Defining normative systems for qualitative argumentation

Simon Parsons[1 2]

[1] Advanced Computation Laboratory, Imperial Cancer Research Fund,
P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom.
[2] Department of Electronic Engineering, Queen Mary and Westfield College,
Mile End Road, London E1 4NS, United Kingdom.
S.Parsons@qmw.ac.uk

Abstract. Inspired by two different approaches to providing a qualitative method for reasoning under uncertainty—qualitative probabilistic networks and systems of argumentation—this paper attempts to combine the advantages of both by defining systems of argumentation that have a probabilistic semantics.

1 Introduction

In the last few years there have been a number of attempts to build systems for reasoning under uncertainty that are of a qualitative nature—that is they use qualitative rather than numerical values, dealing with concepts such as increases in belief and the relative magnitude of values. In particular, two types of qualitative system have become well established, namely qualitative probabilistic networks (QPNs) [4, 18], and systems of argumentation [8, 11, 12]. While the former are built as an abstraction of probabilistic networks where the links between nodes are only modelled in terms of the qualitative influence of the parents on the children, and therefore have an underlying probabilistic semantics, some of the latter lack such a sound foundation. Instead they offer a greater degree of resolution, allowing more precise deductions to be made.

In this paper we present several normative systems of argumentation. These are systems of argumentation which have a probabilistic semantics, and are thus normative in that they behave according to the norms of probability theory. Such systems aim to extend both QPNs in the sense of reducing the degree of abstraction of the former, and argumentation in the sense of providing it with a probabilistic semantics whilst using only qualitative or semi-qualitative information[3]. Of course this extension might not always be desired, but may be useful at times to ensure that a given system reasons within probabilistic norms. The systems are built upon the framework introduced by Fox, Krause and their colleagues [8, 11], and we begin by introducing this framework.

[3] If we don't have any commitment to qualitative information, we can use ordinary probabilities as suggested by Krause *et al.* [11].

2 Introducing systems of argumentation

In classical logic, an argument is a sequence of inferences leading to a conclusion. If the argument is correct, then the conclusion is true. Consider the simple database Δ_1 which expresses some very familiar information in a Prolog-like notation in which variables are capitalised and ground terms and predicate names start with small letters.

$$\begin{array}{ll} f1 : human(socrates). & \Delta_1 \\ r1 : human(X) \rightarrow mortal(X). & \end{array}$$

The argument $\Delta_1 \vdash mortal(socrates)$ may be correctly made from this database because $mortal(socrates)$ follows from Δ_1 given the usual logical axioms and rules of inference. Thus a correct argument simply yields a conclusion which in this case could be paraphrased '$mortal(socrates)$ is true in the context of $f1$ and $r1$'. In the system of argumentation proposed by Fox and colleagues [11] this traditional form of reasoning is extended to allow arguments to indicate support and doubt in propositions, as well as proving them, by assigning labels to arguments which denote the confidence that the arguments warrant in their conclusions. This form of argumentation may be summarised by the following schema:

$$\text{Database } \vdash_{ACR} (\text{Sentence}, \text{Grounds}, \text{Sign})$$

where $\vdash_{ACR}$ is a suitable consequence relation. Informally, Grounds (G) are the facts and rules used to infer Sentence (St), and Sign (Sg) is a number or a symbol drawn from a dictionary of possible numbers or symbols which indicate the confidence warranted in the conclusion.

To formalise this kind of reasoning we start with a language, and we will take $\mathcal{L}$, a set of propositions, including $\perp$, the contradiction. We also have a set of connectives $\{\rightarrow, \neg\}$[4], and the following set of rules for building the well formed formulae of the language:

- If $l \in \mathcal{L}$ then l is a well formed formula (*wff*).
- If $l \in \mathcal{L}$ then $\neg l$ is a *wff*.
- If $l, m \in \mathcal{L}$ then $l \rightarrow m$, $l \rightarrow \neg m$, $\neg l \rightarrow m$ and $\neg l \rightarrow \neg m$ are *wff*s.
- Nothing else is a *wff*.

The members of $\mathcal{W}$, the set of all *wff*s that may be defined using $\mathcal{L}$, may then be used to build up a database Δ where every item $d \in \Delta$ is a triple $(i : l : s)$ in which i is a token uniquely identifying the database item (for convenience we will use the letter 'i' as an anonymous identifier), l is a wff, and s is a sign. With this formal system, we can take a database and use the argument consequence relation $\vdash_{ACR}$ given in Figure 1 to build arguments for propositions in $\mathcal{L}$ that we are interested in.

[4] Note that both the set of connectives and the rules for building *wff*s are more restrictive than for other similar systems of argumentation [11]. A normative system which does not suffer from these limitations is discussed in [13].

$$\text{Ax}\frac{}{\Delta \vdash_{ACR} (St, \{i\}, Sg)} \; (i : St : Sg) \in \Delta$$

$$\to\text{-E}\frac{\Delta \vdash_{ACR} (St, G, Sg) \quad \Delta \vdash_{ACR} (St \to St', G', Sg')}{\Delta \vdash_{ACR} (St', G \cup G', \mathsf{comb}(Sg, Sg'))}$$

$$\to\text{-I}\frac{\Delta \cup (St, \emptyset, Sg) \vdash_{ACR} (St', G, Sg')}{\Delta \vdash_{ACR} (St \to St', G, \mathsf{comb}'(Sg, Sg'))}$$

Fig. 1. Argumentation Consequence Relation

Typically we will be able to build several arguments for a given proposition, and so, to find out something about the overall validity of the proposition, we will *flatten* the different arguments to get a single sign.

Together $\mathcal{L}$, the rules for building the formulae, the connectives, and $\vdash_{ACR}$ define a formal system of argumentation, which, for want of a name we will call $\mathcal{SA}$. In fact, $\mathcal{SA}$ is really the basis of a family of systems of argumentation, because one can define a number of variants of $\mathcal{SA}$ by using different dictionaries of signs. Each dictionary will have its own combination functions comb and comb', and its own means of flattening arguments, and the meanings of the signs, the flattening function, and the combination function delineate the semantics of the system of argumentation. Thus $\mathcal{SA}$ gives us a general framework for expressing logical facts which can incorporate different models of uncertainty by varying the signs and their associated combination and flattening functions as well as a means of representing default information and of handling inconsistent information [15].

3 A first normative system

One commonly used system of argumentation within the framework of $\mathcal{SA}$ is one in which the dictionary consists of three symbols, $+$, $-$ and 0, which represent the notion of an increase, a decrease and no change in belief respectively. When a proposition is labelled with $+$, it is taken to represent the fact that there is an increase in belief in the proposition, while labelling the rule:

$$human \to mortal$$

with a $+$ is taken to represent the fact that showing that there is an increase in belief in the proposition "*human*" causes an increase in belief in the proposition "*mortal*". The combination function comb for this system of argumentation is $\otimes$ of Table 1, while comb' is its inverse $\otimes^{-1}$, also given in Table 1—blank spaces mark impossible combinations. As with all combinator papers in this paper, the first argument of the function described by the table is drawn from the first column, and the second argument is drawn from the first row. The flattening

$\otimes$	$+$	0	$-$	$?$
$+$	$+$	0	$-$	$?$
0	0	0	0	0
$-$	$-$	0	$+$	$?$
$?$	$?$	0	$?$	$?$

$\otimes^{-1}$	$+$	0	$-$	$?$
$+$	$+$	0	$-$	$?$
0		$?$		
$-$	$-$	0	$+$	$?$
$?$		0		$?$

Table 1. The functions $\otimes$ and $\otimes^{-1}$

function is one that implements a form of improper linear model with uniform weights and no constant term [3]. This counts the number of $+$ and $-$ weighted arguments for a proposition, takes the sign that occurs most often and makes that the sign of the proposition, thus taking the sum of all the arguments while giving each argument equal weight. We will refer to the system of argumentation which uses this dictionary and set of functions along with the argument building capabilities of $\mathcal{SA}$ as $\mathcal{NA}_1$. The question that faces us here is how $\mathcal{NA}_1$ may be given a probabilistic semantics. Now, the use of $+$ and $-$ to represent changes in belief suggests a link between this system of argumentation and QPNs [18] since the latter make use of a similar notion. Indeed, it turns out that we can modify the notion of a probabilistic influence in a QPN to give our database facts and rules a probabilistic interpretation. In particular we take triples $(i : l : +)$, where $l \in \mathcal{W}$ and l does not include the connective $\rightarrow$, to denote the fact that $\Pr(l)$ is known to increase, and similar triples $(i : l : -)$, to denote the fact that $\Pr(l)$ is known to decrease. Triples $(i : l : 0)$, clearly denote the fact that $\Pr(l)$ is known to neither increase nor decrease. With this interpretation facts correspond to the nodes in a QPN, and as in QPNs we deal with changes in their probability.

Database rules can similarly be given a probabilistic interpretation by making the triple $(i : n \rightarrow m : +)$, where m and n are members of $\mathcal{W}$ which do not include the connective $\rightarrow$, denote the fact that:

$$\Pr(m \,|\, n, X) \geq \Pr(m \,|\, \neg n, X)$$

for any $X \in \{x, \neg x\}$ for which there is a triple $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$ (where s is any sign), while the triple $(i : n \rightarrow m : -)$ denotes the fact that:

$$\Pr(m \,|\, n, X) \leq \Pr(m \,|\, \neg n, X)$$

again for any X for which there is a triple $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$. We do not make use of triples such as $(i : n \rightarrow m : 0)$ since such rules have no useful effect. As a result a rule $(i : n \rightarrow m : +)$ means that there is a probability distribution over the propositions m and n such that an increase in the probability of n makes m more likely to be true, and a rule $(i : n \rightarrow m : -)$ means that there is a probability distribution over the propositions m and n such that an increase in the probability of n makes m less likely to be true. With this interpretation, rules correspond to qualitative influences in QPNs. It should be noted that the effect of declaring that there is a rule $(i : n \rightarrow m : +)$

is to create considerable constraints on the probability distribution over m and n to the extent that the effect of other rules relating m and n are determined absolutely. That is, a necessary consequence of $(i : n \to m : +)$ is that we have constraints on the conditional probability distributions across the propositions in the database equivalent to the rules $(i : \neg n \to m : -)$, $(i : n \to \neg m : -)$ and $(i : \neg n \to \neg m : +)$, and similar restrictions are imposed by rules like $(i : n \to m : -)$.

With this interpretation of rules and facts, the combination function $\otimes$ has a natural probabilistic interpretation as the function by which changes in probability are combined with probabilistic influences. Indeed $\otimes$ is the function used to combine the two in QPNs, and $\otimes^{-1}$ remains as defined above. The flattening function also has an obvious probabilistic interpretation in terms of calculating the overall change in probability of a proposition. However, in order for the improper linear model to make sense probabilistically, it is necessary to apply a restriction to the sizes of changes in probability represented by $(i : l : +)$ and $(i : l : -)$. In particular, it requires that all arguments, irrespective of how many steps they contain, have the same strength. This clearly places a great restriction on the number of probabilistic models that can be captured by this system.

In the kind of minimal logic we have taken as the base language for our system, any negated formula $\neg l$ is taken as shorthand for a formula $l \to \bot$. Thus in our system we should replace any formula $(i : \neg l : s)$ with $(i : l \to \bot : s)$, and any formula $(i : \neg l \to m : s)$ with $(i : (l \to \bot) \to m : s)$ before constructing any arguments. However, our probabilistic semantics give us an alternative way of replacing negated propositions, since $(i : \neg l : +) \equiv (i : l : -)$ and $(i : \neg l \to m : +) \equiv (i : l \to m : -)$, which does not involve introducing the contradiction (from which $x \to \bot$ follows where x is any proposition), and this is the method we prefer.

As an example of the kind of reasoning that can be performed in $\mathcal{NA}_1$, consider the following simple database Δ_2 of propositional rules and facts. What these rules say is that there are three events that may have an effect on whether or not I lose my job—I have a high research output, I am ill, and I am a good tutor to my students. The first and third have a negative influence on the probability of me losing my job, while the other has a positive influence. The database facts say that there is evidence which leads to an increased probability of me being ill and having a high research output, and a decreased probability of me being a good tutor:

$f1 : high\_research\_output : +.$ $\qquad \Delta_2$
$f2 : ill : +.$
$f3 : good\_tutor : -.$
$r1 : high\_research\_output \to lose\_job : -.$
$r2 : ill \to lose\_job : +.$
$r3 : good\_tutor \to lose\_job : -.$

From Δ_2 we can build the arguments:

$$\Delta_2 \vdash_{ACR} (lose\_job, (f1, r1), -).$$
$$\Delta_2 \vdash_{ACR} (lose\_job, (f2, r2), +).$$
$$\Delta_2 \vdash_{ACR} (lose\_job, (f3, r3), +).$$

And the improper linear model will flatten them to come up with the overall conclusion that there is an increase in the probability of me losing my job after the facts of my situation are known.

4 A second normative system

As stated above, $\mathcal{NA}_1$ is highly restrictive because its flattening function requires all arguments to have the same strength. To relax this restriction we clearly need a new flattening function. One suitable function is that used by QPNs for combining the effect of several influences on one variable. This function is $\oplus$ as specified in Table 2. The use of this function to define a new system of argumentation $\mathcal{NA}_2$ is straightforward after the dictionary of signs is extended to become $\{+, -, 0, ?\}$ where labelling a fact with ? indicates that the change in probability of that fact is unknown, and a rule $(i : n \rightarrow m : ?)$ denotes that the relationship between $\Pr(m \mid n, X)$ and $\Pr(m \mid \neg n, X)$ is unknown, so that if the probability of n increases it is not possible to say how the probability of m will change.

With this interpretation, there is a direct correspondence between a database of formulae drawn from $\mathcal{W}$ and a qualitative probabilistic network, and it is quite easy to see that any conclusion drawn by $\mathcal{NA}_2$ from a database would also be drawn by the corresponding QPN. The fact that qualitative multiplication distributes over addition ensures that the fact that argumentation builds separate arguments for the same proposition and then flattens them does not mean that it gives a different answer to the equivalent QPN.

To illustrate the difference between $\mathcal{NA}_1$ and $\mathcal{NA}_2$, consider what $\mathcal{NA}_2$ would conclude from Δ_2. Firstly it would build the same arguments as $\mathcal{NA}_1$:

$$\Delta_2 \vdash_{ACR} (lose\_job, (f1, r1), -).$$
$$\Delta_2 \vdash_{ACR} (lose\_job, (f2, r2), +).$$
$$\Delta_2 \vdash_{ACR} (lose\_job, (f3, r3), +).$$

$\oplus$	+	0	−	?
+	+	+	?	?
0	+	0	−	?
−	?	−	−	?
?	?	?	?	?

Table 2. The function $\oplus$

But this time the flattening function would conclude that the overall change in belief in the proposition *lose_job* was ?, indicating that it cannot be accurately identified. This is, of course, probabilistically correct—without information on the relative effects of the various causes of a loss of job, the way in which its probability will change cannot be predicted.

5 A more subtle normative system

Now, in the kind of applications for which $\mathcal{SA}$ was developed [7, 9], it is necessary to represent information of the form "X is known to be true", and "If X is true then Y is true"—information that we might term categorical. It is therefore interesting to investigate if $\mathcal{NA}_2$ can be extended to cover categorical relationships. To do so we first extend the dictionary of signs to be $\{++, +, -, --\}$ as suggested in [8, 11], where $++$ and $--$ are labels for categorical information. It then turns out that we can give $++$ and $--$ a probabilistic semantics, giving a system of argumentation $\mathcal{NA}_3$ which is $\mathcal{NA}_2$ extended by allowing triples such as $(i : l : ++)$ and $(i : l : --)$ and rules such as $(i : n \rightarrow m : ++)$ and $(i : n \rightarrow m : --)$.

The meaning of $(i : l : ++)$, where l is a *wff* which does not contain $\rightarrow$, is that the probability of l becomes 1, and $(i : l : --)$ means that the probability of l decreases to 0, and to make this clear, we write $(i : l : \overline{\uparrow})$ for $(i : l : ++)$, and $(i : l : \underline{\downarrow})$ for $(i : l : --)$. The meaning of the rules is slightly more complicated. We want a rule $(i : n \rightarrow m : ++)$, where neither m or n contain $\rightarrow$, to denote a constraint on the probability distribution across m and n such that if $\Pr(n)$ becomes 1, so does $\Pr(m)$. This requires that:

$$\Pr(m \,|\, n, X) = 1$$

for all $X \in \{x, \neg x\}$ such that the database contains $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$. [13]. Similarly, a probabilistic interpretation of a rule $(i : n \rightarrow m : --)$ requires that:

$$\Pr(m \,|\, n, X) = 0$$

for all $X \in \{x, \neg x\}$ such that the database contains $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$. Considering the constraints on the conditional probabilities imposed by $++$ and $--$ rules, a further pair of rules are suggested. These are a rule $(i : n \rightarrow m : -+)$ which requires that:

$$\Pr(m \,|\, \neg n, X) = 1$$

for all $X \in \{x, \neg x\}$ such that the database contains $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$ (s now being able to take any value in the set $\{++, +-, +, -, -+, --\}$), and a rule $(i : n \rightarrow m : +-)$ which requires that:

$$\Pr(m \,|\, \neg n, X) = 0$$

for all $X \in \{x, \neg x\}$ such that the database contains $(i : x \rightarrow m : s)$ or $(i : \neg x \rightarrow m : s)$. Once again, the introduction of such rules imposes restrictions on other

$\otimes_*$	$++$	$+-$	$+$	0	$-$	$-+$	$--$	$?$
$\overline{\uparrow}$	$\overline{\uparrow}$	$+$	$+$	0	$-$	$-$	$\underline{\downarrow}$	$?$
$+$	$+$	$+$	$+$	0	$-$	$-$	$-$	$?$
0	0	0	0	0	0	0	0	0
$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$?$
$\underline{\downarrow}$	$-$	$\underline{\downarrow}$	$-$	0	$+$	$\overline{\uparrow}$	$+$	$?$
$?$	$?$	$?$	$?$	0	$?$	$?$	$?$	$?$

$\otimes_*^{-1}$	$\overline{\uparrow}$	$+$	0	$-$	$\underline{\downarrow}$	$?$
$\overline{\uparrow}$	$++$	$+$	0	$-$	$--$	$?$
$+$		$+$	0	$-$		$?$
0			0			
$-$		$-$	0	$+$		$?$
$\underline{\downarrow}$	$-+$	$-$	0	$+$	$+-$	$?$
$?$			0			$?$

Table 3. Variants of $\otimes$ and $\otimes^{-1}$

rules involving the same propositions so that $(i : n \to m : ++)$ implies that there must be restrictions equivalent to the rules $(i : \neg n \to m : --)$, $(i : n \to \neg m : --)$ and $(i : \neg n \to \neg m : ++)$, and similar restrictions are imposed by the other rules. As before, having introduced new qualitative values and ensured that they have a probabilistic meaning, we have to give a suitably probabilistic means of combining them if we want the whole system to be normative. It is reasonably clear that suitable functions comb and comb' are those variants of $\otimes$ and $\otimes^{-1}$ given in Table 3 [13]. Note the asymmetry in the tables. Once again, the base logic compels us to replace negated literals before constructing any arguments, and again we do this by replacing facts and rules. It is clear that facts and rules with signs $+$ and $-$ are handled as before, and that $(i : \neg l : \overline{\uparrow}) \equiv (i : l : \underline{\downarrow})$. Categorical rules are also handled by the appropriate substitution using, for instance, the equivalencies $(i : l \to \neg m : ++) \equiv (i : l \to m : --)$, $(i : \neg l \to m : ++) \equiv (i : l \to m : -+)$ and $(i : \neg l \to \neg m : ++) \equiv (i : l \to m : +-)$ [13].

The correct way to flatten normative arguments, some of which are categorical, is a little complex. The problem is that the very strong constraint that a rule $(i : n \to m : ++)$ puts on the distribution over m and n greatly restricts the values of other rules whose consequent is m. In fact, if we have $(i : n \to m : s)$, $s \in \{++, -+\}$ then for any other $(i : x \to m : s')$, $s' \in \{++, +, -, -+\}$ and if we have $(i : n \to m : s)$, $s \in \{+-, --\}$ then for any other $(i : x \to m : s')$, $s' \in \{+-, +, -, --\}$ [13, 14]. This means that we have a revised flattening op-

$\oplus_*$	$\overline{\uparrow}$	$+$	0	$-$	$\underline{\downarrow}$	$?$
$\overline{\uparrow}$	$\overline{\uparrow}$	$\overline{\uparrow}$	$\overline{\uparrow}$	$\overline{\uparrow}$	U	$\overline{\uparrow}$
$+$	$\overline{\uparrow}$	$+$	$+$	$?$	$\underline{\downarrow}$	$?$
0	$\overline{\uparrow}$	$+$	0	$-$	$\underline{\downarrow}$	$?$
$-$	$\overline{\uparrow}$	$?$	$-$	$-$	$\underline{\downarrow}$	$?$
$\underline{\downarrow}$	U	$\underline{\downarrow}$	$\underline{\downarrow}$	$\underline{\downarrow}$	$\underline{\downarrow}$	$\underline{\downarrow}$
$?$	$\overline{\uparrow}$	$?$	$?$	$?$	$\underline{\downarrow}$	$?$

Table 4. A new flattening function

erator $\oplus_*$ as given in Table 4 where the symbol U indicates that the result is not defined. U may also be taken to indicate that if this is the result of flattening, then the database on which its deduction is based violates the laws of probability. Equipping $\mathcal{NA}_3$ with these extensions ensures that it is normative in the sense that all its conclusions will either be in accordance with probability theory or indicate that there has been a violation of the theory.

To see how the system incorporates categorical knowledge, consider the following variation on our example, which includes the categorical rule that being found to embezzle funds from my organisation would lead to me losing my job:

$$\begin{array}{ll} f1 : embezzle\_funds : \overline{\uparrow}. & \Delta_3 \\ f2 : good\_research\_output : \overline{\uparrow}. & \\ r1 : embezzle\_funds \rightarrow lose\_job : ++ . & \\ r2 : good\_research\_output \rightarrow lose\_job : -. & \end{array}$$

From this using $\mathcal{NA}_3$ we can build the arguments:

$$\Delta_3 \vdash_{ACR} (lose\_job, (f1, r1), \overline{\uparrow}).$$
$$\Delta_3 \vdash_{ACR} (lose\_job, (f2, r2), -).$$

which will flatten to tell us that I will definitely lose my job since the categorical positive effect of embezzling outweighs the negative effect of not being ill.

6 Using order of magnitude information

As the example of Δ_3 demonstrated, $\mathcal{NA}_3$ extends the kind of representation and reasoning provided by QPNs by allowing the explicit handling of categorical information. This is not the only extension that is possible. Another is to use some form of order of magnitude reasoning. This would make it possible to say, for instance, that because $\Pr(a)$ increases much more than $\Pr(b)$, and $\Pr(a)$ influences $\Pr(c)$ much more strongly than $\Pr(b)$ influences $\Pr(d)$, it is clear that $\Pr(c)$ will undergo a much larger change in value than $\Pr(d)$. A particularly appropriate system for performing this kind of reasoning, known as ROM[K], is provided by Dague [1]. ROM[K] works by manipulating expressions about the relative size of two quantities Q_1 and Q_2. There are four possible ways of expressing this relation: Q_1 is *negligible with respect to* Q_2, $Q_1 \ll Q_2$, Q_1 is *distant from* Q_2, $Q_1 \not\simeq Q_2$, Q_1 is *comparable to* Q_2, $Q_1 \sim Q_2$, and Q_1 is *close to* Q_2, $Q_1 \approx Q_2$. Once the relation between pairs of quantities is specified, it is possible to deduce new relations by applying the axioms and properties of ROM[K], some of which are reproduced in Figure 2.

We can use ROM[K] to define a system of argumentation $\mathcal{NA}_4$ which extends $\mathcal{NA}_2$ with relative order of magnitude reasoning about the size of the changes in probability with which the system deals. As usual, we need to define combination and flattening functions, though here they differ from those of other systems in that they are comparative and additional to those used by $\mathcal{NA}_2$. Once the argument is established as being $+$ or $-$ using the function $\otimes$ from $\mathcal{NA}_2$, this

(A1) $A \approx A$	(A9) $A \sim 1 \rightarrow [A] = [+]$
(A2) $A \approx B \rightarrow B \approx A$	(A10) $A \ll B \leftrightarrow B \approx (B + A)$
(A3) $A \approx B, B \approx C \rightarrow A \approx C$	(A11) $A \ll B, B \sim C \rightarrow A \ll C$
(A4) $A \sim B \rightarrow B \sim A$	(A12) $A \approx B, [C] = [A] \rightarrow (A + C) \approx (B + C)$
(A5) $A \sim B, B \sim C \rightarrow A \sim C$	(A13) $A \sim B, [C] = [A] \rightarrow (A + C) \sim (B + C)$
(A6) $A \approx B \rightarrow A \sim B$	(A14) $A \sim (A + A)$
(A7) $A \approx B \rightarrow C.A \approx C.B$	(A15) $A \not\approx B \leftrightarrow (A - B) \sim A$ or $(B - A) \sim B$
(A8) $A \sim B \rightarrow C.A \sim C.B$	
(P3) $A \ll B \rightarrow C.A \ll C.B$	(P26) $A \sim B \rightarrow B \sim A$
(P35) $A \not\approx B \rightarrow C.A \not\approx C.B$	(P38) $A \not\approx B, C \approx A, D \approx B \rightarrow C \not\approx D$

Fig. 2. Some of the axioms and properties of ROM[K]

new combination function comb_* gives the relation between the changes based on the strength of the influences that cause the change while comb'_* may be used to identify the relation between the influences of two rules based upon the relation between the changes in probability of their antecedents and the change in probability of their consequents. Similarly, the new flattening function identifies the greatest influence on a given hypothesis allowing a ? caused by two conflicting arguments to resolved into a + or a −. The combination function comb_* is defined in Table 5—if the change in $\Pr(a)$ stands in relation rel_1 to the change in $\Pr(b)$ (where rel_1 is one of the relations of ROM[K]) and the strength of the influence of $\Pr(a)$ on $\Pr(c)$ stands in relation rel_2 to the strength of the influence of $\Pr(b)$ on $\Pr(d)$ (rel_2 also being one of the relations of ROM[K]), then the relation rel_3 between the changes in $\Pr(c)$ and $\Pr(d)$ is given by the combinator table. Note that Table 5 only covers the cases in which the change in $\Pr(a)$ is less than or equal to that in $\Pr(b)$ and the strength of the influence between $\Pr(a)$ and $\Pr(c)$ is less than or equal to that of $\Pr(b)$ on $\Pr(d)$. Obvious permutations of the table will cover the other cases. Also note that the letter V indicates that rel_3 may not be determined from the particular values of rel_1 and rel_2 because to make any prediction would be to step outside the bounds of probability.

Table 5 also defines comb'_*, which as before is the inverse of comb_* with rel_1 and rel_3 determining rel_2. Note that for some combinations of input, the output

	rel$_2$			
comb_*	$\approx$	$\sim$	$\not\approx$	$\ll$
rel$_1$ $\approx$	$\approx$	$\sim$	$\not\approx$	$\ll$
rel$_1$ $\sim$	$\sim$	$\sim$	V	$\ll$
rel$_1$ $\not\approx$	$\not\approx$	V	V	$\ll$
rel$_1$ $\ll$	$\ll$	$\ll$	$\ll$	$\ll$

	rel$_3$			
comb'_*	$\approx$	$\sim$	$\not\approx$	$\ll$
rel$_1$ $\approx$	$\approx$	$\sim$	$\not\approx$	$\ll$
rel$_1$ $\approx$	$\approx, \sim$			$\ll$
rel$_1$ $\not\approx$			$\approx$	
rel$_1$ $\ll$				$\approx, \sim, \not\approx, \ll$

Table 5. Combining ROM[K] relations.

is ambiguous.

For the flattening function, if the change in $\Pr(a)$ stands in relation rel_4 to the change in $\Pr(b)$ (where rel_4 is one of the relations of ROM[K]) and the strength of the influence of $\Pr(a)$ on $\Pr(c)$ stands in relation rel_5 to the strength of the influence of $\Pr(b)$ on $\Pr(c)$ (rel_5 also being one of the relations of ROM[K]), the sign of the change in $\Pr(c)$ is given in Table 6 (where $[\Delta \Pr(b)]$ indicates the sign of the change in $\Pr(b)$). Note that Table 6 only covers the cases in which the change in $\Pr(a)$ is less than or equal to that in $\Pr(b)$ and the strength of the influence between $\Pr(a)$ and $\Pr(c)$ is less than or equal to that of $\Pr(b)$ on $\Pr(c)$. Obvious permutations of the table will cover the other cases.

As an example of the kind of reasoning that may be performed in $\mathcal{NA}_4$, consider the following variant of our running example.

$$
\begin{array}{ll}
f1 : ill : +. & \Delta_4 \\
f2 : embezzle\_funds : -. & \\
r1 : ill \rightarrow lose\_job : +. & \\
r2 : ill \rightarrow hospital : +. & \\
r3 : embezzle\_funds \rightarrow lose\_job : +. &
\end{array}
$$

In addition, consider we know that the relationship between the strengths of $r1$ and $r2$ is $\ll$, while the changes in probability implied by $f1$ and $f2$ stand in relation $\sim$. From the database we can build the arguments:

$$
\begin{array}{l}
\Delta_4 \vdash_{ACR} (lose\_job, (f1, r1), -). \\
\Delta_4 \vdash_{ACR} (hospital, (f1, r2), +). \\
\Delta_4 \vdash_{ACR} (lose\_job, (f2, r2), +).
\end{array}
$$

using the combination function from $\mathcal{NA}_2$. Considering the first two arguments, the comb_* may then be used to establish which has stronger support. Since both arguments are based upon the same fact, rel_1 is '$\approx$', so that we can conclude that the relation rel_3 between the changes in probability of 'hospital' and $lose\_job$ must be '$\ll$' so that the increase in belief that I will lose my job is much smaller than the increase in belief that I will go to hospital. Similarly, flattening the arguments for $lose\_job$ with the old flattening function will give ?, while the new flattening function will establish that the probability of $lose\_job$ will increase as can be seen by looking at the intersection of $\not\approx$ and $\approx$ in Table 6.

		rel5 $\approx$	$\sim$	$\not\approx$	$\ll$
rel4	$\approx$	?	?	$[\Delta \Pr(b)]$	$[\Delta \Pr(b)]$
	$\sim$	?	?	?	$[\Delta \Pr(b)]$
	$\not\approx$	$[\Delta \Pr(b)]$	?	?	$[\Delta \Pr(b)]$
	$\ll$	$[\Delta \Pr(b)]$	$[\Delta \Pr(b)]$	$[\Delta \Pr(b)]$	$[\Delta \Pr(b)]$

Table 6. How to flatten arguments in ROM[K]

7 Using numerical information

Further precision may be obtained by incorporating numerical information about the size of changes in probability and the strengths of influences. Inspired by Dubois *et al.* [6], we build a new system of argumentation $\mathcal{NA}_5$ with the same base language as the other systems, but which has a dictionary which includes a set of "linguistic" [5] labels, each of which is an identifier for an interval probability, and may be used to give the strength of rules. A suitable set is:

Strongly Positive	$\geq$	Weakly Positive	$\geq$	Zero	$\geq$	Weakly Negative	$\geq$	Strongly Negative
(SP)		(WP)		(Z)		(WN)		(SN)
$(1, \alpha]$	$\geq$	$[\alpha, 0)$	$\geq$	0	$\geq$	$(0, -\alpha]$	$\geq$	$[-\alpha, 1)$

though we could take any set of intervals we desire—a larger set will give us a finer degree of resolution but be more tedious to use as an example. Note that the open intervals explicitly do not allow the modelling of categorical influences (if these are required we can simply add additional labels at either end of the scale). The dictionary also includes a second set of labels which quantify changes in probability:

Complete Positive	$\geq$	Big Positive	$\geq$	Medium Positive	$\geq$	Little Positive	$\geq$	Zero
(CP)		(BP)		(MP)		(LP)		(Z)
1	$\geq$	$(1, 1-\beta]$	$\geq$	$[1-\beta, \beta]$	$\geq$	$[\beta, 0)$	$\geq$	0

The definition of the changes Little Negative (LN), Medium Negative (MN), Big Negative (BN) and Complete Negative (CN) are symmetrical, and again we could use a different set if desired. Like the other systems of argumentation, $\mathcal{NA}_5$ uses the argument consequence relation $\vdash_{ACR}$ to build arguments for hypotheses, and so in order to be able to determine the strength of arguments we must define combination functions comb and comb' which say how to combine the "linguistic" labels. To do so we must first choose suitable values of α and β, and on the grounds that we would like our intervals to be evenly sized, we choose $\beta \approx 0.33$ and $\alpha \approx 0.5$. This then gives us the combination functions of Table 7 where [MP, LP] stands for the interval whose upper limit is the upper limit of MP and whose lower limit is the lower limit of LP. Results of combining with negative influences and changes can be obtained by symmetry.

To combine several arguments for one proposition we need a suitable flattening function, and this is provided by interval addition. Furthermore, if we are to use the precision of the system we need a way to compare intervals in order to identify which arguments have the greatest support. This may be done using $\leq_{int}$ where $[a, b] \leq_{int} [c, d]$ iff $a \leq c$ and $b \leq d$ [5]. To illustrate the use of $\mathcal{NA}_5$

[5] The scare quotes denoting that no claim is being made that the probability intervals with which we deal are in any way related to interpretations of natural language—we are just adopting Dubois *et al.*'s terminology.

consider the database:

$$
\begin{array}{ll}
f1 : embezzle\_funds : CP. & \qquad \Delta_5 \\
f2 : ill : BP. & \\
r1 : embezzle\_funds \rightarrow lose\_job : SP. & \\
r2 : ill \rightarrow lose\_job : WP. & \\
r3 : ill \rightarrow hospital : SP. &
\end{array}
$$

From this we can build the arguments:

$$\Delta_5 \vdash_{ACR} (lose\_job, (f1, r1), [BP, MP]).$$
$$\Delta_5 \vdash_{ACR} (lose\_job, (f2, r2), [MP, LP]).$$
$$\Delta_5 \vdash_{ACR} (hospital, (f2, r3), [BP, MP]).$$

The two arguments for *lose_job* may be flattened to give the overall value of [CP, MP] and using $\leq_{int}$ we learn that the increase in probability of *hospital* is less than or equal to that of *lose_job*.

8 Discussion

This paper began with the claim that it would present a number of normative systems of argumentation, taking this to mean that they have a probabilistic semantics, and that they would thus be an improvement on non-normative systems of argumentation for those cases in which such norms are desirable. Furthermore, the claim was made that these systems would also be an improvement on qualitative systems for reasoning with probability such as QPNs since they would allow more precise predictions to be made. In the event five different systems, $\mathcal{NA}_1$–$\mathcal{NA}_5$, which meet these objectives to varying degrees, have been presented.

$\mathcal{NA}_1$, uses a probabilistic notion of qualitative influences between variables to give meaning to logical rules. The fact that $\mathcal{NA}_1$ has a strict probabilistic semantics means that it is an extension of non-normative systems of argumentation. However, the restrictions on the meaning of the rules imposed by the improper linear model mean that $\mathcal{NA}_1$ is not an extension of QPNs. $\mathcal{NA}_2$ is a system of argumentation which is roughly equivalent to QPNs. Thus $\mathcal{NA}_2$ whilst an extension of non-normative systems of argumentation on which it is based, is not an extension of QPNs.

The problem of extending QPNs was addressed by $\mathcal{NA}_3$. The extension takes the form of allowing the representation of categorical influences between variables. Giving these a qualitative representation and a probabilistic meaning

comb	CP	BP	MP	LP	Z
SP	[BP, MP]	[BP, MP]	[MP, LP]	LP	Z
WP	[MP, LP]	[MP, LP]	[MP, LP]	LP	Z
Z	Z	Z	Z	Z	Z

comb′	[BP, MP]	[MP, LP]	LP	Z
SP	[CP, BP]	MP	LP	Z
WP		[CP, MP]	LP	Z
Z				Z

Table 7. Combining "linguistic" labels

makes $\mathcal{NA}_3$ a system which is both normative and can represent and reason with a wider range of information than is possible in a QPN whilst retaining the latter's qualitative nature. Thus it meets overall objectives of the paper. Two further extensions were introduced in the form of $\mathcal{NA}_4$ and $\mathcal{NA}_5$ which use order of magnitude and interval information respectively.

9 Relation to other work

There are a number of connections with the work of other authors. The close relation between qualitative approaches to probabilistic reasoning in networks and probabilistic systems based on logic was suggested by Wellman [17] while the idea of a database of influences which is equivalent to a probabilistic network has been discussed by, among others, Poole [16] and Wong [19]. The attempt to give an essentially logical system a probabilistic semantics makes our efforts similar to Goldszmidt's work on normative systems for defeasible reasoning [10]. This clearly has some similarities with our work, but differs in its intent. Goldszmidt aims to build defeasible systems whose behaviour is justified by their probabilistic semantics while we are intent on a more general system. The use of a probabilistic semantics is not our only goal—we are just interested in being able to provide a normative system when one is required, with the choice of alternative combination and flattening functions allowing a broad range of possible systems to be adopted. In addition, our work has strong connections with that of Darwiche [2], this time differing in the way it is approached. His aim was "...to relax the commitment to numbers while retaining the desirable features of probability theory", which is rather different to the aim of the work described here. We started from the opposite position, taking a completely abstract model of reasoning and seeing how it could be instantiated to behave in a probabilistic way if so desired (which often it won't be since probability theory often imposes overly strict constraints for the kind of reasoning that argumentation was designed to provide), and the fact that we did so suggests that the work presented here and that in [2] are to some extent complementary.

Acknowledgements

This work was partially supported by Esprit Basic Research Action 6156 DRUMS II (Defeasible Reasoning and Uncertainty Management Systems). Many thanks to Paul Krause, John Fox, Kathy Laskey and Jack Breese for making me think more clearly about some of the issues discussed in this paper.

References

1. Dague, P. (1993) Symbolic reasoning with relative orders of magnitude, *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, Chambery, France.

2. Darwiche, A. (1993) A symbolic generalization of probability theory, Ph.D. Thesis, Stanford.
3. Dawes, R. (1979) The robust beauty of improper linear models, *American Psychologist*, **34**, 571–582.
4. Druzdzel, M. J. and Henrion, M. (1993) Efficient reasoning in qualitative probabilistic networks, *Proceedings of the 11th National Conference on Artificial Intelligence*, Washington.
5. Dubois, D. and Prade, H. (1979) Fuzzy real algebra: some results, *Fuzzy sets and systems*, **2**, 327–348.
6. Dubois, D., Prade, H., Godo, L., and Lopez de Mantaras, R. (1992) A symbolic approach to reasoning with linguistic quantifiers, *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, Stanford.
7. Fox, J. (1990) Automating assistance for safety critical decisions, *Philosophical Transactions of the Royal Society*, B, **327**, 555–567.
8. Fox, J., Krause, P. and Ambler, S. (1992) Arguments, contradictions and practical reasoning, *Proceedings of the 10th European Conference on Artificial Intelligence*, Vienna.
9. Glowinski, A., O'Neil, M., and Fox, J. (1987) Design of a generic information system and its application to primary care, *Proceedings of AIME Conference*, Marseille.
10. Goldszmidt, M. (1992) Qualitative probabilities: a normative framework for commonsense reasoning, Ph.D. Thesis, UCLA.
11. Krause, P., Ambler, S., Elvang-Gøransson, M., and Fox, J. (1995) A logic of argumentation for reasoning under uncertainty, *Computational Intelligence*, **11**, 113–131.
12. Loui, R. P. (1987) Defeat among arguments: a system of defeasible inference, *Computational Intelligence*, **3**, 100–106.
13. Parsons, S. (1996) Normative argumentation and qualitative probability, Technical Report 317, Advanced Computation Laboratory, Imperial Cancer Research Fund.
14. Parsons, S. (1995) Refining reasoning in qualitative probabilistic networks, *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, Montreal.
15. Parsons, S. and Fox, J. (1994) A general approach to managing imperfect information in deductive databases, *Proceedings of the Workshop on Uncertainty in Databases and Deductive Systems*, Ithaca, NY.
16. Poole, D. (1991) Representing Bayesian networks within probabilistic horn abduction, in *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*, Los Angeles, CA.
17. Wellman, M. P. (1994) Some varieties of qualitative probability, *Proceedings of the 5th International Conference on Information Processing and the Management of Uncertainty*, Paris.
18. Wellman, M. P. (1990) *Formulation of tradeoffs in planning under uncertainty*, Pitman, London.
19. Wong, S. K. M., Xiang, Y., and Nie, X. (1994) Representation of Bayesian networks as relational databases, *Proceedings of the 5th International Conference on Information Processing and the Management of Uncertainty*, Paris.