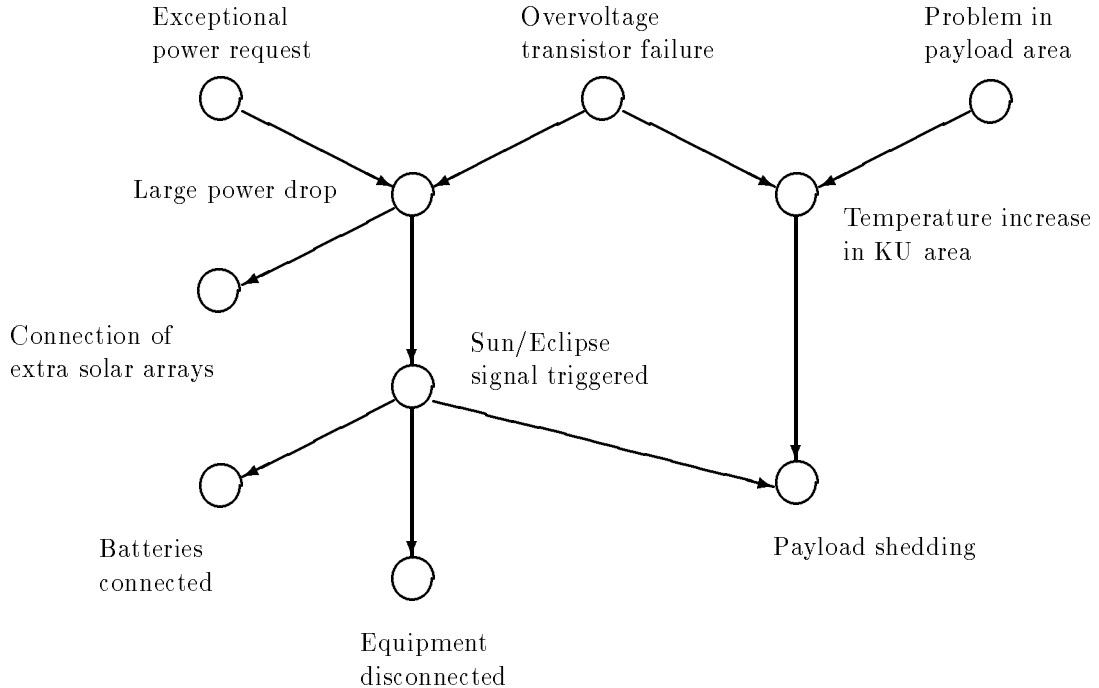


Figure 1: A fragment of network for diagnosing satellite faults

bination of the effect of the causes to be calculated by any discrete function, and Heckerman and Breese [5] go beyond the Noisy OR model altogether by suggesting a new, temporal, model for combining disparate causes.

Our recent work on diagnosing faults in satellite equipment has borne out Pearl’s predictions about the difficulty of assessing the joint effect of a set of causes. For instance, in the illustrative fragment of the knowledge base in Figure 1, the satellite engineers would assess the effects of “Exceptional power request” and “Overvoltage transistor failure” on “Large power drop”, by considering each of the causes independently, leaving the problem of combining their estimates of the effects in a suitable way. Thus it was clear that we needed to use some form of canonical model, and the Noisy OR model was an obvious candidate.

However, the engineers also preferred to express their feelings about the strength of the influences between cause and effect using possibility measures since they felt that the behaviour of possibility measures best fitted their understanding of how they themselves solved the diagnostic problem. As a result we considered a number of ways in which the Noisy OR model could be extended into the framework of possibility theory, and our conclusions are the subject of this paper. It is, of course, possible to extend the other generalisations of the Noisy OR mentioned above into possibility theory, and this is done in [8]

2 PROBABILISTIC NOISY OR

Before launching into a description of how one may build a Noisy OR model in possibility theory, it is worth remembering what the probabilistic Noisy OR model is. The basic situation is depicted in Figure 2. We have a binary variable E , whose possible values are e and $\neg e$, which is influenced by a number of binary causes C_i , each of which has possible values c_i and $\neg c_i$. We have a series of conditional probabilities $\Pr(e | c_i)$, and from these we build up conditional probabilities such as $\Pr(e | c_1, c_2, \dots, c_n)$ as follows:

$$\Pr(e | \mathbf{C}) = 1 - \prod_{c_i \in \mathbf{C}} 1 - \Pr(e | c_i) \quad (1)$$

where \mathbf{C} is a truth assignment to the set of all possible causes of E , and \mathbf{c} is the set of all causes of E that are known to be present. Thus, for instance, $\Pr(e | c_1, c_2, \neg c_3, \dots, \neg c_n) = 1 - (1 - \Pr(e | c_1))(1 - \Pr(e | c_2))$. This can then be extended to take account of the fact that E may occur in the absence of any of the C_i by taking:

$$\Pr(e | \mathbf{C}) = 1 - (1 - p_0) \prod_{c_i \in \mathbf{C}} \left(\frac{1 - \Pr(e | c_i)}{1 - p_0} \right)$$

where p_0 is $\Pr(e | \neg c_1, \neg c_2, \dots, \neg c_n)$ and models the fact that the gate is “leaky” and can give an output even when there are no inputs. This is the model as presented by Henrion [7], and it is the extension of this to possibility theory that is the subject of this paper.

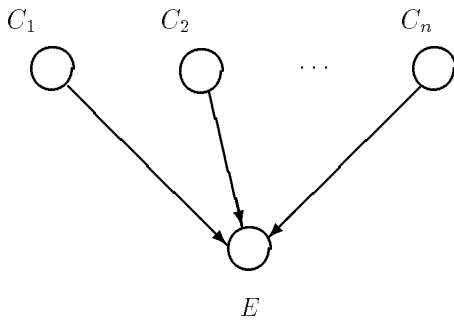


Figure 2: The basic Noisy OR model

3 TOWARDS A POSSIBILISTIC NOISY OR

We approach the construction of a possibilistic Noisy OR in a stepwise fashion. First, within the framework of Zadeh’s possibility theory [11], we consider how a simple noiseless OR model could be built, and then complicate this model by adding in first uncertain causes and then uncertain influences. Then we take Dubois and Prade’s possibilistic logic [2], and show how this may be used, in a number of different ways, to achieve much the same end.

3.1 A NOISELESS OR

In Zadeh’s theory, for each cause, C_i , we have a conditional possibility $\Pi(e | c_i)$, and from these we need to build a conditional possibility $\Pi(e | c_1, c_2, \dots, c_n)$, which should behave in such a way that $\Pi(e)$ is 1 as soon as any $\Pi(c_i)$ is 1. Now, for an OR model, the possibility of e conditional on any single cause, $\Pi(e | c_i)$, must be 1, while the possibility of e conditional on the negation of any cause, $\Pi(e | \neg c_i)$, must be zero. Therefore, a suitable complete conditional distribution may be constructed using:

$$\Pi(e | \mathbf{c}) = \max_i \Pi(e | c_i) \quad (2)$$

taking e given \mathbf{c} to be the disjunction of e given all the separate c_i , a move which seems intuitive. From these conditional possibilities $\Pi(e)$ may be established by using:

$$\Pi(e) = \sup_{\mathbf{c} \subset \mathbf{C}} \min \left(\Pi(e | \mathbf{c}) \Pi(\mathbf{c}) \right) \quad (3)$$

where

$$\Pi(\mathbf{c}) \leq \min_{c_i \in \mathbf{c}} \Pi(c_i) \quad (4)$$

since \mathbf{c} is the conjunction of all the causes which are known to have occurred, and \mathbf{C} , as before, is a truth assignment to the set of all possible causes of E . This clearly gives the necessary result, yielding zero if the possibility of every cause $\Pi(c_i)$ is zero, and 1 otherwise assuming, as we do, that we are dealing with normalised possibility distributions so that either $\Pi(c_i)$ or $\Pi(\neg c_i)$ has value 1.

Now, as it stands, this model is not terribly helpful, since all it does is to model a simple, deterministic, disjunctive relationship using possibility values. Its usefulness is to be found in the fact that it can be easily extended to cover both the situation in which the certainty of the causes is not known, and the situation in which the uncertainty in the cause/effect influence is acknowledged.

3.2 UNCERTAIN CAUSES

The first extension of the simple model which we consider is that to cover the case in which the causes are uncertain. That is, at the point at which we wish to assess the possibility of the event E , we do not have sufficient information to say which of the causes have definitely occurred. For those causes C_j which we know to have occurred, because, for example, we have observed them directly, we know that it is impossible that they take their negative value $\neg c_j$, and so the necessity of those values, $N(\neg c_j)$ must be zero. This then means that the possibilities of those values $\Pi(c_j)$ are known to be 1 by applying the identity $\Pi(x) = 1 - N(\neg x)$.

However, for those causes C_k which have not been directly observed, the only information that we have is some general information about their disposition to occur. This may be suitably modelled by a possibility distribution across their possible values. Since each cause has only two values, c_k and $\neg c_k$, this distribution is rather restricted (because it is normalised) so that at most one of $\Pi(c_k)$ and $\Pi(\neg c_k)$ can be less than 1. Nevertheless, varying the possibility values of some of the c_k allows us to represent what knowledge we have of their relative possibility of occurrence, and these values may be used in (4) to calculate the possibilities of sets of causes, and these latter values may then be used in (3).

3.3 UNCERTAIN RELATIONS

The fact that a given cause, C_i may not be sufficient to ensure that E happens may simply be modelled by allowing $\Pi(e | c_i) < 1$. Thus, considering C_i alone, the observation of c_i , an event which results in $\Pi(c_i)$ being set to 1, makes it less than completely possible that e occurs. Indeed, the possibility of e only rises to a value that reflects the degree to which the occurrence of the cause makes the effect possible. Consider then, the effect of combining several uncertain causal relations in (2). The effect is clearly to make the conditional possibility of the effect given all the causes that of the strongest causal relation which is exactly what we would expect from a disjunctive relation. When we establish all the conditional possibilities and substitute their values into (3) to obtain $\Pi(e)$, the effect is to make $\Pi(e)$ as small as the combination of the strongest link and the most possible set of causes, or to put it another way, it has the possibility value of the least weak combination of cause and influence, and expressed in this way, it is clear that the model has the

right kind of behaviour.

Modelling uncertain relations, of course, gives us a means of including a “leak possibility” if we wish, to model the situation in which e can occur even when it is known that none of the c_i have occurred. To do this we merely give a non-zero value to the conditional value $\Pi(e | \neg c_1, \dots, \neg c_n)$ ($\Pi(e | \emptyset)$ if you like) which by (2) would otherwise be zero, and use this in (3).

We could, of course, also quantify the model of Figure 2 using necessity measures rather than possibility measures. If this were the case the complete conditional distribution would be built using:

$$N(e | \mathbf{c}) \geq \max_i N(e | c_i) \quad (5)$$

and the necessity of e would be established using:

$$N(e) = \inf_{\mathbf{c} \subset \mathbf{C}} \max \left(N(e | \mathbf{c}) N(\mathbf{c}) \right)$$

where

$$N(\mathbf{c}) = \min_{c_i \in \mathbf{c}} N(c_i)$$

Finally, a leak necessity could be introduced by giving $N(e | \emptyset)$ a non-zero value, although the idea of a model in which the leak is somewhat certain to occur is a rather strange one.

It should also be noted that (2) and (5) may be rewritten as:

$$\Pi(e | \mathbf{c}) \geq 1 - \max_i \left(1 - \Pi(e | c_i) \right)$$

and

$$N(e | \mathbf{c}) = 1 - \max_i \left(1 - N(e | c_i) \right)$$

respectively by exploiting the duality of possibility and necessity, to make the connection with (1) quite clear.

4 USING POSSIBILISTIC LOGIC

The approach based on numerical possibility theory discussed in the previous section is, of course, just one of the ways in which the Noisy OR may be addressed in the possibilistic framework. The other obvious approach is to use possibilistic logic [2], modelling the cause and effect as logical propositions, and the influences between them as implications, and that is the approach investigated in this section.

4.1 THE BASIC APPROACH

When using a logical approach we need to modify the model slightly. We could describe the basic noiseless OR model in logical terms with the following statements $\{c_1 \supset e, c_2 \supset e, \dots, c_n \supset e\}$, but it is more correct to describe it as:

$$\begin{array}{rcl} c_1 & \supset & e_1 \\ c_2 & \supset & e_2 \\ & & \vdots \\ c_n & \supset & e_n \\ e_1 \vee \dots \vee e_n & \equiv & e \end{array}$$

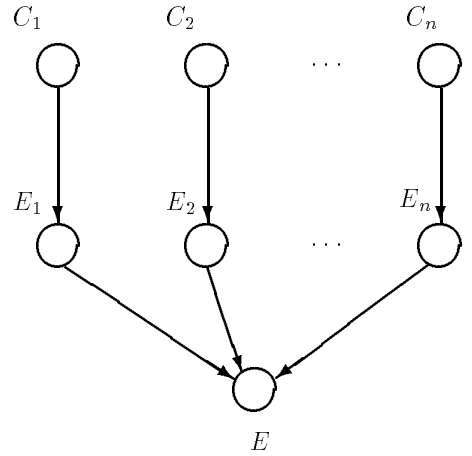


Figure 3: A more general Noisy OR model

making clear the fact that E is the result of some combination of intermediate events E_i which are caused by the C_i . Thus the situation we are modelling is that depicted in Figure 3, which is exactly that used as the basis of the most general Noisy OR models developed using probability theory [5, 6, 10].

With our logical model, as soon as any cause is known to be true, the truth of the event in which we are interested is established. When the both the causes and the implications are less than completely certain they may be assigned suitable possibility or necessity measures, and the resolution principles of possibilistic logic used to perform deductive reasoning to establish the possibility or necessity of e .

For instance, if the causes and implications are quantified with necessity measures then we can use the usual pattern of resolution [2] to establish the necessities of the e_i :

$$N(e_i) \geq \min \left(N(c_i), N(c_i \supset e_i) \right)$$

and these may then be used to obtain an upper bound on the necessity of e :

$$N(e) \geq \max_i N(e_i) \quad (6)$$

Alternatively, if we have a mixture of possibility and necessity measures, we can establish the possibility values of the e_i using another result from [2]. We find that if the causes have necessity values and the implications have possibility values, then:

$$\begin{aligned} \Pi(e_i) &\geq \Pi(c_i \supset e_i) \text{ if } \Pi(c_i \supset e_i) + N(c_i) > 1 \\ &= 0 \text{ otherwise} \end{aligned}$$

while if the causes have possibility values and the implications have necessity values, then:

$$\begin{aligned} \Pi(e_i) &\geq \Pi(c_i) \text{ if } \Pi(c_i) + N(c_i \supset e_i) > 1 \\ &= 0 \text{ otherwise} \end{aligned}$$

The $\Pi(e_i)$ may then be combined to get the possibility of e :

$$\Pi(e) = \max_i \Pi(e_i) \quad (7)$$

Obviously it is also possible to have a situation in which we have possibility values for some of the e_i , and necessity values for some of the others. In such a case, we can always establish a possibility value for e since every e_i with a known non-zero necessity value must have a possibility of 1. In addition we may be able to establish a necessity value for e since if the possibility of e is less than 1, $N(e)$ is zero by definition. Thus, if $e_1 \dots e_j$ have known possibility values, and $e_{j+1} \dots e_n$ have known necessity values, then from (6):

$$N(e) \geq \max_{i=j+1 \text{ to } n} N(e_i)$$

while from (7) $\Pi(e) = 1$, provided that at least one e_i has a non-zero necessity measure, otherwise:

$$\Pi(e) \leq \min_{i=1 \text{ to } j} \Pi(e_i)$$

Finally, it should be noted that the concept of a “leaky cause” could of course be incorporated into this model by introducing a “leaky cause” c_0 which has a possibility and necessity of 1 (meaning that it is always certain to occur), and which has a suitably quantified implication relating it to e_0 .

4.2 USING ASSUMPTIONS

There is a further approach to providing a possibilistic Noisy OR model. This is basically a logical model, but makes explicit use of assumptions of failure. That is, rather than taking the approach of the previous section and writing:

$$\begin{array}{rcl} c_1 & \supset & e_1 \\ c_2 & \supset & e_2 \\ & \vdots & \\ c_n & \supset & e_n \\ e_1 \vee \dots \vee e_n & \equiv & e \end{array}$$

and then handling the fact that a given c_i is not a certain cause of e by attaching a possibility measure to the implication $c_i \supset e$, we instead acknowledge that if c_i occurs but does not cause e , it is because there is some other event δ_i which is defeating the mechanism by which c_i causes e . Thus we have some certain information:

$$\begin{array}{rcl} c_1 \wedge \delta_1 & \supset & e_1 \\ c_2 \wedge \delta_2 & \supset & e_2 \\ & \vdots & \\ c_n \wedge \delta_n & \supset & e_n \\ e_1 \vee \dots \vee e_n & \equiv & e \end{array}$$

and the uncertainty in the model is characterised by attaching a possibility or necessity measure to the defeaters δ_i .

Now, the reason for adopting this kind of model is that in some cases we may need to perform abduction on the Noisy OR model, reasoning from observations of the event E in order to determine something about

the causes C_i . If we do this, then explicitly including the defeaters allows us to establish the reasons for the occurrence or otherwise of E in terms of the occurrence of the C_i and whether or not they have been defeated. When we are reasoning abductively we also need to explicitly take account of the causes of $\neg e$ —those sets of conditions under which the event does not take place—since we will often only have observations of the form “ E does not occur”. In order to use this kind of information, we need to build a model of the causes of $\neg e$. This will clearly be constrained by information about the causes of e .

For instance, consider the situation in which E has just two causes, C_1 and C_2 , so that the model for the occurrence of e is:

$$\begin{array}{rcl} c_1 \wedge \delta_1 & \supset & e_1 \\ c_2 \wedge \delta_2 & \supset & e_2 \\ e_1 \vee e_2 & \equiv & e \end{array}$$

Assuming that these are the only possible causes of E (so that there is no “leak” possibility), we have:

$$c_1 \wedge \delta_1 \vee c_2 \wedge \delta_2 \equiv e$$

which may be negated and manipulated to give:

$$\begin{aligned} \neg e & \equiv (\neg c_1 \wedge \neg c_2) \vee (\neg c_2 \wedge \neg \delta_1) \\ & \quad \vee (\neg c_2 \wedge \neg \delta_1) \vee (\neg \delta_1 \wedge \neg \delta_2) \end{aligned}$$

This expression immediately gives four implications which involve $\neg e$, and are thus suitable for abductive reasoning from the observation that E has not occurred. These are:

$$\begin{array}{rcl} \neg c_1 \wedge \neg c_2 & \supset & \neg e, \\ \neg c_2 \wedge \neg \delta_1 & \supset & \neg e \\ \neg c_2 \wedge \neg \delta_1 & \supset & \neg e \\ \neg \delta_1 \wedge \neg \delta_2 & \supset & \neg e \end{array}$$

Of these only $\neg \delta_1 \wedge \neg \delta_2 \supset \neg e$ does not have an obvious interpretation and this may be transformed [8] into:

$$c_1 \wedge c_2 \wedge \neg \delta_1 \wedge \neg \delta_2 \supset \neg e$$

to complete the logical side of our model. The possibility and necessity measures are then simple to add. We started off with the causes being certain and each defeater having a possibility or a necessity measure. From these values we can easily determine the possibilities and necessities of the negations of the defeaters for the new rules, and from these the measures to be attached to the rules that involve them. The new rules can then be used with the modus tollens forms of possibilistic logic [2] to establish the possibility and necessity of the causes c_i given information about the events e_i .

If both events and implications have necessity values, then the necessity values of the causes may be established:

$$N(c_i) \geq \min \left(N(e_i), N(c_i \supset e_i) \right)$$

If the events have necessity values and the implications have possibility values, then we can establish the possibility values of the causes:

$$\begin{aligned} \Pi(c_i) &\geq \Pi(c_i \supset e_i) \text{ if } \Pi(c_i \supset e_i) + N(e_i) > 1 \\ &= 0 \text{ otherwise} \end{aligned}$$

as we can if the events have possibility values and the implications have necessity values:

$$\begin{aligned} \Pi(c_i) &\geq \Pi(c_i) \text{ if } \Pi(c_i) + N(c_i \supset e_i) > 1 \\ &= 0 \text{ otherwise} \end{aligned}$$

5 DISCUSSION

This paper has examined a number of different approaches to building canonical Noisy OR models in the framework of possibility theory—there are doubtless others since we have not tried to be exhaustive (nor have we space to be)—concentrating upon the use of the two main approaches to possibilistic reasoning, Zadeh’s theory of possibility and Dubois and Prade’s possibilistic logic. In both approaches we began by looking at how a noiseless OR model could be built, and then iteratively complicated the model by adding in uncertain causes and uncertain influences between cause and effect, building up to a model that is sufficiently rich to capture the detail present in most real-world situations.

Whilst there are strong similarities between the models developed using the different approaches, there are slight differences of emphasis. The model built using Zadeh’s theory is intended as a means of constructing full conditional possibility and necessity distributions from a series of partial distributions in order to simplify the building of a possibilistic model. Once this model has been built, the theory tells us how to infer the possibility and/or necessity of the event given the causes. In this respect the intention of model is very like the probabilistic Noisy OR. On the other hand, when the model is developed in possibilistic logic, there are two main points of focus. The first is on how to make the correct combination of the effects of each cause on its own, and the second is how to build a complete model (in the sense that it will support abduction) from a partial model.

Finally, it should be noted that the possibilistic Noisy OR model differs in one important way from the probabilistic Noisy OR model. As stated in the introduction, the development of the probabilistic Noisy OR model was predicated on the idea of causal independence. This assumption is not necessary for the development of the possibilistic Noisy OR model, so that it is possible to model situations in which there is some correlation between the causes. This may be particularly useful when considering how to model the defeaters mentioned in the abduction model, since these may be correlated even when the causes are not.

Acknowledgements

This work was partially supported by ESPRIT Basic Research Action 6156 DRUMS II, and ESPRIT Project 6083 UNITE. Thanks to Philippe Smets for some thought-provoking comments on an early version of this paper.

References

- [1] Diez, F. J. (1993) Parameter adjustment in Bayes networks. The generalised noisy OR-gate, in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington D. C.
- [2] Dubois, D. and Prade, H. (1990) Resolution principles in possibilistic logic, *International Journal of Approximate Reasoning*, **4**, 1–21.
- [3] Heckerman, D. (1993) Causal independence for knowledge acquisition and inference, in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington D. C.
- [4] Heckerman, D. (1990) A tractable algorithm for diagnosing multiple diseases, in *Uncertainty in Artificial Intelligence 5*, M. Henrion, R. Shachter, L. N. Kanal and J. F. Lemmer (eds.), Elsevier Science Publishers B. V., Amsterdam, pp 163–171.
- [5] Heckerman, D. and Breese, J. (1994) A new look at causal independence, in *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence*, Seattle, WA.
- [6] Heckerman, D. and Shwe, M. (1993) Diagnosis of multiple faults: a sensitivity analysis, in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington D. C.
- [7] Henrion, M. (1989) Some practical issues in constructing belief networks, in *Uncertainty in Artificial Intelligence 3*, L. N. Kanal, T. S. Levitt and J. F. Lemmer (eds.), Elsevier Science Publishers B. V., Amsterdam, pp 161–173.
- [8] Parsons, S. and Bigham, J. (1995) Non-probabilistic generalisations of the Noisy OR model, Technical Report, Department of Electronic Engineering, Queen Mary and Westfield College.
- [9] Pearl, J. (1988) *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufman, San Mateo, CA.
- [10] Srinivas, S. (1993) A generalisation of the Noisy-Or Model, in *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington D. C.
- [11] Zadeh, L. A. Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and Systems* **1** (1978) 1–28.