International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems
Vol. 3, No. 1 (1995) 000—000
© World Scientific Publishing Company

FURTHER RESULTS IN QUALITATIVE UNCERTAINTY

SIMON PARSONS
Advanced Computation Laboratory, Imperial Cancer Research Fund,
P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX.

Received (June 1994)
Revised (March 1995)

This paper extends previous work on propagating qualitative uncertainty in networks in which a general approach to qualitative propagation was discussed. The work presented here includes results that make it possible to perform evidential and intercausal reasoning, in addition to the predictive reasoning already covered, in networks quantified with probability, possibility and Dempster-Shafer belief values. The use of these forms of reasoning, which include the phenomenon of "explaining away", is illustrated with the use of a medical example.

Keywords: Qualitative behaviour, evidential reasoning, intercausal reasoning, explaining away, probability theory, possibility theory, Dempster-Shafer theory.

1. Introduction

In the past few years there has been considerable interest in qualitative approaches to reasoning under uncertainty—approaches which do not make use of precise numerical values of the type used by conventional probability theory. These approaches range from systems of argumentation[1,2,3] to systems for nonmonotonic reasoning[4] and abstractions of precise quantitative systems[5,6]. Qualitative abstractions of probabilistic networks, in particular, have proved popular, finding use in areas in which the full numerical formalism is either unnecessary or inappropriate. Applications have been reported in planning[6], explanation[7], diagnosis[8] and engineering design[9].

In qualitative probabilistic networks, the focus is rather different from that of ordinary probabilistic systems. Whereas in probabilistic networks[10] the main goal is to establish the probabilities of hypotheses when particular observations are made, in qualitative systems the main aim is to establish how values change. Thus, given information that a patient has a fever, and given that we are interested in whether the patient has measles, the aim in a qualitative probabilistic system is to establish how the probability of measles changes rather than what the probability of measles is. Since the approach is qualitative, the size of the change is not required. The only consideration is whether the probability increases, in which case the change is positive, written as $[+]$, decreases in which case the change is negative $[-]$, or does not change in which case the change is zero [0]. In some cases it is not possible to resolve the change with any precision so that its value remains unknown, and it is written as [?]. Clearly this information is rather weak, but as the applications show it is sufficient for some tasks. Furthermore, reasoning with qualitative probabilities is much more efficient than reasoning with precise probabilities, since computation is quadratic in the size of the network[5] rather than NP-hard[11].

The popularity of qualitative probabilistic networks prompted work on abstractions of other uncertainty handling formalisms[12,13,14], providing what is essentially a generalisation of the approach provided by qualitative probabilistic networks[15]. This approach uses techniques from qualitative reasoning[16] to determine the behaviour of the formalisms. Given two hypotheses H and G whose probabilities are interesting, the approach relates $p(H)$ to $p(G)$ by establishing the qualitative value of $dp(H)/dp(G)$. Initial results[12] demonstrated how this approach could be used to propagate qualitative probability, possibility[17,18] and Dempster-Shafer belief[19] values in singly connected networks in a predictive direction—the direction in which the conditional values were elicited—and suggested how this propagation might be used to integrate information expressed in the different formalisms. This paper extends the work in a number of ways.

After a brief statement of some basic ideas in Section 2, Section 3 gives results that make use of Bayes' rule and its extensions to other formalisms to enable evidential reasoning. Another new reasoning pattern, intercausal reasoning[20,21], is introduced in Section 4 allowing the propagation of values between the ancestors of a node which represents a variable that is known to be true. To my knowledge this is the first time that this style of reasoning has been explored in possibility and evidence theories. The use of the results is shown on a medical example in Section 5, and the solution of the example necessitates a discussion of a means of integrating the different formalisms.

2. Basic notions

This work is set in the framework of singly connected networks in which the nodes represent variables, and the edges represent explicit dependencies between the variables. When the edges of such graphs are quantified with probability values they are those studied by Pearl[10], when possibility values are used the graphs are those of Fonck and Straszecka[22] and when belief values are used the graphs are those studied by Shafer et al.[23] and Smets[24]. Since we deal with values that may be probabilities, possibilities or beliefs we need a general way of referring to them, and so we define a *certainty value*:

Definition 1 *The certainty value of a variable X taking value x, $val(x)$, is either the probability of x, $p(x)$, the possibility of x, $\Pi(x)$, or the belief in x, $bel(x)$.*

Later on we will also need to distinguish between *upper certainty values*, written $val^*(\cdot)$ which like possibility measure the upper bound on the degree to which a hypothesis might occur, and *lower certainty values*, written $val_*(\cdot)$ like belief which measure the lower bound on the degree to which a hypothesis must occur. This distinction is necessary in order to relate the different formalisms discussed in this paper, since for these the upper certainty value of a hypothesis may be established from the lower certainty value of the complementary hypothesis[25]:

$$val^*(x) = 1 - val_*(\neg x) \tag{1}$$

Thus a belief value may be related to a plausibility value[19], and a possibility value may be related to a necessity value[17]. In the case of probability theory, of course, the upper certainty value and the lower certainty value coincide.

Each node in a graph represents a binary valued variable. We use the convention that the name of the node is a capital letter, often related to the name of the variable it represents, and that the possible values taken by the variable are indicated by lower case letters, usually the lower case letters appropriate to the name of the node. Thus a node X represents some variable, say "Xerxes is alive" whose possible values are x and $\neg x$ with the usual implication that x stands for the value "Xerxes is alive is true" while $\neg x$ stands for the value "Xerxes is alive is false". The set of values $\{x, \neg x\}$ is sometimes written as $\mathbf{X}$. The probability values associated with X are written as $p(x)$ and $p(\neg x)$, and the possibility values associated with X as $\Pi(x)$ and $\Pi(\neg x)$. Belief values may be assigned to any subset of the values of X, so it is possible to have up to three beliefs associated with X—$bel(\{x\})$, $bel(\{\neg x\})$ and $bel(\{x, \neg x\})$. For simplicity these will be written as $bel(x)$, $bel(\neg x)$ and $bel(x \cup \neg x)$, and we will abuse even this rather lapse notation later on by using expressions such as $X \in \{x, x \cup \neg x\}$, $x_i \in X$ to mean that x_i can take the values x or $\neg x$. The use of binary values is purely a matter of convenience since it makes the results simpler to understand and easier to obtain. It is possible to generalise the results[14].

Given two nodes in the network, A and C, which are connected we are interested in the way in which a change in $val(a)$, say, influences $val(c)$ and $val(\neg c)$. We can model the impact of evidence that affects the value of A in terms of the change in certainty value of a and $\neg a$, relative to their value before the evidence was known, and use knowledge about the way that a change in $val(a)$ affects $val(c)$ and $val(\neg c)$ to propagate the effect of the evidence. We define the following relationships that describe how the value of a variable X changes when the value of a variable Y is altered by new evidence:

Definition 2 *The certainty value of a variable X taking value x is said to* follow *the certainty value of variable Y taking value y if $val(x)$ increases when $val(y)$ increases, and $val(x)$ decreases when $val(y)$ decreases.*

Definition 3 *The certainty value of a variable X taking value x is said to* vary inversely *with the certainty value of variable Y taking value y if $val(x)$ decreases when $val(y)$ increases, and $val(x)$ increases when $val(y)$ decreases.*

Definition 4 *The certainty value of a variable X taking value x is said to be* independent *of the certainty value of variable Y taking value y if $val(x)$ does not change as $val(y)$ increases and decreases.*

If it is not possible to determine which of these relationships between $val(x)$ and $val(y)$ hold, then the relationship between $val(x)$ and $val(y)$ is said to be *indeterminate*. These relationships are distinguished because of the way that they relate to the keystone of our method which is the use of qualitative derivatives. The relationship between $val(x)$ and $val(y)$ is charaterised by the derivative $dval(x)/dval(y)$. When the value of the derivative is known, the change in $val(x)$ can be established from the value of the change in $val(y)^{\dagger}$:

$$\Delta val(x) = \frac{dval(x)}{dval(y)} \Delta val(y) \tag{2}$$

Now, we are only interested here in the direction of the change, so we are only interested in the qualitative values[16] of the above terms:

$$[\Delta val(x)] = \left[\frac{dval(x)}{dval(y)}\right] \otimes [\Delta val(y)] \tag{3}$$

where $[x]$ is $[+]$ if x is positive, $[-]$ if x is negative, $[0]$ if x is zero, and $[?]$ if it is not possible to determine whether x is any of the above. The symbol "$\otimes$" denotes qualitative multiplication[6,16]. Clearly $val(x)$ follows $val(y)$ when $[dval(x)/dval(y)] = [+]$, $val(x)$ varies inversely with $val(y)$ when $[dval(x)/dval(y)] = [-]$ and is independent of $val(y)$ when $[dval(x)/dval(y)] = [0]$. If the relationship between $val(x)$ and $val(y)$ is indeterminate then the derivative can take any value and we write $[dval(x)/dval(y)] = [?]$. It is also possible to relate $val(x)$ and $val(y)$ using partial derivatives:

$$[\Delta val(x)] = \left[\frac{\partial val(x)}{\partial val(y)}\right] \otimes [\Delta val(y)] \oplus \left[\frac{\partial val(x)}{\partial val(\neg y)}\right] \otimes [\Delta val(\neg y)] \tag{4}$$

where "$\oplus$" is qualitative addition[6,16]. The difference between the partial derivatives and the (total) derivatives used above is that the latter take account of changes in $val(\neg y)$.

Despite the appeal of the relationship between the terms defined in Definitions 2–4 and the qualitative value of the derivatives, these terms are not sufficient to describe every relationship we come across. We also require the following:

Definition 5 *The certainty value of a variable X taking value x is said to* follow *the certainty value of variable Y taking value y* up *if $val(x)$ increases when $val(y)$ increases, and $val(x)$ does not change when $val(y)$ decreases.*

†Of course, this is only generally true for infinitesimal changes in $val(y)$. However, it turns out[14] that the second derivatives, $d^2 val(x)/dval(y)^2$, of all relations relating certainty values are zero so that the first derivative (which is what we are using) is constant and (2) holds for any change in $val(y)$.

Definition 6 *The certainty value of a variable X taking value x is said to* follow *the certainty value of variable Y taking value y* down *if $val(x)$ does not change when $val(y)$ increases, and $val(x)$ decreases when $val(y)$ decreases.*

Clearly the relationships between $val(x)$ and $val(y)$ described by these two terms are related to those introduced previously since $val(x)$ follows $val(y)$ if and only if it both follows $val(y)$ up and follows it down. We can also introduce the idea that the relationship between $val(x)$ and $val(y)$ is not known exactly, so that it may be the case that $val(x)$ changes when $val(y)$ does but it may also be the case that $val(x)$ does not change when $val(y)$ does. For this we need four new definitions, the first two amend the definitions of "follows" and "vary inversely with":

Definition 7 *The certainty value of a variable X taking value x is said to be related to the certainty value of variable Y taking value y such that $val(x)$* may follow *$val(y)$ if $val(x)$ either follows $val(y)$ or is independent of it.*

Definition 8 *The certainty value of a variable X taking value x is said to be related to the certainty value of variable Y taking value y such that $val(x)$* may vary inversely with *$val(y)$ if $val(x)$ either varies inversely with $val(y)$ or is independent of it.*

The next two defintions amend the definitions of "follow up" and "follow down".

Definition 9 *The certainty value of a variable X taking value x is said to be related to the certainty value of variable Y taking value y such that $val(x)$* may follow *$val(y)$* up *if $val(x)$ either follows $val(y)$ up or is independent of it.*

Definition 10 *The certainty value of a variable X taking value x is said to be related to the certainty value of variable Y taking value y such that $val(x)$* may follow *$val(y)$* down *if $val(x)$ either follows $val(y)$ down or is independent of it.*

We could also introduce further relationships relating to the sub-parts of the "varies inversely" relation and their "maybe" counterparts, but these have not been found necessary to date.

If $val(x)$ follows $val(y)$ up then we write $[dval(x)/dval(y)] = [\uparrow]$, and if $val(x)$ follows $val(y)$ down then we write $[dval(x)/dval(y)] = [\downarrow]$. When we say that $val(x)$ may follow $val(y)$ up then we are saying that $[dval(x)/dval(y)]$ is $[\uparrow]$ or $[0]$ and we write this as $[dval(x)/dval(y)] = [\uparrow, 0]$. Similarly, if $val(x)$ may follow $val(y)$ down then we are saying that $[dval(x)/dval(y)] = [\downarrow, 0]$. Such derivatives naturally lead to changes which are represented by combinations of $[+]$ and $[-]$ with $[0]$ since they are possible increases and decreases.

To handle the new values $[\uparrow]$ and $[\downarrow]$ it is necessary to extend the qualitative multiplication operator to get:

$\otimes$	$[+]$	$[0]$	$[-]$	$[?]$	$[\uparrow]$	$[\downarrow]$
$[+]$	$[+]$	$[0]$	$[-]$	$[?]$	$[+]$	$[0]$
$[0]$	$[0]$	$[0]$	$[0]$	$[0]$	$[0]$	$[0]$
$[-]$	$[-]$	$[0]$	$[+]$	$[?]$	$[0]$	$[-]$
$[?]$	$[?]$	$[0]$	$[?]$	$[?]$	$[+]$	$[-]$

The results are trivial to establish, as they were in the original papers in which they were introduced[6,16]. The product of an increase in value (left column), and a positive derivative (top row) is an increase in value, while the product of an increase in value and a derivative that indicates the relationship in question is "follows down" is no change. The effect of "may follow" and similar derivatives may be established by considering the two effects that they represent. Thus the effect of combining $[0,\uparrow]$ with $[+]$ is $[0] \otimes [+]$ or $[\uparrow] \otimes [+]$. This comes to $[0]$ or $[+]$, which we may write as $[0,+]$. Note that because qualitative addition only ever combines two changes in value, and because changes never have the value $[\uparrow]$ or $[\downarrow]$, we can use the original operator:

$\oplus$	$[+]$	$[0]$	$[-]$	$[?]$
$[+]$	$[+]$	$[+]$	$[?]$	$[?]$
$[0]$	$[+]$	$[0]$	$[-]$	$[?]$
$[-]$	$[?]$	$[-]$	$[-]$	$[?]$
$[?]$	$[?]$	$[?]$	$[?]$	$[?]$

Given this background, it is clear that to determine the change at node C given the change at node A we may use either total or partial derivatives:

$$\begin{bmatrix} [\Delta val(c)] \\ [\Delta val(\neg c)] \end{bmatrix} = \begin{bmatrix} \left[\frac{d\,val(c)}{d\,val(a)}\right] \\ \left[\frac{d\,val(\neg c)}{d\,val(a)}\right] \end{bmatrix} \otimes [\Delta val(a)] \tag{5}$$

$$\begin{bmatrix} [\Delta val(c)] \\ [\Delta val(\neg c)] \end{bmatrix} = \begin{bmatrix} \left[\frac{\partial val(c)}{\partial val(a)}\right] & \left[\frac{\partial val(c)}{\partial val(\neg a)}\right] \\ \left[\frac{\partial val(\neg c)}{\partial val(a)}\right] & \left[\frac{\partial val(\neg c)}{\partial val(\neg a)}\right] \end{bmatrix} \otimes \begin{bmatrix} [\Delta val(a)] \\ [\Delta val(\neg a)] \end{bmatrix} \tag{6}$$

When the value at a node is influenced by changes at several other nodes, we may calculate the overall change by using the principle of superposition[26] which allows us to obtain the compound change by simply summing the changes that would be induced by each influencing node on its own. Thus if there is a node D which is affected by changes at C and another node B, the overall change at D is given by the sum of the change at D due to the change at C and the change at D due to the change at B. This sum may be established by the use of either total or partial derivatives:

$$\begin{bmatrix} [\Delta val(d)] \\ [\Delta val(\neg d)] \end{bmatrix} = \begin{bmatrix} \left[\frac{dval(d)}{dval(b)}\right] \\ \left[\frac{\partial val(\neg d)}{\partial val(b)}\right] \end{bmatrix} \otimes [\Delta val(b)] \oplus \begin{bmatrix} \left[\frac{dval(d)}{dval(c)}\right] \\ \left[\frac{\partial val(\neg d)}{\partial val(c)}\right] \end{bmatrix} \otimes [\Delta val(c)] \quad (7)$$

$$\begin{bmatrix} [\Delta val(d)] \\ [\Delta val(\neg d)] \end{bmatrix} = \begin{bmatrix} \left[\frac{\partial val(d)}{\partial val(b)}\right] & \left[\frac{\partial val(d)}{\partial val(\neg b)}\right] \\ \left[\frac{\partial val(\neg d)}{\partial val(b)}\right] & \left[\frac{\partial val(\neg d)}{\partial val(\neg b)}\right] \end{bmatrix} \otimes \begin{bmatrix} [\Delta val(b)] \\ [\Delta val(\neg b)] \end{bmatrix} \oplus \begin{bmatrix} \left[\frac{\partial val(d)}{\partial val(c)}\right] & \left[\frac{\partial val(d)}{\partial val(\neg c)}\right] \\ \left[\frac{\partial val(\neg d)}{\partial val(c)}\right] & \left[\frac{\partial val(\neg d)}{\partial val(\neg c)}\right] \end{bmatrix} \otimes \begin{bmatrix} [\Delta val(c)] \\ [\Delta val(\neg c)] \end{bmatrix} \quad (8)$$

In this paper we use total derivatives to manipulate probability and belief values and, because of the difficulty of determining how $\Pi(x)$ affects $\Pi(\neg x)$[14], we use partial derivatives to manipulate possibility values. In fact, strictly speaking, it is not possible to establish any kind of derivative in possibility theory since the maximum and minimum combinators used by the theory may not be differentiated. However, it is possible to establish the sign of $\delta val(x)/\delta val(y)$ (the way in which a small change in value rather than an infinitesimal change in value is propagated) and this is what will be manipulated in the case of possibility theory. I trust that the reader will excuse the slight abuse of notation that allows $\partial val(x)/\partial val(y)$ to stand for $\delta\Pi(x)/\delta\Pi(y)$ since it makes the partial nature of the relationship clear.

3. Bayes' rule and its variants

We can use the above approach to analyse the propagation of values in networks quantified using probability, possibility and belief values, obtaining results[12] which allow propagation in the direction implied by the conditional values with which the influences between nodes are quantified. Thus for a direct link between nodes A and C, where the conditional value table that controls propagation over the link is written in terms of values such as $val(c\,|\,a)$, qualitative propagation may be carried out from A to C. Now, such tables are usually elicited in terms of causal influences since if A causes C it is easier to establish $val(c\,|\,a)$ than $val(a\,|\,c)$. However, when performing many types of reasoning, it is often necessary to reason evidentially from effects (which may be observed) to causes (which one wishes to establish). Thus it is desirable to establish a means of obtaining the qualitative behaviour of evidential propagation from the known qualitative behaviour of the causal propagation.

3.1. *Evidential reasoning in probability theory*

In probability theory we can make use of Bayes' rule to establish values such as $p(a\,|\,c)$ from values such as $p(c\,|\,a)$ and thus to establish how qualitative probabilities may be propagated in an evidential direction. If we write the link connecting A and C as $A \rightarrow C$ we have:

Theorem 1 *For $A \rightarrow C$ and for all $x \in \{c, \neg c\}$, $y \in \{a, \neg a\}$, $p(y)$ follows $p(x)$ iff $p(x)$ follows $p(y)$, $p(y)$ varies inversely with $p(x)$ iff $p(x)$ varies inversely with $p(y)$, and $p(y)$ is independent of $p(x)$ iff $p(x)$ is independent of $p(y)$.*

Proof:[27] The qualitative value of $dp(c)/dp(a)$ is determined by $p(c\,|\,a) - p(c\,|\,\neg a)$[12] and so that of $dp(a)/dp(c)$ is determined by $p(a\,|\,c) - p(a\,|\,\neg c)$. Applying Bayes' rule gives us $p(a\,|\,c) - p(a\,|\,\neg c) = p(c\,|\,a)p(a)/p(c) - p(\neg c\,|\,a)p(a)/p(\neg c)$. Writing $p(c)$ as $p(c\,|\,a)p(a) + p(c\,|\,\neg a)p(\neg a)$, $p(\neg c)$ as $p(\neg c\,|\,a)p(a) + p(\neg c\,|\,\neg a)p(\neg a)$ and simplifying we get $p(a \mid c) - p(a \mid \neg c) = \frac{p(a)p(\neg a)}{p(c)p(\neg c)}\Big[p(c \mid a)p(\neg c \mid \neg a) - p(\neg c \mid a)p(c \mid \neg a)\Big]$. Now, because $p(c \mid a) = 1 - p(\neg c \mid a)$, if $p(c \mid a) > p(c \mid \neg a)$ then $p(\neg c \mid \neg a) > p(\neg c \mid a)$, and so the qualitative value of $p(a\,|\,c) - p(a\,|\,\neg c)$ is equal to the qualitative value of $p(c\,|\,a) - p(c\,|\,\neg a)$. From this and similar results for the variation of $p(a)$ with $p(\neg c)$ and $p(\neg a)$ with $p(c)$ and $p(\neg c)$ the result follows.□

Thus a probabilistic influence propagates in the reverse direction exactly as it does in the forward direction, a result which agrees with that of Wellman[6] as well as intuition.

3.2. *Evidential reasoning in possibility theory*

For influences quantified using possibility values, it is possible to perform evidential propagation using a possibilistic version of Bayes' rule provided by Dubois and Prade[28]. This states that the joint possibility distribution over A and C is the same whichever end of the link joining A and C the calculation of the joint value starts from:

$$\min\Big(\Pi(c\,|\,a), \Pi(a)\Big) = \min\Big(\Pi(a\,|\,c), \Pi(c)\Big) \tag{9}$$

$$\min\Big(\Pi(\neg c\,|\,a), \Pi(a)\Big) = \min\Big(\Pi(a\,|\,\neg c), \Pi(\neg c)\Big) \tag{10}$$

This result enables us to obtain:

Theorem 2 *For $A \rightarrow C$ and for all $x \in \{c, \neg c\}$ and $y \in \{a, \neg a\}$, $\Pi(y)$ may follow $\Pi(x)$ up iff $\Pi(x)$ follows $\Pi(y)$ or $\Pi(x)$ may follow $\Pi(y)$ up, and $\Pi(y)$ may follow $\Pi(x)$ down iff $\Pi(x)$ may follow $\Pi(y)$ down or $\Pi(x)$ is independent of $\Pi(y)$.*

Proof: The lengthy proof of this theorem may be found in a previous paper[13].□

Thus reversing a possibilistic link is similar to reversing a probabilistic link, but gives less precise results. When a link that propagates forward so that its consequent follows its antecedent is reversed, it is only possible to say that its antecedent may follow its consequent. Similarly, when a link that blocks propagation, so that the consequent is unaffected by changes in the antecedent, is reversed, the antecedent may follow the consequent. Just as in the probabilistic case the result agrees with intuition—the evidential propagation is roughly the same as the predictive propagation—but is complicated by the use that possibility theory makes of max and min. As a result it is always possible that a change in value at one end of a link will not be propagated because the change in value was not sufficient to pass the threshold at which the value at the other end alters. Similarly, just because a change in the predictive direction is blocked by a link, a change in the evidential direction need not be.

3.3. *Evidential reasoning in evidence theory*

To enable evidential reasoning across influences that are quantified with belief functions, Smets' has provided a generalisation of Bayes' theorem[24,29] which makes it possible to calculate conditional beliefs such as $bel(a \mid c)$ from conditional beliefs such as $bel(c \mid a)$. For $A \rightarrow C$ we have, for any $a \subseteq \mathbf{A}$, $c \subseteq \mathbf{C}$:

$$bel(a \mid c) = \prod_{y \in \tilde{a}} bel(\tilde{c} \mid y) - \prod_{y \in A} bel(\tilde{c} \mid y) \qquad (11)$$

where $\tilde{x}$ is the set-theoretic complement of x with respect to $\mathbf{X}$, namely $\mathbf{X} - x$. Using this rule gives:

Theorem 3 *For $A \rightarrow C$ and for all $x \in \{c, \neg c\}$, $y \in \{a, \neg a\}$, $bel(y)$ may follow $bel(x)$.*

Proof: There are two possible ways of combining belief values when propagating from C to A. If the disjunctive rule of combination[24] is used then $bel(a)$ follows $bel(c)$ whatever the conditionals[12]. If Dempster's rule[19] is used then the qualitative value of $dbel(a)/dbel(c)$ is $[bel(a \mid c) - bel(a \mid c \cup \neg c)]^{12}$. From the generalisation of Bayes' theorem, $bel(a \mid c) = bel(\neg c \mid \neg a) - bel(\neg c \mid a) bel(\neg c \mid \neg a)$ and $bel(a \mid c \cup \neg c) = bel(\emptyset \mid \neg a) - bel(\emptyset \mid a) bel(\emptyset \mid \neg a) = 0$. Since both $bel(\neg c \mid a)$ and $bel(\neg c \mid \neg a)$ are between 0 and 1 by definition it is clear that $bel(a \mid c) \geq bel(a \mid c \cup \neg c)$ and the derivative has value $[+, 0]$ meaning that $bel(a)$ may follow $bel(c)$. From these and similar results for the variation of $bel(a)$ with $bel(\neg c)$ and $bel(\neg a)$ with $bel(\neg c)$ and $bel(\neg a)$ the theorem follows.□

Thus reversing evidential links with Smets' version of Bayes' rule has the effect of making the relationship between the antecedent of the original link and its consequent such that the belief in the antecedent may follow changes in the belief in the consequent. This seems sensible when the forward propagation is carried out using the disjunctive rule since using this rule means that the consequent always follows the antecedent[12] making it sensible that the antecedent may follow consequent on reversal. However, when Dempster's rule is used the behaviour of this belief function version of Bayes' theorem seems less satisfactory. Using Dempster's rule, it

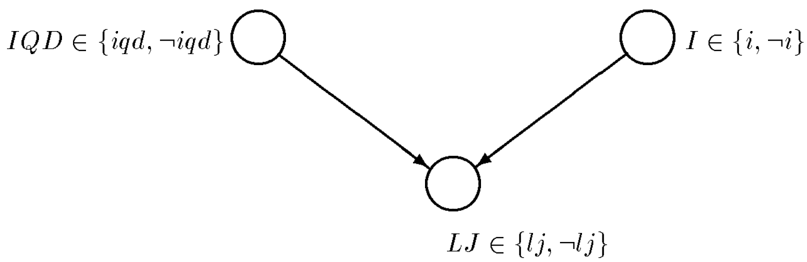

Figure 1: A network describing Jack's beliefs

is possible for $bel(c)$ to vary inversely with $bel(a)$ when reasoning predictively. In such a case it would seem odd that $bel(a)$ may follow $bel(c)$ when the link is reversed. Thus the result suggests that further investigation of Smets' generalised Bayes' theorem would be fruitful, and it is possible that other forms of Bayes' rule that conform to evidence theory and yet have different effects may be proposed.

3.4. *Examples of evidential reasoning*

Late last year, Cody Pomeroy wrote a long letter to his friend Jack Dulouz in which he mentioned that he had just taken a job fixing tyres at the local tyre shack. In the letter he complained about the manager of the tyre shack, a most unscrupulous sort, who immediately sacks any employee if they become ill, or if he finds out that they do not have the relevant certificate of competance in fixing tyres. This made Jack worry about Cody's prospects, since he was well aware both of Cody's fragile health and the fact that Cody had no certificate, and he went to the lengths of drawing up the directed graph in Figure 1 in order to analyse the problem. In the graph node I represents the variable "Ill", IQD represents the variable "Invented qualification discovered", and LJ represents the variable "Lose Job".

Some time later, earlier this year in fact, in a telephone call to Evelyn, Cody's wife, Jack learnt that Cody had indeed lost his job. This made him re-analyse the situation using his direct graph model, to see what it told him about the reasons for Cody's dismissal.

Since Jack has no numerical information about the behaviour of Cody's employer, he is forced to use qualitative methods to carry out his reasoning. As a result he builds his model on the basis that an adequate description of the manager's behaviour is that the certainty value of the proposition "Lose job" must follow the certainty values of the propositions "Ill" and "Invented qualification discovered". This knowledge is sufficient, along with the results obtained above, to permit Jack to update his model with the knowledge that Cody has lost his job. If Jack reasons using probability theory, he can use Theorem 1 to discover that the probabilities of both "Ill" and "Invented qualification discovered" follow that of "Lose job", so that they increase with the new knowledge. Alternatively, if Jack uses belief functions to quantify his model, he can use Theorem 3 to discover that the belief in both

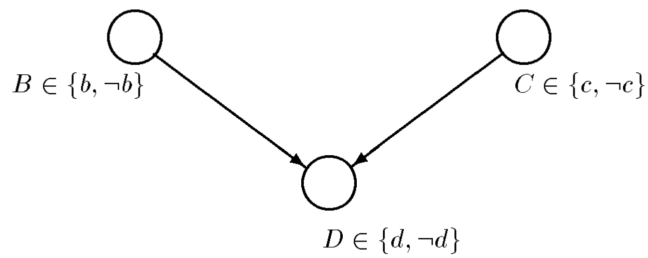

Figure 2: A network for intercausal reasoning

"Ill" and "Invented qualification discovered" may follow that in "Lose job", so that belief in both propositions may increase with the new knowledge—that is it either increases or does not change. Finally, Jack might use possibility theory in his model. In this case, the fact that the possibility of "Lose job" follows the possibilities of "Ill" and "Invented qualification discovered", means that applying Theorem 2 tells Jack that the possibilities of both "Ill" and "Invented qualification discovered" may follow that of "Lose job" up. Thus knowledge of Cody's dismissal leads to the fact that the possibilities of "Ill" and "Invented qualification discovered" may increase.

4. Intercausal reasoning

The results in Section 3 along with those presented previously[12] make it possible to perform both causal and evidential reasoning using qualitative probability, possibility and belief values in singly connected networks. These modes of reasoning are sufficient to deal with many problems, but it is worth considering another important style of reasoning—intercausal reasoning. This is a pattern of reasoning between causes that are dependent on an observed common effect. In this section the approach introduced above is applied to analyse intercausal reasoning in probability, possibility and evidence theories.

The basic network in which intercausal reasoning takes place is that of Figure 2 which we will refer to this network as $B\&C \rightarrow D$. Here B and C are conditionally independent when the value of D is not known, and both are causes of D. When D is observed to take the value d, it is often the case that as evidence is obtained for b the degree of support for c is altered since $val(c)$ and $val(b)$ are no longer independent. Thus, to take the classic[10] example, b could be the hypothesis "The sprinkler was on", c could be the hypothesis "It rained last night" and d could be the hypothesis "The grass is wet". If the grass is known to be wet one morning, then observing something that makes it more likely that the sprinkler was on, say by tripping over the hose on the lawn, makes it less likely that it rained. In this case the relationship between b and c is a negative one, and evidence for b is said to *explain away* c. However, it is also possible that evidence for b might make c more certain or fail to influence it.

4.1. *Intercausal reasoning in probability theory*

To analyse the behaviour of the network we can write down an expression for $p(d)$ in terms of $p(b)$ and $p(c)$ and use this to relate $p(b)$ to $p(c)$ when D has been observed to take the value d, so that $p(d) = 1$. With this approach we obtain:

Theorem 4 *In the network $B\&C \to D$, and for all $x \in \{b, \neg b\}$ and $y \in \{c, \neg c\}$, when $p(d) = 1$:*
(1) $p(x)$ varies inversely with $p(y)$ iff $p(d)$ follows both $p(x)$ and $p(y)$, or $p(d)$ varies inversely with both $p(x)$ and $p(y)$;
(2) $p(x)$ follows $p(y)$ iff $p(d)$ follows $p(x)$ and varies inversely with $p(y)$, or $p(d)$ varies inversely with $p(x)$ and follows $p(y)$;
(3) $p(x)$ is independent of $p(y)$ iff $p(d)$ is independent of $p(y)$ and is not independent of $p(x)$;
(4) Under all other conditions, the relationship between $p(x)$ and $p(y)$ is indeterminate.

Proof: $p(d) = p(d \mid b)p(b) + p(d \mid \neg b)p(\neg b)$. Since $p(d) = 1$, $p(b) = (1 - p(d \mid \neg b)p(\neg b))/(p(d \mid b))$ and since B and C are independent, $p(d \mid b) = p(d \mid b, c)p(c) + p(d \mid \neg b, \neg c)p(\neg c)$. Thus we can write an expression for $p(b)$ in terms of $p(c)$ and differentiate it using the product and quotient rules. This gives $dp(b)/dp(c) = \Big(p(d \mid b)\Big)^{-2} \cdot \Big[p(d \mid b)\frac{d}{dp(c)}\Big(1 - p(d \mid \neg b)p(\neg b)\Big) - \Big(1 - p(d \mid \neg b)p(\neg b)\Big)(dp(d \mid b)/dp(c))\Big]$. Writing K for $p(d \mid b)^{-1}$, we then have $dp(b)/dp(c) = -K^2 \cdot \Big(p(d \mid b)\Big(p(\neg b)(dp(d \mid \neg b)/dp(c)) + p(d \mid \neg b)\Big(dp(\neg b)/dp(c)\Big)\Big) + \Big(1 - p(d \mid \neg b)p(\neg b)\Big)\Big(dp(d \mid b)/dp(c)\Big)\Big)$. Since $dp(b)/dp(c) = -dp(\neg b)/dp(c)$, $\Big(dp(b)/dp(c)\Big) \cdot p(d \mid b) \cdot \Big[p(d \mid b) - p(d \mid \neg b)\Big] = K^2 \cdot \Big(p(d \mid b)p(\neg b)\Big(p(d \mid \neg b, \neg c) - p(d \mid \neg b, c)\Big) + \Big(1 - p(d \mid \neg b)p(\neg b)\Big)\Big(p(d \mid b, \neg c) - p(d \mid b, c)\Big)\Big)$. Now, $1 - p(d \mid \neg b)p(\neg b) = 1 - p(d) + p(d \mid b)p(b)$. Therefore, $\Big(dp(b)/dp(c)\Big) \cdot p(d \mid b) \cdot \Big[p(d \mid b) - p(d \mid \neg b)\Big] = K^2 \cdot \Big(p(d \mid b)p(\neg b)\Big(p(d \mid \neg b, \neg c) - p(d \mid \neg b, c)\Big) + \Big(1 - p(d)p(\neg b)\Big)\Big(p(d \mid b, \neg c) - p(d \mid b, c)\Big) + \Big(p(d \mid b)p(b)\Big)\Big(p(d \mid b, \neg c) - p(d \mid b, c)\Big)\Big)$. Noting that $p(d) = 1$, and combining the two remaining terms on the right hand side, we get $\Big(dp(b)/dp(c)\Big) \cdot p(d \mid b) \cdot \Big[p(d \mid b) - p(d \mid \neg b)\Big] = K^2 \cdot \Big(p(d \mid b)\Big(p(d \mid \neg b, \neg c)p(\neg b) - p(d \mid \neg b, c)p(\neg b) + p(d \mid b, \neg c)p(b) - p(d \mid b, c)p(b)\Big)\Big)$, which gives us $\Big(dp(b)/dp(c)\Big) \cdot \Big[p(d \mid b) - p(d \mid \neg b)\Big] = K^2 \cdot \Big[p(d \mid \neg c) - p(d \mid c)\Big]$. Thus the qualitative value of $dp(b)/dp(c)$ depends upon the value of the two expressions $p(d \mid b) - p(d \mid \neg b)$ and $p(d \mid c) - p(d \mid \neg c)$, being negative when both have the same qualitative value, being positive when they have opposite qualitative values, and being zero when $p(d \mid c) - p(d \mid \neg c)$ is zero and $p(d \mid b) - p(d \mid \neg b)$ is not. But these expressions are exactly those which determine how $p(d)$ varies with $p(b)$ and $p(c)$[12], so the theorem is true for $x \equiv b$ and $y \equiv c$. Similar results for $p(b)$ varying with $p(c)$ and $p(\neg b)$

varying with $p(c)$ and $p(\neg c)$ complete the proof. $\square$

Thus the qualitative relation between the probabilities of B and C is determined by the qualitative relation between the probabilities of B and D and the qualitative relation between the probabilities of C and D. Clearly the case in which explaining away occurs is that in which $p(b)$ varies inversely with $p(c)$, and it is worth noting that the relationship between B and C is symmetrical so that if, for instance, $p(b)$ is explained away by $p(c)$, then $p(c)$ is explained away by $p(b)$. The conditions seem entirely reasonable, and may be justified by the following argument. If the probability of d tends to follow that of $p(c)$, and $p(c)$ increases, then the joint probability of b, c and d increases as $p(c)$ increases and when $p(d)$ is not fixed, this will cause it to increase. If, however, the probability of d is fixed, there must be some change in $p(b)$ to offset the change in the joint value, and if $p(d)$ follows $p(b)$ this means that $p(b)$ must decrease. Similarly, if $p(d)$ follows $p(c)$ and varies inversely with $p(b)$, $p(b)$ must increase as $p(c)$ increases in order to offset the change that would otherwise occur in the joint probability.

Whilst this explanation seems an adequate justification of the kind of intercausal relationship implied by Theorem 4, the fact that the conditions are on the relationship between the probability of D and its causes, rather than simply on the product of the conditional values of D given its causes, makes it clear that the notion of "explaining away" that is captured here is rather different to that of other authors such as Druzdzel, Henrion and Wellman[20,21].

As an aside, it should be pointed out that our explanation of the way in which the probabilities alter is similar to that used in the argument put forward by Tzeng[30] in his re-establishment of Henrion and Wellman's[21] result. He examines Pearl's[10] method for probability propagation, and considers the flow of probability between the nodes. His result, and its proof are reconstructed below using our approach to qualitative uncertainty:

Theorem 5 *In the network $B\&C \rightarrow D$, and for all $x \in \{b, \neg b\}$ and $y \in \{c, \neg c\}$, $p(x)$ varies inversely with $p(y)$ if $p(d \,|\, x, y)p(d \,|\, \neg x \neg y) < p(d \,|\, \neg x, y)p(d \,|\, x, \neg y)$ The conditions under which $p(x)$ follows $p(y)$ and is independent of $p(y)$ may be obtained analagously.*

Proof: When D is known to be true, the evidential flow of probability into node B is such that the ratio of the change in $p(b)$ to that in $p(\neg b)$ is given by the ratio of $p(d \,|\, b)$ to $p(d \,|\, \neg b)$. Now, since B and C are independent, $p(d) = \sum_{B \in \{b, \neg b\}, C \in \{c, \neg c\}} p(d \,|\, B, C)p(B)p(C)$ and $p(d \,|\, b) = p(d \,|\, b, c)p(c) + p(d \,|\, b, \neg c)p(\neg c) = p(c)\Big[p(d \,|\, b, c) - p(d \,|\, b, \neg c)\Big] + p(d \,|\, b, \neg c)$ and a similar expression may be written for $p(d \,|\, \neg b)$. Now, we are interested in the way in which this ratio alters as $p(c)$ changes. If the ratio increases, $p(b)$ increases as $p(c)$ increases, and if the ratio decreases, $p(c)$ explains $p(b)$ away. $\frac{d}{dp(c)}\Big(p(d \,|\, b)/p(d \,|\, \neg b)\Big) = \frac{d}{dp(c)}\Big(p(c)\Big[p(d \,|\, b, c) - p(d \,|\, b, \neg c)\Big] + p(d \,|\, b, \neg c)\Big) / \Big(p(c)\Big[p(d \,|$

$\neg b, c)-p(d \mid \neg b, \neg c)\Big]+p(d \mid \neg b, \neg c)\Big) = K\Big[p(d \mid b, c)p(d \mid \neg b, \neg c)-p(d \mid \neg b, c)p(d \mid b, \neg c)\Big]$ where $K = p(d \mid b)^{-2}$. This latter can be ignored since it is always positive, so the sign of the derivative, and hence the behaviour of $p(b)$ depends upon $p(d \mid b, c)p(d \mid \neg b, \neg c) - p(d \mid \neg b, c)p(d \mid b, \neg c)$ alone. From similar results for the variation of $p(b)$ with $p(\neg c)$ and $p(\neg b)$ with $p(c)$ and $p(\neg c)$ the result follows.□

4.2. *Intercausal reasoning in possibility theory*

When the network is quantified with possibility values, the observation that D takes value d is modelled by setting the value of $\Pi(\neg d)$ to 0. Since the values are normalised, $\max(\Pi(d), \Pi(\neg d)) = 1$ and so the effect of the observation on the possibility of d is to make $\Pi(d) = 1$. This yields the following result:

Theorem 6 *In the network $B\&C \rightarrow D$, and for all $x \in \{b, \neg b\}$ and $y \in \{c, \neg c\}$, when $\Pi(d) = 1$, $\Pi(x)$ varies inversely with $\Pi(y)$ when $\Pi(d)$ follows $\Pi(y)$, $\Pi(d \mid \neg b, \neg c) < \Pi(d \mid b, \neg c) = 1$, and initially $\Pi(\neg b) = 1$, $\Pi(c) = 1$, and $\Pi(b) < 1$. Otherwise $\Pi(x)$ is independent of $\Pi(y)$.*

Proof: The relationship between $\Pi(d)$, $\Pi(b)$ and $\Pi(c)$ may be determined from $\Pi(d) = \sup_{B \in \{b, \neg b\}, C \in \{c, \neg c\}} \Pi(d, B, C)$ where $\Pi(d, B, C) = \min\big(\Pi(d \mid B, C)\Pi(B)\ \Pi(C)\big)$. The only time that a change in $\Pi(c)$ can require a change in $\Pi(b)$ is when $\Pi(c)$ changes from a value that determines $\Pi(d)$, meaning that $\Pi(d)$ must follow $\Pi(c)$ so that initially $\Pi(c) = 1$ and either (1) $\Pi(b) = \Pi(d \mid b, c) = 1$ or (2) $\Pi(\neg b) = \Pi(d \mid \neg b, c) = 1$, or both. If (1) is the case, then $\Pi(b)$ cannot be forced to change and so we have the requirement that $\Pi(b) < 1$ for there to be any intercausal reasoning. In case (2) we also require $\Pi(d \mid \neg b, \neg c) < 1$ to ensure that changes in $\Pi(c)$ have any effect. Then, if $\Pi(c)$ falls, provided that $\Pi(d \mid b, \neg c) = 1$ the fact that $\Pi(d)$ is held to 1 will mean that $\Pi(b)$ increases to 1. From similar results for the variation of $\Pi(b)$ with $\Pi(\neg c)$ and $\Pi(\neg b)$ with $\Pi(c)$ and $\Pi(\neg c)$ the result may be obtained.□

Thus we can have a form of explaining away in possibility theory, although it is a rather limited one. The inverse relationship between $\Pi(b)$ and $\Pi(c)$ can only be expressed in such a way that $\Pi(b)$ increases as $\Pi(c)$ falls. Thus it is the case that evidence for C not taking value c *explains* B taking value b rather than evidence for C taking value c explaining away B taking value b. In addition there cannot be a positive relationship between B and C so that $\Pi(b)$ can never follow $\Pi(c)$. For this form of intercausal reasoning between B and C to occur in possibility theory, the conditional possibilities must be such that it is less possible for B and C to respectively take values $\neg b$ and $\neg c$ than to take values b and $\neg c$ suggesting some kind of exclusivity between the values. It is also necessary that $\Pi(d)$ would be affected by a change in value of $\Pi(c)$ were it not fixed, and as discussed above this seems an entirely reasonable restriction.

It is worth noting that, unlike the case in probability theory, prediction of explaining away in possibility theory requires explicit knowledge of the quantitiative possibility values as well as some information about qualitative relations.

4.3. *Intercausal reasoning in evidence theory*

When modelling the situation depicted in Figure 2 using evidence theory, there are a number of possible ways of combining the influences of B and C on D[12,24]. One may either use conditional values of the form $bel(d \,|\, b)$ or conditional values of the form $bel(d \,|\, b, c)$ and one may combine conditionals using either Dempster's rule of combination[19] which involves the use of conditionals such as $bel(d \mid b, c \cup \neg c)$, or Smets' disjunctive rule[24] which does not need such values, replacing $bel(d \,|\, b, c \cup \neg c)$ with $bel(d \,|\, b, c) bel(d \,|\, b, \neg c)$. Firstly, for Dempster's rule we have:

Theorem 7 *In the network $B\&C \rightarrow D$, and for all $x \in \{b, \neg b\}$ and $y \in \{c, \neg c\}$, when $bel(d) = 1$:*
(1) $bel(x)$ varies inversely with $bel(y)$ iff $bel(d)$ follows both $bel(x)$ and $bel(y)$, or $bel(d)$ varies inversely with both $bel(x)$ and $bel(y)$;
(2) $bel(x)$ follows $bel(y)$ iff $bel(d)$ follows $bel(x)$ and varies inversely with $bel(y)$, or $bel(d)$ varies inversely with $bel(x)$ and follows $bel(y)$;
(3) $bel(x)$ is independent of $bel(y)$ iff $bel(d)$ is independent of $bel(y)$ and is not independent of $bel(x)$;
(4) Under all other conditions, the relationship between $bel(x)$ and $bel(y)$ is indeterminate.

Proof: There are two cases. In the first, we have conditionals such as $bel(d \,|\, b, c)$. Since B and C are conditionally independent, $bel(d) = \sum_{B \subseteq \{b, \neg b\}} bel(d \,|\, B) m(B)$, and $bel(d \,|\, B) = \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, B, C) m(C)$. Given that $bel(d)$ is known to be 1, we have $1 = \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, b, C) m(C) m(b) + \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, \neg b, C) m(C) m(\neg b)$ $+ \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B, C) m(C) m(B)$. Since $bel(b) = m(b)$ this gives us:

$$bel(b) = \frac{1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B, C) m(C) m(B)}{\sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, b, C) m(C)}$$

Taking the derivative of this with respect to $bel(c)$, and writing $\left(\sum_{C \subseteq \{c, \neg c\}} bel(d \mid b, C) m(C) \right)^{-1} = bel(d \mid b)^{-1}$ as K we have $dbel(b)/dbel(c) = K^2 \cdot \left(\left(\sum_{C \subseteq \{c, \neg c\}} bel(d \mid b, C) m(C) \right) \frac{d}{dbel(c)} \left(1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \mid B, C) m(C) m(B) \right) - \left(1 - \sum_{C \subseteq \{c, \neg c\},\ B \in \{\neg b, b \cup \neg b\}} bel(d \mid B, C) m(C) m(B) \right) \frac{d}{dbel(c)} \left(\sum_{C \subseteq \{c, \neg c\}} bel(d \mid b, C) m(C) \right) \right)$. Now, $\frac{d}{dbel(c)} \left(\sum_{C \subseteq \{c, \neg c\}} bel(d \mid b, C) m(C) \right) = \left[bel(d \mid b, c) - bel(d \mid b, c \cup \neg c) \right]$, while $\frac{d}{dbel(c)} \left(1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \mid B, C) m(C) m(B) \right) = - \sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \mid B, c) m(B) + \sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \mid B, c \cup \neg c) m(B) + \left(dbel(b)/dbel(c) \right) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, b \cup \neg b, C) m(C)$. Thus $\left(dbel(b)/dbel(c) \right) \left[bel(d \,|\, b) - bel(d \,|\, b \cup \neg b) \right] = K^2 \cdot \Big(- bel(d \,|\, b) \Big(\sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B, c) m(B) + \sum_{B \in \{\neg b, b \cup \neg b\}}$

$bel(d \,|\, B, c \cup \neg c)m(B)\Big) - \Big(1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B, C)m(C)m(B)\Big)\big[bel(d \,|\, b, c) - bel(d \,|\, b, c \cup \neg c)\big]\Big)$. Now, it is clearly possible to write $1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B, C)m(C)m(B)$ as $1 - bel(d) + bel(d \,|\, b)m(b)$, which, since $bel(d) = 1$ becomes simply $bel(d \mid b)m(b)$. This allows us to write $\Big(dbel(b)/dbel(c)\Big)\big[bel(d \mid b) - bel(d \mid b \cup \neg b)\big] = K^2 \cdot \Big(-bel(d \,|\, b)\Big(bel(d \,|\, c) - bel(d \,|\, b, c)bel(d \,|\, b)m(b) + bel(d \,|\, c \cup \neg c) - bel(d \,|\, b, c \cup \neg c)bel(d \,|\, b)m(b)\Big) - bel(d \,|\, b)m(b)[bel(d \,|\, b, c) - bel(d \,|\, b, c \cup \neg c)]\Big)$, which neatly reduces to $\Big(dbel(b)/dbel(c)\Big)\big[bel(d \,|\, b) - bel(d \,|\, b \cup \neg b)\big] = -\big[bel(d \,|\, c) - bel(d \,|\, c \cup \neg c)\big]$. Thus the value of the derivative is controlled by the qualitative value of two expressions: (1) $bel(d \mid b) - bel(d \mid b \cup \neg b)$ and (2) $bel(d \mid c) - bel(d \mid c \cup \neg c)$. The derivative is positive when one expression is positive and one negative, negative when both expressions are positive, or both are negative, and zero when (2) is zero and (1) is not. However, (1) and (2) are exactly the expressions that determine the relationship between $bel(d)$ and $bel(b)$ and $bel(c)$[12], such that $bel(d)$ follows $bel(b)$ when (1) is positive, varies inversely with $bel(b)$ when (1) is negative, follows $bel(c)$ when (2) is positive and varies inversely with $bel(c)$ when (2) is negative[12]. Thus the result follows for $x \equiv b$ and $y \equiv c$ and since similar results may be obtained for the variation of $bel(b)$ with $bel(\neg c)$ and $bel(\neg b)$ with $bel(c)$ and $bel(\neg c)$, we have proved the result for the first case.

In the second case, we have conditionals such as $bel(d \,|\, b)$, and $bel(d) = \sum_{B \subseteq \{b, \neg b\}, C \subseteq \{c, \neg c\}} bel(d \,|\, B)bel(d \,|\, C)m(B)m(C)$. Thus:

$$bel(b) = \frac{1 - \sum_{B \in \{b, \neg b\}, C \subseteq \{c, \neg c\}} bel(d \,|\, B)bel(d \,|\, C)m(B)m(C)}{bel(d \,|\, b) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, C)m(C)}$$

Taking the derivative of this with respect to $bel(c)$, and writing $\Big(bel(d \,|\, b) \sum_{C \subseteq \{c, \neg c\}} bel(d \mid C)m(C)\Big)^{-1}$ as K we have $dbel(b)/dbel(c) = K^2 \cdot \Big(\Big(bel(d \mid b) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, C)m(C)\Big)\frac{d}{dbel(c)}\Big(1 - \sum_{B \in \{b, \neg b\}, C \subseteq \{c, \neg c\}} bel(d \,|\, B)bel(d \,|\, C)m(B)m(C)\Big) - \Big(1 - \sum_{B \in \{b, \neg b\}, C \subseteq \{c, \neg c\}} bel(d \mid B)bel(d \mid C)m(B)m(C)\Big)\frac{d}{dbel(c)}\Big(bel(d \,|\, b) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, C)m(C)\Big)\Big)$. Now, $\frac{d}{dbel(c)}\Big(bel(d \,|\, b) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, C)m(C)\Big) = bel(d \,|\, b)\big[bel(d \,|\, c) - bel(d \,|\, c \cup \neg c)\big]$, while $\frac{d}{dbel(c)}\Big(1 - \sum_{B \in \{b, \neg b\}, C \subseteq \{c, \neg c\}} bel(d \,|\, B)bel(d \,|\, C)m(B)m(C)) = -\sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B)bel(d \,|\, c)m(B) + \sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \,|\, B)bel(c \cup \neg c)m(B) + \Big(dbel(b)/dbel(c)\Big) \sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, b \cup \neg b)bel(C)m(C)$. Thus, since $\sum_{C \subseteq \{c, \neg c\}} bel(d \,|\, c)m(c) = 1$, which it must be to allow $bel(d) = 1$, $\Big(dbel(b)/dbel(c)\Big)\big[bel(d \mid b) - bel(d \mid b \cup \neg b)\big] = K^2 \cdot \Big(-bel(d \mid b)\Big(\sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \mid B)bel(d \mid c)m(B) + \sum_{B \in \{\neg b, b \cup \neg b\}} bel(d \mid B)bel(d \mid c \cup \neg c)m(B)\Big) - \Big(1 - \sum_{C \subseteq \{c, \neg c\}, B \in \{\neg b, b \cup \neg b\}} bel(d \mid B)bel(d \,|\, C)m(C)m(B)\Big)\big[bel(d \,|\, b)bel(d \,|\, c) - bel(d \,|\, b)bel(d \,|\, c \cup \neg c)\big]\Big)$. As above, we can write $1 - \sum_{C\subseteq\{c,\neg c\},B\in\{\neg b,b\cup\neg b\}} bel(d \mid B)bel(d \mid C)m(C)m(B)$ as $1 - bel(d) + bel(d \mid b)m(b)$, which, since $bel(d) = 1$ becomes simply $bel(d \mid b)m(b)$. This allows us to write $\Big(dbel(b)/dbel(c)\Big)\big[bel(d \mid b) - bel(d \mid b \cup \neg b)\big] = K^2 \cdot \Big(- bel(d \mid b)\big(bel(d \mid c) - bel(d \mid c)bel(d \mid b)m(b) + bel(d \mid c \cup \neg c) - bel(d \mid c \cup \neg c)bel(d \mid b, c \cup \neg c)m(b)\big) - bel(d \,|\, b)^2 m(b)[bel(d \,|\, c) - bel(d \,|\, c \cup \neg c)]\Big)$, which neatly reduces to $\Big(dbel(b)/dbel(c)\Big)\big[bel(d \,|\, b) - bel(d \,|\, b \cup \neg b)\big] = -\big[\big(bel(d \,|\, c) - bel(d \,|\, c \cup \neg c)\big]$, exactly as for the first case. Thus in both cases the relationship between $bel(b)$ and $bel(c)$ is subject to the same conditions, and the theorem is proved.□

Thus for explaining away to take place in evidence theory, the conditions that must be met are analagous to those for probability theory, and suggest that the same kind of mechanism is at work. If $bel(d)$ follows $bel(c)$ and $bel(b)$ when it is not fixed, then when it is fixed the inflow of belief into the joint distribution over D, B and C from increasing $bel(c)$ must be matched by a decrease in $bel(b)$. Similarly, if $bel(d)$ follows $bel(c)$ and varies inversely with $bel(b)$, then when $bel(d)$ is fixed, the increased belief over all three variables in question that results from an increase in $bel(c)$ must be offset by a decrease in $bel(b)$.

As mentioned above, it is also possible to combine the effects of B and C on D using Smets' disjunctive rule[24]. The idea behind this rule is that it should establish the belief in the disjunction of two events for which the belief in their occurence is known in the same way that Dempster's rule[19] establishes the belief in the conjunction of the events. When the disjunctive rule is used we indicate its adoption by referring to the network of Figure 2 as $B \vee C \to D$, and find:

Theorem 8 *For the network $B \vee C \to D$, for all $x \in \{b, \neg b\}$ and $y \in \{c, \neg c\}$, when $bel(d) = 1$, $bel(x)$ may varies inversely with $bel(y)$.*

Proof: Again we have two cases. In the first we have conditionals such as $bel(d \,|\, b, c)$ and $bel(d) = \sum_{B\subseteq\{b,\neg b\},C\subseteq\{c,\neg c\}} m(B)m(C) \prod_{b_i\in B,c_j\in C} bel(d \,|\, b_i, c_j)$. Thus:

$$bel(b) = \frac{1 - \sum_{B\in\{b,\neg b\},C\subseteq\{c,\neg c\}} m(B)m(C) \prod_{b_i\in B,c_j\in C} bel(d \,|\, b_i, c_j)}{\sum_{C\subseteq\{c,\neg c\}} m(C) \prod_{c_j\in C} bel(d \,|\, b, c_j)}$$

which with $K = \left[\sum_{C\subseteq\{c,\neg c\}} m(C) \prod_{c_j\in C} bel(d \,|\, b, c_j)\right]^{-1}$ gives us the derivative $dbel(b)/dbel(c) = K^2 \cdot \Big(\sum_{C\subseteq\{c,\neg c\}} m(C) \prod_{c_j\in C} bel(d \mid b, c_j) \frac{d}{dbel(c)} \Big(1 - \sum_{B\in\{b,\neg b\}, C\subseteq\{c,\neg c\}} m(B)m(C) \prod_{b_i\in B,c_j\in C} bel(d \,|\, b_i, c_j)\Big) - \Big(1 - \sum_{B\in\{b,\neg b\},C\subseteq\{c,\neg c\}} m(B)m(C) \prod_{b_i\in B,c_j\in C} bel(d \mid b_i, c_j)\Big) \frac{d}{dbel(c)} \Big(\sum_{C\subseteq\{c,\neg c\}} m(C) \prod_{c_j\in C} bel(d \mid b, c_j)\Big)\Big)$. Now, we have $\frac{d}{dbel(c)} \Big(\sum_{C\subseteq\{c,\neg c\}} m(C) \prod_{c_j\in C} bel(d \,|\, b, c_j)\Big) = m(b)bel(d \,|\, b, c) - m(b) \prod_{c_j\in C} bel(d \,|\, b, c_j)$. Since the product will contain the term $bel(d \,|\, b, c)$, and all belief values are not greater than 1, this derivative can never be negative. We also have $\frac{d}{dbel(c)}\Big(1-$

$\sum_{B\in\{b,\neg b\},C\subseteq\{c,\neg c\}} m(B)m(C)\prod_{b_i\in B,c_j\in C} bel(d \mid b_i,c_j)\Big) = -\sum_{B\in\{b,b\cup\neg b\}} m(B)$ $\prod_{b_i\in B} bel(d\,|\,b_i,c)+\sum_{B\in\{b,b\cup\neg b\}} m(B)\prod_{b_i\in B,c_j\in\{c,\neg c\}} bel(d\,|\,b_i,c_j)-dbel(b)/dbel(c)$ $\Big(\sum_{C\subseteq\{c,\neg c\}}\prod_{b_i\in\{b,\neg b\},c_j\in C} bel(d\,|\,b_i,c_j)\Big)$. Once again, the second term is a product of terms including the first, and in this case the limitations on the possible values of beliefs mean that the sum of the first two terms cannot be positive. As a result, we can say that $\Big(dbel(b)/dbel(c)\Big)\Big(K-\sum_{c\subseteq\{c,\neg c\}}\prod_{b_i\in\{b,\neg b\},c_j\in C} bel(d\,|\,b_i,c_j)\Big)\leq 0$. Having done this we can recall the value represented by K and observe yet again that we have the difference of two terms where the second cannot be larger than the first, so that $dbel(b)/dbel(c)$ is either negative or zero, and this, along with the obvious symmetric results for the variation of $bel(b)$ with $bel(\neg c)$ and $bel(\neg b)$ with $bel(c)$ and $bel(\neg c)$ gives us the necessary result for this first case.

In the second case we have conditionals such as $bel(d \mid b)$ and $bel(d \mid b,c)$ and $bel(d)=\sum_{B\subseteq\{b,\neg b\},C\subseteq\{c,\neg c\}} m(B)m(C)\prod_{b_i\in B,c_j\in C} bel(d\,|\,b_i)bel(d\,|\,c_j)$. This is the same as the expression for $bel(b)$ obtained above, but with $bel(d\,|\,b_i,c_j)$ replaced by the product of two conditional beliefs. This substitution will not change the qualitative value of the derivative which is thus never positive. From similar results for the variation of $bel(b)$ with $bel(\neg c)$ and $bel(\neg b)$ with $bel(c)$ and $bel(\neg c)$ the result may be obtained. □

Thus if there is any intercausal reasoning when the disjunctive rule is used, it is in the form of explaining away. Given the behaviour reported in Theorem 7 for combination using Dempster's rule, and the fact that under the disjunctive rule $bel(d)$ will always follow $bel(b)$ and $bel(c)$[12], this result is not surprising. It does, however, have some consequences for the expressiveness of the networks that one may build using belief functions and the disjunctive rule. Indeed, the practical result of Theorem 8 is that it is not possible to construct a network of the form $B\vee C\rightarrow D$ where B, C and D are binary valued, in which evidence for the values of C causes belief in the values of B to increase. This is something of a restriction, and may have important consequences for Xu and Smets' evidential networks[31] which use the disjunctive rule in a similar way to that analysed here.

4.4. *Examples of intercausal reasoning*

Following his initial thoughts about Cody's dismissal, Jack telephones him to discuss the matter. During the conversation, in which Cody talks of his desire to become a brakeman on the railroad, Jack learns that, although he is not sure of the matter, Cody reckons that the manager did not find out about his lack of qualification. Armed with this new information, Jack sits down to reason once again about the situation using the same model as before (for convenience repeated in Figure 3).

Given his initial probabilistic model, in which "Lose job" follows both "Ill" and "Invented qualification discovered", Jack can apply Theorem 4 to determine that, given he knows "Lose job" is true, evidence against "Invented qualification

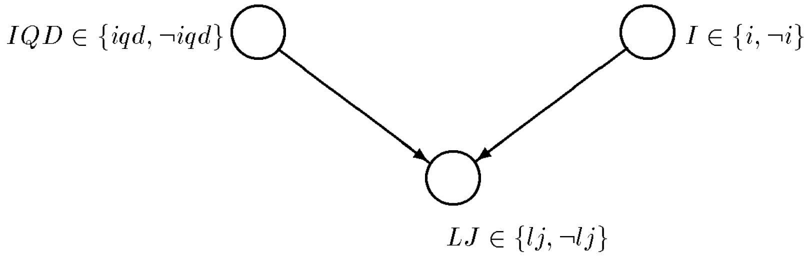

Figure 3: Jack's network revisited

discovered" is evidence for "Ill" since the probability of the latter varies inversely with that of the former. Thus, since on Cody's evidence the probability of "Invented qualification discovered" may fall, the probability that Cody was ill may increase. He get similar results with his evidence theory model. The fact that in the model his belief in "Lose job" follows his belief in both "Ill" and "Invented qualification discovered" means that whether he combines his beliefs with Dempster's rule or Smets' disjunctive rule, Theorems 7 and 8 tell him that his belief in "Ill" may vary inversely with his belief in "Invented qualification dicovered" giving the same result as in the probabilistic case.

Things are a little different if Jack chooses to use a possibilistic model. In this case, he cannot get away without using some numerical values since these values themselves are needed to use Theorem 6. After some thought, he settles on the possibility values in Table 1 which fit with his feelings about Cody's employer as well as the health and educational status of the tyre shack employees while ensuring that the possibility of "Lose job", $\Pi(lj)$, follows both those of "Ill", $\Pi(i)$, and "Invented qualificiation discovered", $\Pi(iqd)$, in accordance with his initial information. Now,

Table 1: Possibilities for the example

$\Pi(lj \,\vert\, i, iqd)$	$=$	1	$p(i)$	$=$	0.1
$\Pi(lj \,\vert\, \neg i, iqd)$	$=$	1	$p(iqd)$	$=$	0.1
$\Pi(lj \,\vert\, i, \neg iqd)$	$=$	1			
$\Pi(lj \,\vert\, \neg i, \neg iqd)$	$=$	0.05			

the conditional values obey the conditions imposed by Theorem 6 for some form of intercausal relationship to hold between $\Pi(i)$ and $\Pi(iqd)$, but the prior possibilities of the events themselves rule out any such relation. Thus, using his possibilistic model, Jack is forced to conclude that the change in the possibility of "Invented qualification discovered" has no effect upon the change in possibility of "Ill".

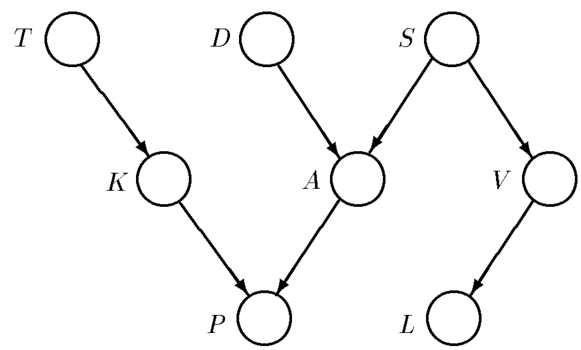

Figure 4: A network representing medical knowledge

5. Integrating formalisms in evidential and intercausal reasoning

One of the applications of predictive reasoning using qualitative uncertainty is the integration of different formalisms[12,14,32]. In this section we demonstrate the use of the results obtained in previous sections in the integration of different formalisms in evidential and intercausal reasoning. To do this we use the following medical example. The network of Figure 4 encodes the medical information that joint trauma (T) leads to loose knee bodies (K), and that these and arthritis (A) cause pain (P). The incidence of arthritis is influenced by dislocation (D) of the joint in question and by the patient suffering from Sjorgen's syndrome (S) . Sjorgen's syndrome affects the incidence of vasculitis (V), and vasculitis leads to vasculitic lesions (L). Consider further a scenario[12,14,32] in which the influences between the nodes are quantified using a mixture of probability, possibility and belief values for exactly the same reasons that a mixture of formalisms are used in MILORD[33,34]—the only quantitative information that is available is expressed in different formalisms. Thus the relationship between T and K, S and V and D, S and A is expressed using probability, that between V and L using possibility theory, and that between K, A and P using evidence theory. Now, we are told that, by applying previous results for the propagation of qualitative uncertainty in a predictive direction[12] one can tell that $p(k)$ follows $p(t)$, $p(a)$ follows $p(s)$, $bel(p)$ follows $bel(a)$ and $bel(k)$, $p(v)$ varies inversely with $p(s)$ and $\Pi(l)$ may follow $\Pi(v)$ down. Given that a particular patient is in pain, how will an observation that suggests that the patient does not have vasculitic lesions affect the probability that they are suffering from joint trauma?

To answer this question we must propagate the effect of the change in the value of $\Pi(l)$ to find the effect on $p(k)$ and to do this we need to combine changes in values expressed in different formalisms. Previously we have suggested that this may be achieved by means of the so-called *monotonicity assumption*, a heuristic which states that:

> If the value of a hypothesis in one formalism increases, the value of the same hypothesis in any other formalism does not decrease.

However, Prade[35] has pointed out that as it stands this assumption is flawed, and to overcome the flaw, we have to restate it in terms that capture the difference between upper and lower certainty values[36]. In particular we need to relate the change in an upper certainty value of a hypothesis to the lower certainty value of the complement of that hypothesis and vice-versa. That is $\Delta val^*(x)$ must be related to $\Delta val_*(\neg x)$ and $\Delta val_*(x)$ must be related to $\Delta val^*(\neg x)$. Taking this need into account, a more correct version of the monotonicity assumption may be obtained:

> If the lower certainty value of a hypothesis in one formalism increases, the lower certainty value of the same hypothesis in any other formalism does not decrease, and the upper certainty value of the complement of the hypothesis in any formalism does not increase.
>
> Similarly, if the upper certainty value of a hypothesis in one formalism increases, the upper certainty value of the same hypothesis in any other formalism does not decrease, and the lower certainty value of the complement of the hypothesis in any formalism does not increase.

A similar pair of statements may be made for decreases in value. What this assumption means for the probabilty, possibility and belief measures that we are dealing with is that given some hypothesis h, if $p(h)$ is known to increase then $bel(h)$ will not decrease, and $\Pi(\neg h)$ will not increase. Similarly, if $\Pi(h)$ increases, then $p(\neg h)$ will not decrease and neither will $bel(\neg h)$. Again, this assumption may fail to hold in some cases—it is, after all, a heuristic. However, it does allow us to make useful deductions in those situations in which we are happy to employ it. The use of the assumption clearly raises questions of semantic coherence and necessitates the adoption of a suitable semantic model such as that of "degrading"[14] which is based upon the idea that, at heart, all numerical methods for handling uncertainty are trying to measure the extent to which it is reasonable to predict that a variable will take a given value.

Now, in our example, the observation suggests the patient does not have vasculitic lesions, so $\Delta\Pi(l) = [-]$. From Theorem 2, $\Pi(v)$ may follow $\Pi(l)$ down, so $\Delta\Pi(v) = [0, -]$. In other words, $\Pi(v)$ may decrease. Applying the new monotonicity assumption gives $\Delta p(\neg v) = [0, +]$ which since $p(v) + p(\neg v) = 1$ means that $\Delta p(v) = [0, -]$ ‡. Now, Theorem 1 tells us $p(s)$ varies inversely with $p(v)$, and as a result $\Delta p(s) = [+, 0]$. Since $p(s)$ may increase, $p(a)$ may increase, and so the monotonicity assumption gives $bel(a) = [+, 0]$. Now, because $bel(p)$ would follow $bel(a)$ and $bel(k)$ if it were not fixed by the knowledge that the patient is in pain, Theorems 7 and 8 tell us that irrespective of whether Dempster's rule or the disjunctive rule is used, $\Delta bel(k) = [0, -]$, and using the monotonicity assumption this means $\Delta p(k) = [0, -]$. From Theorem 1 we know that $p(t)$ follows $p(k)$ and so $\Delta p(t) = [0, -]$. Thus we can answer the original question—when the patient is

‡ Which, incidentally, is exactly the same result as would have been given by the original monotonicity assumption. This fact is a consequence of probabilistic normalisation—the two assumptions give different results when not translating into or out of probability theory.

known to be in pain, evidence against vasculitic lesions may mean that she is less likely to be suffering from joint trauma.

6. Conclusions

The above results generalise the kind of qualitative propagation of values that may be carried out using a mixture of probability, possibility and belief values, making it possible to propagate in an evidential direction and between the causes of an observed effect. The work is useful for two reasons. Firstly this work has provided an analysis of the patterns of intercausal reasoning, such as "explaining away", in possibility and evidence theories— something that has not been previously attempted. This analysis has shown that explaining away occurs under specific, but very similar, circumstances in probability and evidence theories when the latter employes Dempster's rule of combination, and may always occur in evidence theory if Smets' disjunctive rule is employed along with binary variables. This semi-obligatory[§] nature of explaining away when the disjunctive rule is used rules out other forms of intercausal reasoning that are possible in probability theory and when Dempster's rule is used. Intercausal reasoning is also observed in possibility theory, albeit in a limited way such that "explaining away" does not occur, and again this only occurs under specific circumstances. Comparing the results with those of work in qualitative probabilistic networks it seems that the approach discussed here, whilst broadly being a generalisation of the work of Wellman[6] and Druzdzel and Henrion[20] captures a slightly different notion of intercausal reasoning. Secondly, this work extends the range of situations in which it is possible to integrate information expressed in different formalisms from cases of predictive reasoning[12,13] to any situation in which the dependency between variables can be expressed using a singly connected network. This means that the approach now has a much wider scope, and can be applied to a much wider range of problems.

Acknowledgements

This work was partially supported by Ann Radda when I lost my job due to illness, and has benefited from the comments of many people including Didier Dubois, Marek Druzdzel, Pascale Fonck, Frank Klawonn, Rudolph Kruse, Henri Prade, Alessandro Saffiotti, David Spiegelhalter, and Mike Wellman, as well as a number of anonymous referees. Many thanks to all of them.

References

1. Benferhat, S., Dubois, D. and Prade, H. "Argumentative inference in uncertain and inconsistent knowledge bases", *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, 1993, pp 411–419.
2. Darwiche, A. "Argument calculus and networks", *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, 1993, pp 420–427.

[§] Semi-obligatory because it is only obligatory that one cause may vary inversely with the other, rather than one cause having to vary inversely with the other.

3. Fox, J., Krause, P. and Ambler, S. "Arguments, contradictions and practical reasoning", *Proceedings of the 10th European Conference on Artificial Intelligence*, Vienna, 1992, pp 623–627.
4. Goldszmidt, M. "Qualitative probabilities: a normative framework for commonsense reasoning", PhD Thesis, UCLA, 1992.
5. Druzdzel, M. J. and Henrion, M. "Efficient reasoning in qualitative probabilistic networks", *Proceedings of the 11th National Conference on Artificial Intelligence*, Washington, 1993, pp 548–553.
6. Wellman, M. P. *Formulation of tradeoffs in planning under uncertainty*, (Pitman, London, 1990).
7. Henrion, M. and Druzdzel, M. J. "Qualitative propagation and scenario-based approaches to explanation of probabilistic reasoning", *Proceedings of the 6th Conference on Uncertainty in AI*, Boston, 1990, pp 10–20.
8. Henrion, M., Provan, G., Del Favero, B. and Sanders, G. "An experimental comparison of numerical and qualitative probabilistic reasoning", *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, Seattle, 1994, pp 319–326.
9. Michelena, N. "Monotonic influence diagrams: application to optimal and robust design", PhD Thesis, Berkeley, 1991.
10. Pearl, J. *Probabilistic reasoning in intelligent systems: networks of plausible inference* (Morgan Kaufman, San Mateo, 1988).
11. Cooper, G. F. "The computational complexity of probabilistic inference using belief networks", *Artificial Intelligence* **42** (1990) 393–405.
12. Parsons, S. and Mamdani, E. H. "On reasoning with qualitative uncertainty in networks", *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, 1993, pp 435–442.
13. Parsons, S. "Some elements of the theory of qualitative possibilistic networks", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **2** (1994) 67–90.
14. Parsons, S. *Qualitative methods for reasoning under uncertainty* (MIT Press, Cambridge, to appear).
15. Wellman, M. P. Personal communication.
16. Bobrow, D. *Qualitative reasoning about physical systems* (Elsevier Publishers Ltd, Amsterdam, 1984).
17. Dubois, D. and Prade, H. *Possibility Theory: An Approach to Computerised Processing of Uncertainty* (Plenum Press, New York 1988).
18. Zadeh, L. A. "Fuzzy sets as a basis for a theory of possibility", *Fuzzy Sets and Systems* **1** (1978) 1–28.
19. Shafer, G. *A mathematical theory of evidence* (Princeton University Press, Princeton 1976).
20. Druzdzel, M. J. and Henrion, M. "Intercausal reasoning with uninstantiated ancestor nodes", *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, Washington, 1993, pp 317–325.
21. Wellman, M. P. and Henrion, M. "Qualitative intercausal relations, or explaining 'explaining away'", *Proceedings of the 2nd Conference on Knowledge Representation and Reasoning*, 1991, pp 535–546.
22. Fonck, P. and Straszecka, E. "Building influence networks in the framework of possibility theory", *Annales Univ. Sci. Budapest., Sect. Comp.* **12** (1990) 101–106.
23. Shafer, G., Shenoy, P. P., and Mellouli, K. "Propagating belief functions in qualitative Markov trees", *International Journal of Approximate Reasoning*, **1** (1987) 349–400.
24. Smets, Ph. "Belief functions: the disjunctive rule of combination and the generalized bayesian theorem", *International Journal of Approximate Reasoning* **9** (1993) 1–35.

25. Dubois, D. and Prade, H. "Modelling uncertainty and inductive inference: a survey of recent non-additive probability systems", *Acta Psychologica*, **68** (1988) 53–78.
26. Kreyszig, E. *Advanced engineering mathematics*, 5th Edition, (John Wiley & Sons, New York, 1983).
27. Spiegelhalter, D. J. Personal communication.
28. Dubois, D. and Prade, H. "Possibilistic abduction", in *IPMU '92– Advanced Methods in Artificial Intelligence*, eds. B. Bouchon-Meunier, L. Valverde, and R. R. Yager (Springer-Verlag, 1993) pp 3–12.
29. Smets, Ph. "Belief functions", in *Non-Standard Logics for Automated Reasoning*, eds. Ph. Smets, E. H. Mamdani, D. Dubois and H. Prade (Academic Press, London, 1988) pp 253–277.
30. Tzeng, C-H. "A plausible reasoning in bayesian networks", *Proceedings of the 10th European Conference on Artificial Intelligence*, Vienna, 1992, pp 51–52.
31. Xu, H. and Smets, Ph. "Evidential reasoning with conditional belief functions". *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence*, Seattle, 1994, pp 598–605.
32. Parsons, S. and Saffiotti, A. "Integrating uncertainty handling formalisms in distributed artificial intelligence", *Proceedings of the 2nd European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*, Granada, Spain, 1993, pp 304–309.
33. Agustí-Cullell, J., Esteva, F., Garcìa, P., Godò, L., and Sierra, C. "Combining multiple-valued logics in modular expert systems", *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*, Los Angeles, 1991, pp 17–25.
34. Agustí-Cullell, J., Esteva, F., Garcìa, P., Godò, L., Lòpez de Mántaras, R., Puyol, J., and Sierra, C. "Structured local fuzzy logics in MILORD", in *Fuzzy Logic for the Management of Uncertainty*, eds. L. A. Zadeh and J. Kacprzyk (John Wiley & Sons, New York, 1992) pp 523–551.
35. Prade, H. Personal communication.
36. Saffiotti, A. Personal communication.