Comparing normative argumentation to other probabilistic systems

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

Abstract

This paper discusses a system of argumentation with a probabilistic semantics and compares it to two other probabilistic systems—Wellman's qualitative probabilistic networks and Neufeld's probabilistic default reasoning.

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—qualitative probabilistic networks (QPNs) [2, 12], and systems of argumentation [5, 6]. 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. This lack of a probabilistic semantics for argumentation prompted work [8, 9] to provide such a semantics for systems of argumentation of the kind introduced by Krause *et al.* [5] using only qualitative or semi-qualitative information[1]. 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. This paper further extends this work by comparing the most well-developed of these normative systems of argumentation with two similar systems—qualitative probabilistic networks and Neufeld's probabilistic commonsense reasoning [7]. The paper begins with a recap of the systems of argumentation upon which the normative system is built.

[1] If there is no commitment to qualitative information, it is possible to give argumentation a semantics in terms of numerical probabilities [5].

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 $\Delta_1 \vdash mortal(socrates)$ is correct 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 Krause *et al.* 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 $\{\rightarrow, \neg\}$[2], and the following set of rules for building the well formed formulae of the language:

[2] Note that both the set of connectives and the rules for building *wffs* are more restrictive than for other similar systems of argumentation [5], but these restrictions may be lifted [9].

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

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

Figure 1: Argumentation Consequence Relation

- 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 given in Figure 1, to build arguments for propositions in $\mathcal{L}$ that we are interested in.

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 we will call $\mathcal{SA}'$ since it is a cut-down version of the system $\mathcal{SA}$ introduced in [8]. 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 function 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.

3 A NORMATIVE SYSTEM

One commonly used system of argumentation within the framework of $\mathcal{SA}'$ is one in which the dictionary includes 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:

$$human(x) \rightarrow mortal(x)$$

with a $+$ is taken to represent the fact that showing that there is an increase in the belief of something being human causes an increase in belief that it is mortal.

Now, the use of $+$ and $-$ to represent changes in belief suggests a link between this system of argumentation and QPNs [12] 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 $p(l)$ is known to increase, and similar triples $(i : l : -)$, to denote the fact that $p(l)$ is known to decrease. Triples $(i : l : 0)$, clearly denote the fact that $p(l)$ 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 $(i : n \rightarrow m : +)$ is that we have other rules $(i : n \rightarrow \neg m : -)$, $(i : \neg n \rightarrow m : -)$ and $(i : \neg n \rightarrow \neg m : +)$, and similar restrictions are imposed by rules like $(i : n \rightarrow m : -)$.

Now, in some applications [3], 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

$\otimes_*$	$++$	$+-$	$+$	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	$?$	$?$	$?$	$?$

Table 1: The function $\otimes_*$

$\oplus_*$	$\overline{\uparrow}$	$+$	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 2: The function $\oplus_*$

might term categorical. To do this we first extend the dictionary of signs to be $\{++, +, -, --\}$ where $++$ and $--$ are labels for categorical information. It then turns out that we can give $++$ and $--$ a probabilistic semantics, giving a system of argumentation which includes 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 $(i : \neg x \rightarrow m : s)$. [9]. 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 rules involving the same propositions so that $(i : n \rightarrow m : ++)$ implies that there must be restrictions equivalent to the rules $(i : \neg n \rightarrow m : --)$, $(i : n \rightarrow \neg m : --)$ and $(i : \neg n \rightarrow \neg m : ++)$, and similar restrictions are imposed by the other rules.

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 a suitable comb will be the function $\otimes_*$ given in Table 1. Note the asymmetry in the table, and the addition of the sign '?' to indicate a change in probability whose value is not known.

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 $(i : n \rightarrow m : ++)$ puts on the distribution over m and n greatly restricts the values of other rules whose consequent is m. In fact, if we have $(i : n \rightarrow m : s)$, $s \in \{++, -+\}$ then for any other $(i : x \rightarrow m : s')$, $s' \in \{++, +, -, -+\}$ and if we have $(i : n \rightarrow m : s)$, $s \in \{+-, --\}$ then for any other $(i : x \rightarrow m : s')$, $s' \in \{+-, +, -, --\}$ [9]. This means that we have a flattening operator $\oplus_*$ as given in Table 2 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.

We will call the system of argumentation which uses this dictionary and pair of functions along with the argument building capabilities of $\mathcal{SA}'$ as $\mathcal{NA}'_3$ since it bears the same relation to $\mathcal{NA}_3$ [8] as $\mathcal{SA}'$ does to $\mathcal{SA}$.

As an example of the kind of reasoning that can be performed in $\mathcal{NA}'_3$, consider the following simple database Δ_2 of propositional rules and facts. What these rules say is that there are three events that may influence my losing my job—I embezzle funds, I am ill, I am an illegal alien. All of these events have a positive influence on my losing my job, so that if any single one of them on their own becomes more believable, it is more believable that I will lose my job, and, conversely, if they become less believable, it is less believable that I will lose my job.

$f1 : embezzle\_funds : -.$ $\quad\quad\quad \Delta_2$
$f2 : ill : +.$
$f3 : illegal\_alien : -.$
$r1 : embezzle\_funds \rightarrow lose\_job : +.$
$r2 : ill \rightarrow lose\_job : +.$
$r3 : illegal\_alien \rightarrow lose\_job : +.$

The database facts say that there is reason to increase belief in that fact that I am ill, and that there are reasons to decrease belief in that fact that I have embezzled funds, and am an illegal alien. From Δ_2 we can build the arguments:

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

and these will be flattened to conclude that the overall change in belief in the proposition '*lose_job*' was ?, indicating that it cannot be accurately identified. To see how the system incorporates categorical knowledge, consider the following variation on our example:

$$\begin{array}{ll} f1 : embezzle\_funds : \overline{\uparrow}. & \Delta_3 \\ f2 : ill : \underline{\downarrow}. & \\ r1 : embezzle\_funds \rightarrow lose\_job : ++ . & \\ r2 : ill \rightarrow lose\_job : +. & \end{array}$$

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

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

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

In the kind of minimal logic which forms the base language for $\mathcal{NA}'_3$ any negated formula $\neg l$ is taken as shorthand for $l \rightarrow \bot$ so all triples $(i : \neg l : s)$ should be replaced with $(i : l \rightarrow \bot : s)$, and any formula $(i : \neg l \rightarrow m : s)$ with $(i : (l \rightarrow \bot) \rightarrow m : s)$ before constructing any arguments. However, since $(i : \neg l : +) \equiv (i : l : -)$, $(i : \neg l : \overline{\uparrow}) \equiv (i : l : \underline{\downarrow})$, and $(i : \neg l \rightarrow m : +) \equiv (i : l \rightarrow m : -)$ we can avoid introducing the contradiction by using the appropriate substitution. Categorical rules are similarly handled using, for instance, the equivalencies $(i : l \rightarrow \neg m : ++) \equiv (i : l \rightarrow m : --)$, $(i : \neg l \rightarrow m : ++) \equiv (i : l \rightarrow m : -+)$ and $(i : \neg l \rightarrow \neg m : ++) \equiv (i : l \rightarrow m : +-)$ [9].

4 COMPARISON WITH OTHER SYSTEMS

Now, $\mathcal{NA}'_3$ clearly bears some relation to other probabilistic systems for dealing with changes in probability, most notably qualitative probabilistic networks (QPNs) [12], from which it borrows the notion of what consititutes a probabilistic connection between variables, and Neufeld's probabilistic commonsense reasoner [7]. This section aims to establish the form of this relation.

4.1 QUALITATIVE PROBABILISTIC NETWORKS

Formally, a QPN is a pair $G = (V, Q)$, where V is a set of variables or nodes in a graph, denoted by capital letters, and Q is a set of qualitative relations among the variables. There are two types of qualitative relations in Q, "influences" and "synergies". Qualitative

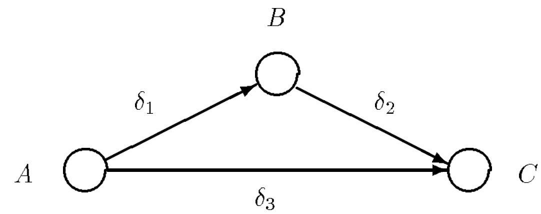

Figure 2: A qualitative probabilistic network

influences define the sign of the direct influence between variables and correspond to arcs in a probabilistic network and are defined as follows. We say that "A positively influences C", written $S^+(A, C)$, iff for all values $a_1 > a_2$, c_0, and X, which is the set of all of C's predecessors other than A:

$$\Pr(c \geq c_0 \,|\, a_1, X) \geq \Pr(c \geq c_0 \,|\, a_2, X)$$

where a_i and c_j are the possible values of A and C. This definition expresses the fact that increasing the value of A makes higher values of C more probable. Negative qualitative influence, S^-, and zero qualitative influence, S^0, are defined analogously by substituting $\leq$ and $=$, respectively, for $\geq$. For binary valued variables, A positively influences C if

$$\Pr(c \,|\, a, X) \geq \Pr(c \,|\, \neg a, X)$$

which means that $S^+(A, C)$ in a QPN has exactly the same meaning as $(i : a \rightarrow c : +)$ in an $\mathcal{NA}'_3$ database. Thus it is possible to represent any given binary valued QPN as a set of facts and rules in $\mathcal{NA}'_3$—nodes in the QPN are propositions, and influences are rules. Given this, the question that it seems reasonable to ask is "can the same conclusions be drawn from a QPN and its representation in $\mathcal{NA}'_3$?" Well, Wellman [12], paraphrasing slightly, specifies that in the situation in Figure 2 the conclusions to be drawn about the change in probability of c, given a change in probability of δ_4 in the probability of a is:

$$((\delta_1 \otimes_* \delta_2) \oplus_* \delta_3) \otimes_* \delta_4$$

where δ_1, δ_2 and δ_3 are the signs of the influences between A and B, B and C, and A and C respectively, since the combination functions given by Wellman are exactly those used by $\mathcal{NA}'_3$, except that they only deal with $+$, 0, $-$ and ?. If this same QPN were represented in $\mathcal{NA}'_3$, we would have:

$$\begin{array}{ll} f1 : a : \delta_4. & \Delta_4 \\ r1 : a \rightarrow b : \delta_1. & \\ r2 : b \rightarrow c : \delta_2. & \\ r3 : a \rightarrow c : \delta_3. & \end{array}$$

from which is possible to build two arguments for c

$$\begin{array}{lll} \Delta_4 & \vdash_{ACR} & (c, (f1, r1, r2), (\delta_1 \otimes_* \delta_2 \otimes_* \delta_4)). \\ \Delta_4 & \vdash_{ACR} & (c, (f1, r3), (\delta_3 \otimes_* \delta_4)). \end{array}$$

Now, flattening these two arguments for c tells us that the overall change in the probability of c is:

$$(\delta_1 \otimes_* \delta_2 \otimes_* \delta_4) \oplus_* (\delta_3 \otimes_* \delta_4)$$

which, since $\otimes_*$ and $\oplus_*$ distribute like ordinary multiplication and addition, is the same as the change established using QPNs. Since all possible QPNs updates can be reduced to that pictured in Figure 2 [12], it is clear that $\mathcal{NA}'_3$ can completely capture binary QPNs where the only relations are qualitative influences. Furthermore, it is clear that $\mathcal{NA}'_3$ go considerably further than QPNs in distinguishing categorical influences. However, $\mathcal{NA}'_3$ can only deal with binary valued variables whereas QPNs can deal with many-valued and continuous variables, and $\mathcal{NA}'_3$ cannot represent qualitative synergies. These latter are hyperrelations which capture the fact that, for instance, the joint effect of variables A and B on C is greater than the sum of their individual effects. In binary QPNs this is expressed by saying that A and B exhibit positive synergy with respect to C, written $Y^+(\{A, B\}, C)$ if:

$$\Pr(c \,|\, a, b) + \Pr(c \,|\, \neg a, \neg b) \geq \Pr(c \,|\, \neg a, b) + \Pr(c \,|\, a, \neg b)$$

negative synergy $Y^-(\{A, B\}, C)$ and zero synergy $Y^0(\{A, B\}, C)$ are similarly defined. It is currently unclear how, if at all, this kind of relation may be captured in $\mathcal{NA}'_3$.

4.2 THE PROBABILISTIC COMMONSENSE REASONER

Neufeld [7] bases his probabilistic commonsense reasoner on a graphical notation which distinguishes the following kinds of relations that may hold between variables:

$$\begin{array}{lll} b \rightarrow a & \text{means} & 1 > \Pr(a \,|\, b) > \Pr(a) \\ b \Rightarrow a & \text{means} & 1 = \Pr(a \,|\, b) > \Pr(a) \\ b \not\rightarrow a & \text{means} & 1 > \Pr(\neg a \,|\, b) > \Pr(\neg a) \\ b \not\Rightarrow a & \text{means} & 1 = \Pr(\neg a \,|\, b) > \Pr(\neg a) \end{array}$$

These are clearly very similar to the rules in $\mathcal{NA}_3$. The semantics of Neufeld's $b \rightarrow a$ are virtually the same as that of $\mathcal{NA}'_3$'s $(i : b \rightarrow a : +)$. To see this consider that $(i : b \rightarrow a : +)$ means:

$$\begin{aligned} \Pr(a \,|\, b, X) &\geq \Pr(a \,|\, \neg b, X) \\ &\geq \Pr(a \,|\, X) \\ &\geq \Pr(a) \end{aligned}$$

Now,

$$\Pr(a \,|\, b) = \sum_{x \in X} \Pr(a \,|\, b, x) \Pr(x)$$

so, if $\Pr(a \,|\, b, X) \geq \Pr(a)$ for all values of X:

$$\begin{aligned} \Pr(a \,|\, b) &\geq \sum_{x \in X} \Pr(a) \Pr(x) \\ &\geq \Pr(a) \sum_{x \in X} \Pr(x) \\ &\geq \Pr(a) \end{aligned}$$

so $(i : b \rightarrow a : +)$ captures the same meaning as Neufeld's $b \rightarrow a$ in that if the restriction on the probabilities of a and b implied by $(i : b \rightarrow a : +)$ is met, the restriction implied by $b \rightarrow a$ is also met (provided, of course, that we know that $\Pr(a \,|\, b, X) \neq \Pr(a)$). In a similar way, $(i : b \rightarrow \neg a : +)$ captures the meaning of $b \not\rightarrow a$, $(i : b \rightarrow a : ++)$ captures the meaning of $b \not\rightarrow a$ and $(i : b \rightarrow \neg a : ++)$ captures the meaning of $b \not\Rightarrow a$. However, there are differences between the systems.

The most obvious is the use of the strict inequality in Neufeld's system, and this is clearly a very minor difference. The use of the non-strict inequality in $\mathcal{NA}'_3$ is really just for convenience, and of course it is the reason that the condition on the probabilities imposed by $(i : b \rightarrow a : ++)$ is stated as:

$$\Pr(a \,|\, b, X) = 1$$

rather than:

$$1 = \Pr(a \,|\, b, X) \geq \Pr(a \,|\, \neg b, X)$$

which would make the connection with $b \Rightarrow a$ clearer—the second expression is a trivial consequence of the first.

The major difference between the two systems, as Neufeld [7] points out (with reference to QPNs), is that he does not force the condition to hold for all the possible events that might influence a other than b. In other words, in $\mathcal{NA}'_3$, if we have $(i : b \rightarrow a : +)$ and $(i : c \rightarrow a : +)$, then:

$$\Pr(a \,|\, b, c) \geq \Pr(a)$$

so that a becomes more probable when a and b are known to be true. Within the framework of normative argumentation and QPNs this is a very useful property since it makes it possible to carry out purely local calculations—if we have $(i : b \rightarrow a : +)$ then we can calculate the effect of a change in $\Pr(b)$ on $\Pr(a)$ without having to simultaneously consider the effects of changes in the probabilities of other propositions that influence it in order to avoid "double-counting" the effects of evidence as we do in standard probability theory (other influences, of course, are not ignored, just considered locally in their turn).

In Neufeld's system, however, this property does not hold, so that it is perfectly possible to have $b \rightarrow a$ and $c \rightarrow a$ and yet have:

$$\Pr(a \,|\, b, c) < \Pr(a)$$

Now, within the framework of Neufeld's system it is this latter property which is seen as being good since it ensures that the system does not draw any unwarranted conclusions about the probability of c given the conjunction of a and b. To see that this may be a good thing, consider the following scenario.

Jack is having a party, and as far as he is concerned the party will be a good one if his best friends come. His two best friends are Cody and Evelyn. With this knowledge we can build a simple model of Jack's beliefs using the variables *good_party*, *cody_comes*, and *evelyn_comes*. The model will contain the two implications (in Neufled's sense) *cody_comes* $\rightarrow$ *good_party* and $evelyn\_comes \rightarrow good\_party$. Now, with Neufeld's scheme it is not possible to infer that:

$$\Pr(good\_party) < \Pr(good\_party \,|\, cody\_comes, evelyn\_comes)$$

which is just as well for Jack since Cody and Evelyn are divorced, and if they come to the same party they will fight and upset him (since he wishes that the two of them still got along). If this kind of conservative reasoning is required, then Neufeld's system has an advantage over $\mathcal{NA}'_3$ since the latter could not be used to represent Jack's beliefs—it could only be used if it were the case that in Jack's view the party will be good as soon as either of Cody or Evelyn turn up, regardless of who else comes. Of course, making it possible to deal with qualitative synergies between variables in $\mathcal{NA}'_3$ would make it possible to model Jack's beliefs since the effect of Cody and Evelyn both coming to the party may be modelled as $Y^-(\{cody\_comes, evelyn\_comes\}, good\_party)$.

5 DISCUSSION

This paper has discussed a normative system of argumentation $\mathcal{NA}'_3$, and explored the relationship between it and two similar systems. In particular, it has shown that $\mathcal{NA}'_3$ can represent binary qualitative probabilistic networks [12] which do not contain any qualitative synergies (or, equivalently, for which all the possible synergies are zero), and can represent the same kind of information as Neufeld's probabilistic commonsense reasoner, although there are slight differences which mean that it can handle situations which defeat Neufeld's system and is defeated by some situations which Neufeld's system can handle.

Clearly these are not the only other systems which $\mathcal{NA}'_3$ ressembles. The attempt to give an essentially logical system a probabilistic semantics prompts recollection of Goldszmidt's work on normative systems for defeasible reasoning [4]. In addition, our work has strong connections with Darwiche's [1] move "...to relax the commitment to numbers while retaining the desirable features of probability theory". Furthermore, the close relation between qualitative approaches to probabilistic reasoning in networks and probabilistic systems based on logic was suggested by Wellman [11] while the idea of a database of influences which is equivalent to a probabilistic network has been discussed by, among others, Poole [10] and Wong [13].

Finally, it should be noted that the base language of the system of argumentation $\mathcal{SA}$ is more restrictive than that of other similar systems. This is because we exclude formulae that include the $\wedge$ and $\vee$ connectives, complex formulae such as $(a \rightarrow b) \rightarrow (c \rightarrow d)$, and only have a very limited set of rules of inference. The latter problem is partly tackled in [8], while the introduction of $\wedge$ and complex formulae is addressed in [9]. Other extensions of the system are the subject of ongoing work.

Acknowledgements

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

References

[1] Darwiche, A. (1993) A symbolic generalization of probability theory, PhD. Thesis, Stanford.

[2] Druzdzel, M. J. and Henrion, M. (1993) Efficient reasoning in qualitative probabilistic networks, *Proceedings of the 11th National Conference on Artificial Intelligence*, Washington.

[3] Fox, J. (1990) Automating assistance for safety critical decisions, *Philosophical Transactions of the Royal Society*, B, **327**, 555–567.

[4] Goldszmidt, M. (1992) Qualitative probabilities: a normative framework for commonsense reasoning, PhD Thesis, UCLA.

[5] Krause, P., Ambler, S., Elvang-Gøransson, M., and Fox, J. (1995) A logic of argumentation for reasoning under uncertainty, *Computational Intelligence*, **11**, 113–131.

[6] Loui, R. P. (1987) Defeat among arguments: a system of defeasible inference, *Computational Intelligence*, **3**, 100–106.

[7] Neufeld, E. (1990) A probabilistic commonsense reasoner, *International Journal of Intelligent Systems*, **5**, 565–594.

[8] Parsons, S. (1996) Defining normative systems for qualitative argumentation, *Proceedings of the International Conference on Formal and Applied Practical Reasoning*, Bonn.

[9] Parsons, S. (1996) Normative argumentation and qualitative probability, Technical Report 317, Advanced Computation Laboratory, Imperial Cancer Research Fund.

[10] Poole, D. (1991) Representing Bayesian networks within probabilistic horn abduction, in *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*, Los Angeles, CA.

[11] Wellman, M. P. (1994) Some varieties of qualitative probability, *Proceedings of the 5th International Conference on Information Processing and the Management of Uncertainty*, Paris.

[12] Wellman, M. P. (1990) *Formulation of tradeoffs in planning under uncertainty*, Pitman, London.

[13] 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.