Arguing about beliefs and actions

John Fox[1] and Simon Parsons[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,
University of London,
London E1 4NS,
United Kingdom.

Abstract. Decision making under uncertainty is central to reasoning by practical intelligent systems, and attracts great controversy. The most widely accepted approach is to represent uncertainty in terms of prior and conditional probabilities of events and the utilities of consequences of actions, and to apply standard decision theory to calculate degrees of belief and expected utilities of actions. Unfortunately, as has been observed many times, reliable probabilities are often not easily available. Furthermore the benefits of a quantitative probabilistic representation can be small by comparison with the restrictions imposed by the formalism. In this paper we summarise an approach to reasoning under uncertainty by constructing arguments for and against particular options and then describe an extension of this approach to reasoning about the expected values of actions.

1 Introduction

Standard decision theory [35] builds on the probabilistic view of uncertainty in reasoning about actions. The costs and benefits of possible outcomes of actions are weighted with their probabilities, yielding a preference ordering on the "expected utility" of alternative actions. However, as Tan and Pearl [40], amongst others, have pointed out, the specification of the complete sets of probabilities and utilities required by standard decision theory make the theory impractical in complex tasks which involve common sense knowledge. This realisation has prompted work on qualitative approaches to decision making which attempt to reduce the amount of numerical information required.

Work on such qualitative decision making techniques has been an established topic of research at the Imperial Cancer Research Fund since the early 80s (see [31] for a review). Our early work was partly concerned with the description of human decision processes [12] and partly with the practical development of decision systems for use in medicine [13]. Whilst the qualitative decision procedures we developed proved to have considerable descriptive value and practical promise, our desire to build decision support systems for safety-critical fields such as medicine raised the concern that our early applications were *ad hoc*. In particular we were concerned that they, in common with all other expert systems being built at the time, were not based on a rigorously defined decision theory. As a result we have put considerable effort into developing a theoretical framework for qualitative decision making. The best developed part of this is an approach to uncertainty and belief based on the idea of *argumentation*. This approach emphasizes the *construction* and *aggregation* of symbolic arguments based on the non-standard logic LA [18, 22]. This provides rules for constructing reasons to believe in and doubt hypotheses, and reasons to believe or doubt arguments.

The generality of the everyday idea of argumentation suggests that a similar approach could be taken to reasoning about actions, for instance in deciding on medical treatments or investigations. We might hope to construct arguments for and against alternative actions in the usual way, avoiding issues about the elicitation and use of numerical utilities by representing the desirability and undesirability of actions symbolically. This suggestion immediately raises two questions:

- How well does our formalisation of support and opposition transfer to reasoning about action?
- Is LA directly applicable to arguments about action or will different logics be required?

This paper attempts to provide some answers to these questions. In particular it argues that while there are similarities between arguments for and against beliefs and arguments for and against actions, there are also significant differences which amount to a requirement for additional rules for assigning values to the outcomes of actions, and for arguing the expected benefits of alternative actions. The paper then makes an initial attempt to suggest a framework for handling such rules, as well as summarising some of the applications developed using argumentation, and discussing one set of tools that are available for building such applications.

However, before starting this work, the paper first sets the discussion in context by recalling the logic of argumentation about beliefs, LA, and its relation to argumentation in general.

2 The logic of argument LA

Our approach to decision making was to seek a rapprochement between the purely quantitative and purely logical traditions, seeking a form of uncertainty management which people find natural, yet one which can be shown to be mathematically sound and general. This approach was based on *argumentation*, the familiar form of reasoning which is based on everyday patterns of debate. It turned out, however, that this approach was not new.

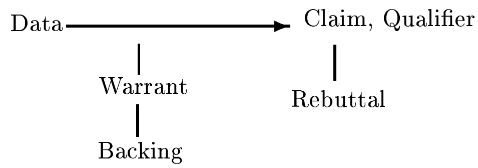

Fig. 1. The Toulmin argument schema

2.1 The nature of arguments

The philosopher Stephen Toulmin explored the question of why traditional formal models of reasoning have apparently little relevance to everyday dispute and debate, concluding that *argumentation* is a human form of reasoning distinct from both probabilistic reasoning and classical deduction. Toulmin characterised argumentation by means of the informal schema in Figure 1. This can be illustrated by the example (the italics are ours):

> in support of the *claim* that Harry is a British subject, we appeal to the *datum* that he was born in Bermuda, and ... (the claim is *warranted* by a sentence such as) ... "A man born in Bermuda may be taken to be a British subject": since, however, questions of nationality are always subject to qualifications and conditions we shall have to insert a *qualifying* "presumably" in front of the conclusion and note the possibility that our conclusion may be *rebutted* in case it turns out that both his parents were aliens or he has since become a naturalised American. Finally, in case the warrant itself is challenged, its *backing* can be put in: this will record the terms and the dates of enactment of the Acts of Parliament and other legal provisions governing the nationality of persons born in the British colonies ([42] p 104).

Two points are prominent here; the idea that in general conclusions are not certain, hence the qualifier "presumably", and that practical reasoning frequently involves contradictions among arguments (the notion of rebuttal). This is in contrast, as we have seen, to the usual approach to modeling uncertainty with a quantitative measure; Toulmin's approach anticipated the interest in symbolic representations of uncertainty in artificial intelligence and logic. Toulmin also anticipated another recent development in artificial intelligence, the desire to get to grips with the concept of contradiction. In classical logic and probability contradiction is eschewed; something cannot be both true and false nor have a probability of 0 and 1.

In attempting to address the practical problems of decision making in medical domains we faced similar problems to those identified by Toulmin. First, we have to make decisions in the face of uncertainty in situations where it is impractical to state the degree of uncertainty. Second, it is common in practical settings to have to deal with apparent contra-indications where from one point of view something is definitely the case whereas from another point of view it is definitely not the case, or at one time something is held to be true while at another it is considered false.

2.2 Arguments about beliefs

Toulmin's analysis was perceptive but from our point of view it is clearly inadequate since it is entirely informal. What we need is a formalisation which preserves the basic ideas while giving it sound mathematical foundations. In this section we work towards such a formalisation by providing an informal account of what our formal system, LA, provides.

We start with the notion of an argument in a standard logic such as propositional logic, first-order predicate calculus, or a modal logic such as T, $S4$ or $S5$ [21]. In such a logic, L, an argument is a sequence of inferences leading to a conclusion. If the argument is correct, then the conclusion is true. An argument:

$$G_1 \ldots G_n \vdash St$$

is correct in the logic L if St may be derived using the rules of inference and axioms of L augmented with $G_1 \ldots G_n$. Therefore a correct argument simply yields a proposition St. This can be paraphrased as

$$St \text{ is true (in the context } G_1 \ldots G_n)$$

In the approach we take, this traditional form of logic based argumentation is extended in two important ways:

1. to allow arguments not only to prove propositions but also to merely indicate support for, or even doubt in, them; and
2. by explictly recording the context in which the deduction holds.

The way we do this is by borrowing from the idea of a labelled deduction system [20].

A labelled deductive system is essentially an enriched logical system, where formulae can be labelled, thereby adding structure to logical theories (usually called databases). Both formulas and labels can be manipulated independently; the exact correspondence is made explicit in the way labelled formulae are constructed by inference rules. To see how this helps, consider a situation in which we have the following pieces of information:

$$\mathit{lost\_weight} : a_1$$
$$\mathit{lost\_weight} \rightarrow \mathit{cancer} : r_1$$

where *cancer* is an abbreviation for "the patient has cancer", and *lost_weight* is an abbreviation for "the patient has lost weight". In a labelled deductive system we can derive the proposition *cancer* and denote this by:

$$\mathit{cancer} : (a_1, r_1)$$

so the label $(a1, r1)$ is a label which represents the proof of *cancer* by identifying the database items used in the proof. This takes care of recording the context of the proof; it is contained in the label.

The other thing that we need to do is to allow arguments to just indicate support for, or doubt in, propositions. Here we just use a second label which designates the confidence warranted by the arguments for their conclusions. There is nothing in the theory of labelled deductive systems which precludes the use of a number of labels, and this simple mechanism allows confidences to be expressed in a variety of representations without modifying the underlying inference system. Thus the result of a derivation is an *argument* of the form:

$$(St : G : Sg)$$

Each argument consists of a triple consisting of a Sentence (St), which is the claim in Toulmin's terminology, Grounds (G), which are the formulae used to justify the argument, and a Sign (Sg), which is a number or a symbol which indicates the confidence warranted in the conclusion. The idea of argumentation from a database may thus be summarised by the following schema:

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

In this schema, $\vdash_{ACR}$ is a consequence relation which defines the inference rules by which we may construct arguments for claims using the information in the database.

The use of confidences rather than logical proofs introduces a slight complication. In classical logic, if we can construct an argument (proof) for St then any further arguments for St are of no interest since St is known to be true. If, however, we only have an indication of support for St then it may be the case that additional information casts doubt on St. Thus we need to consider every distinct argument concerning St and then carry out a process of *aggregation* to combine them. This process is also known as *flattening* since it has the effect of mapping a number of distinct arguments into a single measure. One intuitively plausible way of doing this aggregation is to assume that the more independent grounds we have for St, the greater our confidence in St may reasonably be, and so we assess the strength of confidence in St in some applications of LA by simply summing the number of arguments for St. Ambler [2] gives a rigorous justification for this procedure in category theoretic terms.

2.3 Formalising argumentation about beliefs

Having spoken informally about what LA is attempting to do, we present a formal description of LA. This is broadly the same system as that discussed in [22], but this version is less influenced by Ambler's work on the category theoretic basis of argumentation, and is more influenced by labelled deductive systems and the style of presentation used in recent work on argumentation [27, 28]. However, the differences between the two versions of the system are largely cosmetic.

We start with a set of atomic propositions $\mathcal{L}$ including $\top$ and $\bot$, the ever true and ever false propositions. We also have the set of connectives $\{\neg, \rightarrow, \wedge\}$, and the following set of rules for building the well-formed formulae (*wff*s) of the language:

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

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 $(St : G : Sg)$ in which St is a *wff*, Sg represents confidence in St, and G are the grounds on which the assertion is made. With this formal system, we can take a database and use the argumentation consequence relation $\vdash_{ACR}$ defined in Figure 2 to build arguments for propositions that we are interested in. This consequence relation is defined in terms of rules for building new arguments from old. The rules are written in a style similar to standard Gentzen proof rules, with the antecedents of the rule above the line and the consequent below. Thus if the arguments above the line may be made, then the argument below the line may also be made. In detail the rules are as follows:

- The rule Ax says that if the triple $(St : G : Sg)$ is in the database, then it is possible to build the argument $(St : G : Sg)$ from the database. The rule thus allows the construction of arguments from database items.
- The rule $\wedge$-I says that if the arguments $(St : G : Sg)$ and $(St' : G' : Sg')$ may be built from a database, then an argument for $St \wedge St'$ may also be built. The rule thus says how to introduce arguments about conjunctions.
- The rule $\wedge$-E1 says that if it is possible to build an argument for $St \wedge St'$ from a database, then it is also possible to build an argument for St. Thus the rule allows the elimination of one conjunct from an argument.
- The rule $\wedge$-E2 is analogous to $\wedge$-E1 but allows the elimination of the other conjunct.
- The rule $\rightarrow$-I says that if on adding $(St, \emptyset, Sg)$, $\emptyset$ indicating that the triple has no grounds, to a database it is possible to conclude St', then there is an argument for $St \rightarrow St'$. The rule thus allows the introduction of $\rightarrow$ into arguments.
- The rule $\rightarrow$-E says that from an argument for St and an argument for $St \rightarrow St'$ it is possible to build an argument for St'. The rule thus allows the elimination of $\rightarrow$ from arguments and is analogous to modus ponens in standard propositional logic.

We use the term “ dictionary” to describe a set of symbols which can be used to label a proposition. If we define dictionary $\mathcal{D}$ by:

$$\mathcal{D} =_{\text{def}} \{S_1, \ldots S_n\}$$

then we may write:

$$(St : G : S_i)$$

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

$$\wedge\text{-I}\frac{\Delta \vdash_{ACR} (St : G : Sg) \; \Delta \vdash_{ACR} (St' : G' : Sg')}{\Delta \vdash_{ACR} (St \wedge St' : G \cup G' : \mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ intro}}(Sg, Sg'))}$$

$$\wedge\text{-E1}\frac{\Delta \vdash_{ACR} (St \wedge St' : G : Sg)}{\Delta \vdash_{ACR} (St : G : Sg)}$$

$$\wedge\text{-E2}\frac{\Delta \vdash_{ACR} (St \wedge St' : G : Sg)}{\Delta \vdash_{ACR} (St' : G : Sg)}$$

$$\rightarrow\text{-I}\frac{\Delta, (St : \emptyset : Sg) \vdash_{ACR} (St' : G : Sg')}{\Delta \vdash_{ACR} (St \rightarrow St' : G : \mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}(Sg, Sg'))}$$

$$\rightarrow\text{-E}\frac{\Delta \vdash_{ACR} (St : G : Sg) \; \Delta \vdash_{ACR} (St \rightarrow St' : G' : Sg'))}{\Delta \vdash_{ACR} (St' : G \cup G : \mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}(Sg, Sg'))}$$

Fig. 2. Argumentation Consequence Relation

where S_i is any symbol drawn from $\mathcal{D}$. Where there is the possibility of confusion between dictionaries we write $S_i^{\mathcal{D}}$ to denote the symbol S_i from dictionary $\mathcal{D}$. Among the obvious dictionaries we may consider are sets of numbers. Dictionaries for probabilities, possibilities [8], certainty factors [39], belief functions [37] are thus straightforwardly defined. They are, respectively:

$$\mathcal{D}_{\text{probability}} =_{\text{def}} \{S : S \in [0,1]\}$$
$$\mathcal{D}_{\text{possibility}} =_{\text{def}} \{S : S \in [0,1]\}$$
$$\mathcal{D}_{\text{belief functions}} =_{\text{def}} \{S : S \in [0,1]\}$$
$$\mathcal{D}_{\text{certainty factors}} =_{\text{def}} \{S : S \in [-1,1]\}$$

Systems of argumentation which are based on LA and have semantics in terms of both quantitative probability and possibility values have been defined [22]. However, there is no requirement that we should restrict dictionaries to sets of numbers. For example we have frequently adopted one of a number of simple symbolic dictionaries. The four simplest such dictionaries are described below.

Generic dictionary In a standard logical proof the value "true" is assigned to a sentence if it is possible to construct a proof for it from facts which are held to be "true". However, in practical situations, prior facts, and consequently any conclusions that can be deduced from them, can be in error. To capture this idea, we therefore substitute the sign + for "true" giving the simple dictionary:

$$\mathcal{D}_{\text{generic}} =_{\text{def}} \{+\}$$

We refer to arguments with sign + as *supporting* arguments. The argument:

$$(cancer : lost\_weight : +)$$

simply says "the fact that the patient has lost weight increases my confidence in her having cancer, but I cannot say by how much". Using the generic dictionary thus means that the "force" of different arguments cannot be distinguished. Suppose we have a number of arguments whose signs are drawn from the generic dictionary. Given we cannot distinguish between the force of the arguments, it seems reasonable to assume that:

Assumption A1 If $Args$ is any set of arguments concerning St, then[1]:

$$|Args \cup \{(St : G : +)\}| \geq |Args|$$

where $|Args|$ indicates the force of the set of arguments $Args$. The simple aggregation procedure mentioned above, in which we just count arguments for a proposition to assess our confidence in that proposition, conforms to this assumption.

Bounded generic dictionary With a large database, it will often be possible to construct a large number of arguments for a proposition. Intuitively, however, some arguments are conclusive, that is they leave no room for doubt with respect to their grounds (they may be rebutted on other grounds). To represent this we may define a more specialised dictionary, introducing an additional sign ++:

$$\mathcal{D}_{\text{bounded generic}} =_{\text{def}} \{+, ++\}$$

We refer to arguments with sign ++ as *confirming*. Informally, if we have a conclusive argument for some proposition then this argument will dominate the aggregation procedure. Thus a confirming argument is more forceful than any set of supporting arguments, and a set of confirming and supporting arguments is exactly as forceful as a single confirming argument. Thus the aggregation function is restricted by:

Assumption A2 Let $Args$ be any set of supporting arguments concerning St, and $Args'$ be any set of supporting and confirming arguments concerning St, then:

$$|\{(St : G : ++)\}| > |Args|$$
$$|\{(St : G : ++)\}| = |Args'|$$

as well as A1. This assumption is, of course, consistent with many quantitative calculi including probability and belief functions.

[1] The non-strict inequality allows for limits to the force of a set of arguments, as is the case when using the bounded dictionary introduced below.

Delta dictionary The dictionaries discussed so far have had signs which represent belief values. At times we may also wish to reason about changes in these values. Doing this it is natural to consider both increases and decreases in value, and so the simplest delta dictionary which we make use of is:

$$\mathcal{D}_{\text{delta}} =_{\text{def}} \{+, -\}$$

in which sentences $(St : G : +)$ and $(St : G : -)$ can be interpreted as indicating, respectively, an increase or decrease in confidence in the proposition St, without indicating the degree of the increase or decrease. The use of these signs is similar to their use in qualitative probabilistic networks [44], and qualitative certainty networks [30]. We call arguments with sign $-$ *opposing* arguments.

It is possible to justify a number of aggregation procedures for arguments which use the delta dictionary. Some of these honour A1[2], and it makes sense for such procedures to also make the following assumption:

Assumption A3 Let $Args$ be any set of arguments concerning St, then:

$$|Args \cup \{(St : G : -)\}| \leq |Args|$$

At times when using the delta dictionary, the following rules of inference may also be used:

$$(St : G : -) \Leftrightarrow (\neg St : G : +) \quad (1)$$

$$(St : G : +) \Leftrightarrow (\neg St : G : -) \quad (2)$$

where $\neg St$ is the negation of St. The first of these is read as "if you have $(St : G : -)$ you may infer $\neg(St : G : +)$ and if you have $\neg(St : G : +)$ you may infer $(St : G : -)$". Using this rule means that if we have a negative argument (for instance $(cancer : young : -)$, "the young age of the patient argues against her having cancer") then this increases our overall confidence in the negated conclusion. The second rule is analogous, and together they take account of the fact that there is no rule in $\vdash_{ACR}$ for handling negation. Taken together, the rules are akin to the rule of the excluded middle, and this explains why they are not included in $\vdash_{ACR}$. We don't include them since we want to be able to build systems whose signs do not use the rule of the excluded middle.

Bounded delta dictionary We can also extend the delta calculus with symbols which denote increases to a maximum and decreases to a minimum value:

$$\mathcal{D}_{\text{bounded delta}} =_{\text{def}} \{++, +, -, --\}$$

We call arguments with sign $--$ *excluding* arguments. As with the bounded generic dictionary, the fact that the dictionary is bounded suggests that any flattening function should operate under the assumption:

[2] Note that this involves overloading the assumption by making it apply to arguments about value and arguments about changes in value. However, this seems reasonable since exactly what kind of sign is being used is always clear from the context.

Assumption A4 Let $Args$ be any set of supporting or opposing arguments concerning St, and $Args'$ be any set of supporting, opposing and excluding arguments concerning St, then:

$$|\{(St : G : --)\}| < |Args|$$
$$|\{(St : G : --)\}| = |Args'|$$

Under this assumption, it is inconsistent to have both $(St : G : ++)$ and $(St : G' : --)$ for any St. Furthermore, when using the bounded delta dictionary, if (1) and (2) hold, then so do the following:

$$(St : G : ++) \Leftrightarrow (\neg St : G : --) \quad (3)$$
$$(St : G : --) \Leftrightarrow (\neg St : G : ++) \quad (4)$$

This completes the description of the four simplest dictionaries.

The reason that the generic dictionary is called "generic" is that it can be viewed as an abstraction of a number of quantitative uncertainty handling formalisms. Thus the "+" in the dictionary can be viewed, for instance, as either a probability, possibility or belief value, but one which is not precisely specified. A similar interpretation may be used for the delta dictionary; we can look at the "−" used there as a decrease in probability, possibility or belief without saying how much of a decrease it is. Clearly there is a limit to what can be done without identifying what kind of value is being manipulated, since the theories from which the values are taken will place some constraints on which assumptions may be valid and under what conditions they are valid.

For example, if we give a probabilistic semantics to the delta dictionary so that + represents an increase in probability [27, 28], then we get a system of argumentation which is similar in many ways to qualitative probabilistic networks [44]. With this probabilistic semantics, (1)–(4) are valid, and it is possible to determine the precise conditions under which the simple aggregation procedure of adding up arguments is reasonable [28]. Furthermore, there are delta forms for any quantitative uncertainty representation [30], and it is straightforward to show that in some of these, most notably when the signs are given a semantics in terms of possibility theory or belief functions, (1) and (4) are not valid.

As noted above, we will typically 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. We can describe this in terms of a function $\mathsf{Flat}^{\mathsf{A}}(\cdot)$ which maps from a set of arguments $\mathbf{A}$ for a proposition St from a particular database Δ to the pair of that proposition and some overall measure of validity:

$$\mathsf{Flat}^{\mathsf{A}} : \mathbf{A} \mapsto \langle St, v \rangle$$

where $\mathbf{A}$ is the set of all arguments which are concerned with St, that is:

$$\mathbf{A} = \{(St : G_i : Sg_i) \mid \Delta \vdash_{ACR} (St : G_i : Sg_i)\}$$

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ intro}}$	++	+
++	+	+
+	+	+

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$	++	+
++	++	+
+	+	+

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}$	++	+
++	++	+
+		+

Fig. 3. Combinator tables for LA using the bounded generic dictionary.

and v is the result of a suitable combination of the Sg that takes into account the structure of the arguments. Thus v is the result of applying a flattening function to the grounds and signs of all the arguments in $\mathbf{A}$:

$$v = \mathsf{flat}^{\mathsf{A}}(\{\langle G_i, Sg_i\rangle \mid (St : G_i : Sg_i) \in \mathbf{A}\})$$

Often the signs Sg_i and the overall validity v will be drawn from the same dictionary, but it is perfectly feasible for them to be drawn from different dictionaries (so, for example, a set of arguments with numerical weights may be flattened to give a degree of support drawn from the dictionary $\{high, medium, low\}$).

Thus, if we have a set of arguments $\mathbf{A}$ for a proposition St, then the result of flattening is:

$$\mathsf{Flat}^{\mathsf{A}}(\mathbf{A}) = \langle St, \mathsf{flat}^{\mathsf{A}}\Big(\{\langle G_i, Sg_i\rangle \mid (St : G_i : Sg_i) \in \mathbf{A}\}\Big)\rangle$$

Together $\mathcal{L}$, the rules for building the formulae, the connectives, and $\vdash_{ACR}$ define a formal system of argumentation LA[3]. In fact, LA is really the basis of a family of systems of argumentation, because one can define a number of variants of LA by using different meanings for the connectives, different dictionaries of signs, different meanings for the dictionaries, different functions for combining signs $\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ elim}}$, and implication $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}$ and $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$, and different means of flattening arguments, $\mathsf{flat}^{\mathsf{A}}$. Given the number of possible choices, it is possible to define a bewildering variety of different versions of LA[4]. We now describe a couple of the best understood.

The way we go about defining a new version of LA is to decide three things. First, which dictionary to use. Second, how the signs within that dictionary are to be interpreted. Third, how the connectives are to be interpreted. It should be stressed that these choices are separate; it is possible to use the same dictionary with different meanings for the signs and with different meanings for the connectives. Once the choices are made it is possible to identify how to combine the signs correctly, and to identify which additional rules of inference (such as (1)) hold. Then it is possible to determine how to flatten arguments, and under what conditions the various assumptions about flattening are reasonable.

We start by considering the use of the bounded generic dictionary in which the signs are interpreted using probability theory. In particular, we take + to

[3] The name stands for Logic of Argument [17].

[4] And we can complicate the picture further by defining other systems of argumentation which use different underlying logics, and so have different consequence relations $\vdash_{ACR}$. An example of such a system may be found in [26]

denote a probability of some unknown value, and ++ to denote certainty (a probability of 1). We take $\wedge$ to be logical conjunction, and $\rightarrow$ to be material implication. With this interpretation, the combination functions required by LA are those of Figure 3. These require a little explanation. The table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ intro}}$ gives the sign of the sentence $St \wedge St'$ from the signs of the sentences St and St'. Thus, whatever the signs of St and St', the sign of $St \wedge St'$ is +. The table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$ gives the sign of the sentence St' from the signs of the sentences St and $St \rightarrow St'$. Thus if both the signs of St and $St \rightarrow St'$ are ++ so is that of St', and otherwise the sign of St' is +. The table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}$ follows directly from that for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$ since it gives the sign of $St \rightarrow St'$ from that of St and St'. In the table, St is the value in the leftmost column and St' is the value in the top row. This time the table includes a space, since it is impossible for St' to have sign ++ when St has sign +. Furthermore, it should be noted that when St and St' both have sign +, then the sign of $St \rightarrow St'$ could be either + or ++. Since + includes ++ (since a probability of 1 is also a probability of some value), we give the result as +. This forgiving nature of the signs allows $\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ elim}}$ to be stated as follows:

$$\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ elim}}(Sg) = Sg.$$

It is straightforward to verify that these functions are correct for this interpretation of the signs.

Since these are the only rules of inference we need to consider, we can then proceed to identifying aggregation procedures. Two obvious ones spring to mind. In the first, the function $\mathsf{flat}^{\mathsf{A}}$ examines the signs Sg_i and returns ++ if any of the Sg_i is ++, and otherwise returns +. Thus, formally:

$$v = \begin{cases} ++ \text{ if } Sg_i = ++ \text{ for some } i \\ \;\,+ \;\;\text{otherwise} \end{cases}$$

This flattening function conforms to assumptions A1 and A2 while making no additional assumptions about the strength of arguments which are not implicit in the meaning of the signs. Note that this flattening function, in common with the others detailed in this paper, ignores the grounds. This is possible because of the use of qualitative dictionaries—at this coarse level of granularity, the interactions between arguments captured by the grounds can be ignored. However, when quantitative dictionaries are used, the grounds play an important part in flattening.

The system of argumentation described here, with the combination functions of Figure 3 and the flattening function described above, is basically that discussed in [17], though in the latter paper the meaning of the signs is less explicit than here, and the presentation is slightly different.

The second obvious aggregation procedure is slightly more complex. In this function, the Sg_i come from $\mathcal{D}_{\text{bounded generic}}$, while v is just a positive number which we can consider coming from the dictionary:

$$\mathcal{D}_{\text{aggregation}} =_{\text{def}} \{0, 1, 2, \ldots\}$$

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$	++	+	−	−−
++	++	+	−	−−
+	+	+	−	−
−	−	−	+	+
−−	−	−	+	+

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ intro}}$	++	+	−	−−
++	++	?	?	−−
+	?	?	?	−−
−	?	?	?	−−
−−	−−	−−	−−	−−

$\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}$	++	+	−	−−
++	++	+	−	−−
+		+	−	
−		−	+	
−−		−	+	

Fig. 4. Combinator tables for LA using the bounded delta dictionary.

All the procedure does is to count the number of arguments, again taking into account the fact that once one has one argument with sign ++ in favour of a sentence, all other arguments are irrelevant.

$$v = \begin{cases} \infty & \text{if } Sg_i = {+}{+} \text{ for some } i \\ |\mathbf{A}| & \text{otherwise} \end{cases}$$

where $|X|$ gives the cardinality of the set X. This flattening function also conforms to assumptions A1 and A2 but in addition assumes that all arguments with sign + have equal strength[5]. Using the second aggregation function we get a version of LA which is essentially that used in the system Pro*forma* described in Section 5.

The other system we consider uses the bounded delta dictionary in which, for a formula which does not contain an implication, the sign + denotes an increase in probability, − denotes a decrease in probability, ++ denotes an increase in probability to 1 and −− denotes a decrease in probability to 0. We also have to consider what an implication means in this system, and we take a sign of ++ for $St \rightarrow St'$ to mean that if the probability of St increases to 1 so does that of St'. We also take a sign of + for $St \rightarrow St'$ to mean that if the probability of St increases so does the probability of St', a sign of − for $St \rightarrow St'$ to mean that if the probability of St increases the probability of St' decreases, and a sign of −− for $St \rightarrow St'$ to mean that if the probability of St increases to 1 the probability of St' decreases to zero.

With this semantics, the combinator tables are those in Figure 4, and these can, once again, easily be proved to be correct for changes in probability [25, 28]. There are a couple of things that should be noted. First, the table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{conj\ intro}}$ introduces the sign ? to stand for "++ or + or − or −−". This is a usual feature of qualitative systems—when you deal with abstractions, you find that eventually you need new composite abstractions because it becomes unclear

[5] This additional assumption is taken to be reasonable when there is no knowledge about the comparative strength of arguments.

$\mathsf{flat}^{\mathsf{A}}$	$++$	$+$	$-$	$--$
$++$	$++$	$++$	$++$	
$+$	$++$	$+$	?	$--$
$-$	$++$	?	$-$	$--$
$--$		$--$	$--$	$--$

Fig. 5. Flattening function for LA using the bounded delta dictionary.

which abstraction is the right one. Second, the table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ elim}}$ should be read with the sign of the antecedent being picked from the leftmost column and the sign of the implication being picked from the top column, and the table for $\mathsf{comb}^{\mathsf{A}}_{\mathsf{imp\ intro}}$ should be read with the sign of the antecedent being picked from the leftmost column and the sign of the consequent being picked from the top row. Third, that the spaces in the latter table reflect impossible situations, and fourth that the sign given in this table is always the least specific possible, so when the implication could have sign $+$ or $++$, the table gives $+$.

As before, there are a number of different ways in which one can flatten arguments. One possible flattening function is one which conforms to all the assumptions introduced so far, but makes no additional assumptions. This gives the table of Figure 5, and once again this can be shown to be correct for probability theory. Here, as with the remainder of the systems discussed in this paper we define the flattening function to be binary—to generate v we apply it recursively.

With these combination and flattening functions, the system we have described is essentially the system $\mathcal{NA}''$ described in [28], and similar systems which uses the delta dictionary are $\mathcal{NA}_1$ and $\mathcal{NA}_2$ in [27]. The notation used by these three systems is slightly different from that presented here because the overloading of $++$, $+$, $-$ and $--$ is overcome by the use of additional symbols to represent changes in probability.

2.4 Argumentation and defeasibility

The main focus of this paper is on reasoning under uncertainty in the context of making decisions. However, it is worth making a few remarks about the ways in which LA may be related to systems such as default logic and standard modal logic.

Default logic Suppose we can construct an argument for St on the basis of a default rule. By definition a default is not guaranteed to be correct, so in this calculus the argument has the form:

$$(St : \mathit{default} : +)$$

If we later identify reasons to reject St, because we obtain an argument:

$$(\neg St : G : ++)$$

then aggregation will yield the conclusion $\neg St$ by A4 and (4). Argumentation therefore permits behaviour much like that of standard default logic, but it may also illuminate the relationship between default reasoning and quantitative uncertainty. Suppose we have a reason to doubt St but not to reject it (because we can construct the argument $(\neg St : G : +)$) then, using the flattening function that counts arguments, this balances the default argument, and we are equivocal about whether St or $\neg St$. If we have further arguments against St then the balance of argument turns against it (3) but we can still hold both St and $\neg St$ as possibilities, a behaviour similar to the normal behaviour of probabilistic, possibilistic, belief function and other quantitative calculi.

Modal logic The bounded delta calculus may also accommodate ideas akin to those of of modal logic. Informally, $possible(St)$ holds if we can construct an argument for St, and $necessary(St)$ if we can construct a bounding argument for St:

$$possible(St) \Leftrightarrow (St : G : +)$$
$$necessary(St) \Leftrightarrow (St : G : ++)$$

Suppose we have an argument:

$$(St : G : +)$$

which means that $possible(St)$ holds. Then, if we introduce an additional argument:

$$(\neg St : G : ++)$$

which means that $necessary(\neg St)$ holds, and if we aggregate these arguments constraints, A4 and (4) entail that $necessary(\neg St)$ dominates $possible(St)$. Turning this around if we hold $possible(St)$ then we cannot hold $necessary(\neg St)$, therefore:

$$possible(St) \Leftrightarrow \neg necessary(\neg St)$$

An analogous argument can be followed for the dual rule of modal logic:

$$necessary(St) \Leftrightarrow \neg possible(\neg St)$$

Equating modality to provability in this way echoes work on classifying arguments on the basis of the the arguments which my be built against their grounds [10, 11].

2.5 Soundness and completeness

So far we have neglected to say much about what it means to have an argument for a proposition beyond the fact that an argument is a *tentative proof* of the proposition and so is a proof which can fail if suitably strong arguments against the proposition can be found. However, as with any formal model of reasoning, what we would like to do is to prove that argumentation is in some sense correct, that is it generates all and only correct inferences. In other words, we would like to show that argumentation is complete and sound. To do this, however, we need to say precisely what an argument is. There are a number of ways of doing this, and three different approaches have been taken.

The first approach was based upon the commonalities between argumentation as introduced here and intuitionistic logic first pointed out by Ambler [3]. The idea was that since it is possible to give intuitionistic logic a proof-theoretic semantics in terms of category theory, this should also be possible for argumentation. Indeed this turned out to be the case. The first steps in providing this semantics are detailed in [3] which identifies the structure of the space of arguments, along with the kind of operations possible over them. The rest of the formalisation is provided in [2], which also highlights the link between argumentation and Dempster-Shafer theory [37].

The second approach was to give argumentation a model theoretic semantics. In particular, standard Kripke semantics for modal logic have been adapted by Das [7] to give a possible worlds interpretation for what it means for an argument to support a proposition to some degree.

The final semantics developed so far [27, 28] relates certain types of argumentation to probability theory by taking an argument in favour of a proposition to mean that there is evidence that the probability of the proposition increases (so the proposition becomes more likely to be true). With this interpretation, and using the bounded delta dictionary, it is possible to show that argumentation is sound and complete. Thus argumentation can capture probabilistic reasoning if required, and so it is possible to claim that, under particular conditions, argumentation is a normative theory for handling uncertainty. The probabilistic semantics has another advantage. Because it ties the notion of an argument securely to well-understood ideas about qualitative probability, it is possible to harness a number of useful results concerning qualitative probability [25, 29]. In particular, it is possible to develop a finer-grained representation of what it means to have an argument for a proposition which allows arguments of different strengths to be accommodated [27].

3 Towards arguments about actions

Having described the logic of argumentation LA for reasoning with uncertain information, we now consider some steps towards extending it to deal with actions in order to build a more complete decision theory. As in the previous section we begin with an informal discussion of the kinds of things we are trying to achieve.

3.1 An overview

At an informal level there appears to be a clear isomorphism between arguments for beliefs and arguments for actions. Suppose we wish to construct an argument in favour of treating a patient with cancer by means of chemotherapy. This might run as follows:

> Cancer is an intolerable condition and should be eradicated if it occurs. It is a disease consisting of uncontrolled cell proliferation. Certain chemical agents kill cancer cells and/or reduce proliferation. Therefore we should treat cancer patients with such agents.

The steps in this argument are *warranted*[6] by some generalised (and probably complex) *theory* of the pathophysiological processes involved in cancer, and a *value system* which defines what kinds of things are tolerable, desirable and so on. The argument is not conclusive, however, since the conclusion might be rebutted by counter-arguments, as when chemotherapy is contra-indicated if a patient is frail or pregnant.

Such arguments appear compatible with LA and consequently we might consider using LA to construct such arguments. Suppose we summarise the above example in the notation of LA:

$$(St : G : +)$$

where St is the sentence "the patient should be treated with chemotherapy", G denotes the grounds of the argument (the sequence of steps given), and $+$ indicates that the grounds support action St. However this conceals some significant complexities. The notion of "support" seems somewhat different from the interpretation we have previously assigned to it. For LA we have adopted the interpretation that an argument is a conventional proof, albeit one which it is acknowledged cannot in practice be guaranteed to be correct. An argument in support of some proposition is, in other words, a proof of the proposition which we accept could be wrong. This analysis of "support" does not seem to be entirely satisfactory when reasoning about what we *ought to do* as opposed to what *is the case*. Consider the following simple argument, which is embedded in the above example:

> cancer is an intolerable condition, therefore it should be eradicated

There is a possibility that this argument is mistaken, which would justify signing it with $+$ (a "supporting" argument in LA) but the sense of support seems to be different from that which is intended when we say that the intolerable character of cancer gives support to any action that will eradicate it. In other words when we say "these symptoms support a diagnosis of cancer", and "these conditions support use of chemotherapy" we are using the term "supports" in quite distinct ways. The latter case involves no uncertainty, but depends only upon some sort of statement that intolerable states of affairs ought not to be allowed to continue. If this is correct then it implies that arguing from "value axioms" is not the same thing as arguing under uncertainty and so is it inappropriate to use LA for constructing such arguments.

3.2 The logics of value LV and expected value LEV

How might we accommodate arguments about value within our existing framework? One possibility might be to keep the standard form and elaborate the

[6] The terminology harking back to Toulmin.

The patient has colonic polyps	$(cp : G1 : {++})$	$e1$
polyps may lead to cancer	$(cp \rightarrow ca : G2 : +)$	$e2$
cancer may lead to loss of life	$(ca \rightarrow ll : G3 : +)$	$e3$
loss of life is intolerable	$(\neg ll : av : {++})$	$v1$
surgery preempts malignancy	$(su \rightarrow \neg(cp \rightarrow ca) : G4 : {++})$	$e4$
argument for surgery	$(su : (e1, e2, e3, e4, v1) : +)$	$ev1$
surgery has side-effect se	$(su \rightarrow se : G5 : {++})$	$e5$
$\neg se$ is desirable	$(\neg se : av : +)$	$v2$
argument against surgery	$(\neg su : (e5, v2) : +)$	$ev2$
se is preferable to loss of life	$(pref(se, ll) : (v1, v2) : {++})$	$p1$
no arguments to veto surgery	$(safe(su) : cir : {++})$	$c1$
surgery is preferable to $\neg$ surgery	$(pref(su, \neg su) : (ev1, ev2, p1) : {++})$	$p2$
commit to surgery	$(do(su) : (p2, c1) : {++})$	$a1$

Fig. 6. An example argument

sentence we are arguing about to include a "value coefficient":

$$((St : +) : G : +)$$

Which might be glossed as "there is reason to believe that action St will have a positively valued outcome". This may allow us to take advantage of standard LA for reasoning with sentences about the value of actions, but it does not, of course, solve our problem since it says nothing about the way in which we should assign or manipulate the value coefficients.

As a result, we currently prefer another approach, which is analogous to the decision theoretic notion of expected value. In this approach we construct compound arguments based on distinct steps of constructing and combining belief arguments and value arguments. For example, consider the following argument:

Doing A will lead to the condition C	$(A \rightarrow C : G : +)$
C has positive value	$(C : G' : +)$
Doing A has positive expected value	$(A : G \cup G' : +)$

We can think of this as being composed of three completely separate stages as well as having three steps. The first stage is an argument in LA that C will occur if action A is taken, which could be glossed as "G is grounds for arguing in support of C resulting from action A". The second stage says nothing about uncertainty; it simply requires some mechanism for assigning a value to C, call this LV[7]. The final stage concludes that A has positive expected value; to make this step we shall have to give some mechanism for deriving arguments over sentences in LA and LV, call this LEV[8].

The attraction of this scheme is that it appears to make explicit some inferences which are hidden in the other argument forms. However, it has the additional requirements that we define two new systems—LV and LEV. It seems

[7] The name stands for Logic of Value.

[8] The name stands for Logic of Expected Value.

to us that this is a price worth paying since making the assignment of values and the calculation of expected value explicit gives much more flexibility and so makes it possible to represent quite complex patterns of reasoning. As an example of the kind of reasoning that should be possible consider the following:

> (1) The patient is believed to have colonic polyps which, while presently benign, could become cancerous.
> (2) Since cancer is life-threatening we ought to take some action to pre-empt this threat.
> (3) Surgical excision is an effective procedure for removing polyps and therefore this is an argument for carrying out surgery.
> (4) Although surgery is unpleasant and has significant morbidity this is preferable to loss of life, so surgery ought to be carried out.

Informally we can represent this argument as in Figure 6.

There are six different forms of argument in this example which has a similar scope to the examples considered by Tan and Pearl [40]. The first are those labelled $e1$–$e5$ which are standard arguments in LA. The second are value assignments $v1$ and $v2$ which represent information about what states are desirable and undesirable. The third are expected value arguments $ev1$ and $ev2$ which combine the information in standard and value arguments. The fourth are arguments $p1$ and $p2$ which express preferences between different decision options. The fifth type of argument is the closure argument $c1$ which explicitly states that all possible arguments have been considered, and this leads to the final type of argument, the commitment argument $a1$ which explicitly records the taking of the decision. The following sections discuss some features of these arguments, in particular values and expected values.

4 Systems of argumentation for dealing with values and expected values

Having discussed in general terms what is required from LV and LEV, we can start moving towards an initial formal definition. We require some language for representing values. Notwithstanding the common-sense simplicity of the idea of value its formalisation is not likely to be easy. Value assignments are commonly held to be fundamentally subjective—they are based on the preferences of a decision maker rather than being grounded in some observable state of affairs.

4.1 Arguments about values

There are a number of possible formalisms we might consider. We might, for instance, adopt some set of modal operators, such as $desirable(St)$, where St is some sentence such as "the patient is free of disease". This is the approach adopted by Bell and Huang [4]. Alternatively we might attach numerical coefficients, as in the use of quantitative utilities in traditional decision theory. We propose representing the value of a state or condition St by labelling a proposition describing St with a sign drawn from some dictionary $\mathcal{D}$ just as we do for beliefs. In this discussion we shall only consider qualitative value dictionaries because, as with uncertainty, we can invariably judge whether some state has positive or negative value, or is valueless, though we may not be able to determine a precise point value or precise upper and lower bounds on the value.

Another similarity with our view of uncertainty is that we can frequently assign different values to states from different points of view. For example the use of opiates is bad since they lead to addiction, but good if they are being used as an analgesic. We therefore propose to label value assignment expressions with the grounds for the assignment, for instance $St : G : V$, giving us a "value argument" analogous to the argument expressions of LA. This is not a new idea of course. For example, multi-attribute utility theory also assumes the possibility of multiple dimensions over which values can be assigned. However, the benefits of this sort of formalisation is that it may allow us to cope with situations where we cannot precisely quantify the value of a situation, and it permits explicit representation of the justifications for particular value assignments making it possible to take them into account when reasoning. The basic schema of value assignment is analogous to the standard argumentation schema:

$$\text{Database} \vdash_{VCR} (\text{Condition} : \text{Grounds} : \text{Value}) \tag{5}$$

A Basic Value Argument (BVA) is a triple defining some state, the value assigned to it, and a justification for this particular assignment. The assertions "health is good" might be represented in grounds-labelled form by:

$$(health : va : +)$$

where va is a label representing the justification for the BVA.

Traditionally there has been considerable discussion of the justifications for value assignments. Any discussion has to face the difficulty that values seem to be fundamentally subjective. In discussion of beliefs there is an analogous idea of subjective probability but it is also possible to invoke the idea of long-run frequency to provide an objective basis for probability theory. There has been a similar attempt to identify an objective framework for values, in consensual values (for example social mores and legal systems), but it seems inescapable that values are grounded in opinion rather than some sort of objective estimation analogous to the chances of events. We therefore accept that a value assignment may in the end be warranted by sentences like "because I say so", "because the law says so", and "because the church says so".

In other words we have nothing new to say about the nature of the "value theories" invoked in (5). We shall simply assume that the theory provides a set of basic value assignments. Our task here is not to give or justify any particular set of value assignment sentences (any more than probability theorists are required to provide particular collections of prior or conditional probabilities) but to identify ways in which collections of such value sentences might be manipulated, aiming to take some steps towards the definition of a system LV which is

$$\text{Ax}\frac{(St : G : Sg) \in \Delta}{\Delta \vdash_{VCR} (St : G : Sg)}$$

$$\wedge\text{-I}\frac{\Delta \vdash_{VCR} (St : G : Sg) \ \Delta \vdash_{VCR} (St' : G' : Sg')}{\Delta \vdash_{VCR} (St \wedge St' : G \cup G' : \mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ intro}}(Sg, Sg'))}$$

$$\wedge\text{-E1}\frac{\Delta \vdash_{VCR} (St \wedge St' : G : Sg)}{\Delta \vdash_{VCR} (St : G : \mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}(Sg))}$$

$$\wedge\text{-E2}\frac{\Delta \vdash_{VCR} (St \wedge St' : G : Sg)}{\Delta \vdash_{VCR} (St' : G : \mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}(Sg))}$$

Fig. 7. Value Consequence Relation

analogous to LA but deals with values rather than beliefs. The assumption is that the assignment of values in sentences like "health is good" depends upon a derivation which bottoms out in some set of BVAs and that these will be propagated in the grounds of the relevant arguments.

4.2 Formalising argumentation about values

Having spoken informally about what LV is attempting to do, we present an initial attempt at formalizing it. This, as the observant reader will notice, is virtually identical to the definition of LA. We start with another set of atomic propositions $\mathcal{M}$ including $\top$ and $\bot$, the ever true and ever false propositions. We also have a set of connectives $\{\neg, \wedge\}$, and the following set of rules for building the well-formed formulae (*wff* s) of the language.

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

Note that currently LV does not make use of the connective $\rightarrow$ since it is unclear to us what such a connective might mean. However, Shoham's recent work [38] suggests that some way of expressing conditional values may well be necessary. The set of all *wff* s that may be defined using $\mathcal{M}$, may then be used to build up a database Δ where every item $d \in \Delta$ is a triple $(St : G : Sg)$ in which St is a *wff*, Sg represents the value of St, and G are the grounds on which the assertion is made. With this formal system, we can take a database and use the argumentation consequence relation $\vdash_{VCR}$ defined in Figure 7 to build arguments for propositions that we are interested in. Given the explanation of $\vdash_{ACR}$ the way this works should be clear.

Now, as before, we define dictionary $\mathcal{D}$ by:

$$\mathcal{D} =_{\text{def}} \{S_1, \ldots S_n\}$$

and so we may write:

$$(St : G : S_i)$$

where S_i is any symbol drawn from $\mathcal{D}$. For values there are a couple of obvious dictionary. The first is that of numerical value, measured in whatever currency one chooses, another is that of utiles—the familiar measure of classical decision theory:

$$\mathcal{D}_{\text{money}} =_{\text{def}} \{S : S \in (-\infty, \infty)\}$$
$$\mathcal{D}_{\text{utility}} =_{\text{def}} \{S : S \in (-\infty, \infty)\}$$

However, as was the case with beliefs, our interest is primarily with qualitative dictionaries, so it is worth considering in more detail two value dictionaries which are analogous to the simple qualitative dictionaries we considered for use with LA.

Cost benefit dictionary The simplest useful dictionary of values allows us to talk about states that are good or desirable and states which are bad or undesirable.

$$\mathcal{D}_{\text{cost benefit}} =_{\text{def}} \{+, -\}$$

As with beliefs there are two ways we could interpret these signs. We could take + to mean simply that the state has some absolute (point) positive value, but that the precise value is unknown, or we could take it to mean that we have an argument for the overall value of our goods being increased. For the moment we restrict ourselves to using absolute values, but delta values for values may be required at a later date. It would seem that good and bad states can be related through complementation rules:

$$(St : G : +) \Leftrightarrow (\neg St : G : -) \tag{6}$$
$$(St : G : -) \Leftrightarrow (\neg St : G : +) \tag{7}$$

analogous to (1) and (2) above.

Bounded cost benefit dictionary There also seems to be some benefit in extending the cost benefit dictionary to allow us to talk about maximal amounts of goodness (badness):

$$\mathcal{D}_{\text{bounded cost benefit}} =_{\text{def}} \{++, +, -, --\}$$

However, there seems to be a complication here. It seems straightforward to claim that there is a lower bound on badness—we might gloss this by saying certain conditions are "intolerable" such as death for instance—but an upper bound on "goodness" (for example of a bank balance) is harder to conceive of. However if we accept:

$$(St : G : ++) \Leftrightarrow (\neg St : G : --) \tag{8}$$
$$(St : G : --) \Leftrightarrow (\neg St : G : ++) \tag{9}$$

by analogy with (3) and (4), then we can obtain a reasonable interpretation for the idea of a condition which is maximally desirable as the complement of any condition that is intolerable. Furthermore sentences like "human life is priceless" are held, by their users at least, to have some meaning. From a pragmatic point of view such statements can seem merely romantic, but if we accept the above rules it is a direct consequence of asserting that loss of life is intolerable.

Since values are derived with respect to some value theory we can contemplate different value arguments for the same sentence. In common with LA, such value arguments can be aggregated. We can describe this aggregation, as for LA, in terms of a function $\mathsf{Flat}^{\mathsf{V}}(\cdot)$ which maps from a set of value arguments $\mathbf{A}$ for a proposition St from a particular database Δ to the pair of that proposition and some overall measure of validity:

$$\mathsf{Flat}^{\mathsf{V}} : \mathbf{A} \mapsto \langle St, v \rangle$$

where $\mathbf{A}$ is the set of all arguments which are concerned with St, that is:

$$\mathbf{A} = \{(St : G_i : Sg_i) \mid \Delta \vdash_{VCR} (St : G_i : Sg_i)\}$$

and v is the result of a suitable combination of the Sg that takes into account the structure of the arguments, that is v is the result of applying a flattening function to the grounds and signs of all the arguments in $\mathbf{A}$:

$$v = \mathsf{flat}^{\mathsf{V}}\Big(\{\langle G_i, Sg_i\rangle \mid (St : G_i : Sg_i) \in \mathbf{A}\}\Big)$$

Often the signs Sg_i and the overall validity v will be drawn from the same dictionary, but it is perfectly feasible for them to be drawn from different dictionaries (so that a set of arguments with numerical values might be flattened to a value drawn from the dictionary $\{$*very expensive*, *expensive*, *cheap*$\}$).

There are, of course, a number of possible ways in which we might aggregate values. Numerical values might be aggregated by summation, for instance, and clearly the exact aggregation operation will depend upon the meaning of the value signs. One obvious assumption we might wish to make when using the cost benefit or bounded cost benefit dictionary is that:

Assumption A5 If $Args$ is any set of arguments supporting and opposing arguments, then:

$$|Args| \leq |Args \cup \{(S : G : +)\}|$$

Following previous usage we might refer to the set of arguments as the *case* for S being positively valued, and $|Args|$ as the *force* of these arguments. Now, a condition may be desirable on some grounds and undesirable on others, for instance if we have:

$$\Delta \vdash_{VCR} (St : G : +)$$
$$\Delta \vdash_{VCR} (St : G' : -)$$

This raises the question of how supporting and opposing arguments interact. One possibility is to make the flattening function obey the assumption:

Assumption A6 If $Args$ is any set of supporting and opposing arguments, then:

$$|Args| \geq |Args \cup \{(St : G : -)\}|$$

So that arguments with negative value bring the overall weight of a set of arguments down. In addition, we might want to assume that:

Assumption A7 If $Args$ is any set of supporting and opposing arguments, then:

$$|Args| = |Args \cup \{(St : G : -), (St : G' : +)\}|$$

so that positive and negative arguments cancel one another. This latter assumption is exactly the same as the one encoded in the flattening function for LA which counts the number of arguments. An alternative flattening, which is more in agreement with qualitative versions of classical decision theory [1, 44], is to have complementary value arguments lead to indeterminacy.

This picture is complicated slightly by the use of the bounded cost-benefit dictionary, where we have limits to values. Using this dictionary suggests the adoption of an additional assumption similar to A2 and A4:

Assumption A8 let $Args$ be any set of supporting and opposing arguments concerning St, and $Args'$ be any set of supporting, opposing and confirming arguments concerning St, then:

$$|\{(St : G : ++)\}| > |Args|$$
$$|\{(St : G : ++)\}| = |Args'|$$

let $Args''$ be any set of supporting and opposing arguments concerning St, and $Args'''$ be any set of supporting, opposing and excluding arguments concerning St, then:

$$|\{(St : G : --)\}| < |Args''|$$
$$|\{(St : G : --)\}| = |Args'''|$$

so that an argument with maximal strength is not affected by additional information. Of course, as with A4, this means that it is inconsistent to have both $(St : G : ++)$ and $(St : G' : --)$ for any St.

Having discussed things in abstract terms, let's make things concrete by discussing one possible semantics for the bounded cost-benefit dictionary. In particular, we take $+$ to be some unknown positive value, and $++$ denotes some limiting unknown value (but not infinity[9]). Similarly, $-$ is some unknown negative value, and $--$ is a limiting negative value. If it helps, these can be taken to be qualitative abstractions of monetary value, with $+$ being any credit, $-$ any debit and $++$ the amount of money which if one had it, one would no longer have to worry about working for a living. We take $\wedge$ to be logical conjunction. With this interpretation, the combination function $\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ intro}}$ is that of Figure 8. This again uses ? as an abbreviation for "one of $++$, $+$, $-$ and $--$" (though it would probably suffice to make it just an abbreviation for "$+$ or $-$"). The

[9] We could use infinity, but that would make the combinator tables slightly different.

$\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ intro}}$	$++$	$+$	$-$	$--$
$++$	$++$	$+$	$+$	$?$
$+$	$+$	$+$	$?$	$-$
$-$	$+$	$?$	$-$	$-$
$--$	$?$	$-$	$-$	$--$

Fig. 8. The combinator table for LV using the bounded cost benefit dictionary.

$\mathsf{flat}^{\mathsf{V}}$	$++$	$+$	$-$	$--$
$++$	$++$	$+$	$+$	$\mathcal{C}$
$+$	$+$	$+$	$?$	$-$
$-$	$+$	$?$	$-$	$-$
$--$	$\mathcal{C}$	$-$	$-$	$--$

Fig. 9. The flattening function for LV using the bounded cost benefit dictionary.

function $\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}$ for this interpretation is:

$$\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}(Sg) = \begin{cases} Sg \text{ if } Sg \in \{++, --\} \\ ? \quad \text{otherwise} \end{cases}$$

It is straightforward to verify that these functions are correct for this interpretation of the signs, and the interested reader is encouraged to do so.

We also need to define a function to flatten value arguments with this interpretation. The binary version of this flattening function is that of Figure 9. This is very similar to the table for $\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}$, but differs in that it introduces another new symbol, $\mathcal{C}$. This symbol represents a contradiction, and is at the heart of the difference between the flattening function and $\mathsf{comb}^{\mathsf{V}}_{\mathsf{conj\ elim}}$. If we have two arguments $(St : G : ++)$ and $(St' : G' : --)$ then we can build an argument for $St \wedge St'$. This represents a state of affairs has one component which is completely desirable and another which is completely undesirable, and it seems reasonable to give it a value which is, roughly speaking, the sum of $++$ and $--$, and is therefore somewhere in between. We therefore use the value ?. The intuition here is that the conjunction of a sentence with maximum positive value and one with maximum negative value has some intermediate value. Thus I may find it completely desirable to not have to work, and completely undesirable to have no income but I can put some value on the state in which I don't work and have no income.

However, if St and St' are the same sentence, then the two arguments contradict each other—they say that St is both completely desirable and completely undesirable—no overall value for St can be agreed. The intuition here is that we cannot simply cancel an argument that a condition is absolutely desirable with an argument that the same condition is absolutely undesirable. For example, in discussions of euthanasia we may have an absolute prohibition on killing; this cannot simply be cancelled out by arguing that a loved one's pain is intolerable.

There are, of course, no simple decision rules for such situations and that is why we choose to flag the situation with $\mathcal{C}$ rather than reduce the conflict to some arbitrary value. What we need is to be able to recognise that a conflict has occurred, and then resolve it by means of some form of meta-logical reasoning, something like the opposite of circumscription, in which we introduce new assumptions or theories whose specific role is to overcome such deadlocks. In the euthanasia example, we may appeal to societal "thin end of the wedge" theories for instance in which "society's needs" were not included in the framing of the original decision.

4.3 Formalising argumentation about expected values

The previous section dealt with the problem of aggregation of value arguments. It remains to provide rules for deriving sentences from combinations of belief arguments and value arguments (that is arguments in LA and LV respectively). As an example of this kind of derivation, consider the following argument in LA:

$$(St : G : S)$$

meaning that we can argue for St with sign S. Assume further that we also have the following argument in LV:

$$(St : G' : V)$$

which means that the value of St is V. From these two arguments we wish to derive an expected value argument in LEV:

$$(St : G \cup G' : E)$$

meaning that the expected value of St is E. Now, from a decision making point of view, arguments about expected value of states are of little interest, except in the situation where they are the *outcomes* of actions that we can choose to take or not take. As an example of the kind of thing we would like to reason about, consider combining a sentence about belief with one about action:

$$\begin{array}{l} \neg cancer : v1 : V \\ \underline{surgery \Rightarrow \neg cancer : e1 : S} \\ surgery : v1 \cup e1 : E \end{array}$$

where $\Rightarrow$ is a connective which captures the notion of applying an action so that the sentence $surgery \Rightarrow \neg cancer$ is read "the action of *surgery* leads to the condition of $\neg cancer$" This pattern of reasoning is exactly the same as the previous one, combining a statement about beliefs (that surgery is a means of eradicating cancer, believed to degree S) with a statement about value (that a lack of cancer is a state with value V) to come up with a statement about expected value (that surgery in this case has some expected value E). However, to deal with this kind of reasoning we need to be able to talk about actions and to be able to reason backwards from the effects of actions to their causes. In particular,

$$\text{Ax}\frac{\Delta \vdash_{ACR} (St : G : Sg) \ \Delta' \vdash_{VCR} (St : G' : Sg')}{\Delta \cup \Delta' \vdash_{LEV} (St : G \cup G' : \mathsf{comb}^{\mathsf{L}}_{\mathsf{ax}}(Sg, Sg'))}$$

$$\wedge\text{-I}\frac{\Delta \vdash_{LEV} (St : G : Sg) \ \Delta \vdash_{LEV} (St' : G' : Sg')}{\Delta \vdash_{LEV} (St \wedge St' : G \cup G' : \mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ intro}}(Sg, Sg'))}$$

$$\wedge\text{-E1}\frac{\Delta \vdash_{LEV} (St \wedge St' : G : Sg)}{\Delta \vdash_{LEV} (St : G : \mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ elim}}(Sg))}$$

$$\wedge\text{-E2}\frac{\Delta \vdash_{LEV} (St \wedge St' : G : Sg)}{\Delta \vdash_{LEV} (St' : G : \mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ elim}}(Sg))}$$

Fig. 10. Expected Value Consequence Relation

from a formal point of view, we need to be able to handle the connective $\Rightarrow$. Given the well-known difficulties of building formal systems to reason about action, we will leave this for future work and only deal with combining values and beliefs about states.

With this simplification we can once again start the process of formalising the system of argumentation. We start with a third set of atomic propositions $\mathcal{N}$ including $\top$ and $\bot$, the ever true and ever false propositions. We also have the connectives $\{\neg, \wedge\}$, and the following set of rules for building the well-formed formulae (*wff* s) of the language.

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

The set of all *wff* s that may be defined using $\mathcal{N}$ then defines a legal set of triples $(St : G : Sg)$ in which St is a *wff* of LEV, Sg represents the expected value of St, and G are the grounds on which the assertion is made. However, LEV differs from LA and LV in that we don't build up a database of triples and build arguments from them, but build triples from existing arguments in LA and LV using the consequence relation $\vdash_{LEV}$ defined in Figure 10. Note that the consequence relation of Figure 10 differs from $\vdash_{ACR}$ and $\vdash_{VCR}$ in that the "bootstrap" rule, which allows the creation of an LEV argument from something other than an LEV argument, does not directly involve a tuple from some database. Instead it involves an argument in LA and an argument in LV. This captures the fact that any expected value argument is formed from a belief argument and a value argument.

With the consequence relation fixed, we can move on to identify suitable dictionaries. Now, our choice of dictionary is a little restricted since expected value arguments are based on both belief arguments and value arguments. Thus the meanings of the signs of expected value arguments are completely determined by the meanings of signs of their constituent belief and value arguments. Let us consider a suitable dictionary for expected value arguments built using belief

$\mathsf{comb}^{\mathsf{L}}_{\mathsf{ax}}$	++	+	−	−−
++	++	+	−	−−
+	+	+	−	−

$\mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ intro}}$	++	+	−	−−
++	++	+	+	?
+	+	+	?	−
−	+	?	−	−
−−	?	−	−	−−

Fig. 11. The combinator tables for LEV using the bounded expectation dictionary.

arguments whose signs are drawn from the bounded generic dictionary, and value arguments whose signs are drawn from the bounded cost benefit dictionary. This gives what we might call the bounded expectation dictionary:

$$\mathcal{D}_{\text{bounded expectation}} =_{\text{def}} \{++, +, -, --\}$$

If the belief dictionary in question is taken to be the probabilistic one discussed earlier, then the signs in the bounded expectation dictionary become qualitative abstractions of expectations (and hence the name). Suitable combinations functions are those of Figure 11. That for $\mathsf{comb}^{\mathsf{L}}_{\mathsf{ax}}$ reflects the multiplication of a belief (in the leftmost column) with a value (in the top row), while that for $\mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ intro}}$ is identical to the analogous function for LV. The function for eliminating conjunctions is:

$$\mathsf{comb}^{\mathsf{L}}_{\mathsf{conj\ elim}}(Sg) = \begin{cases} Sg \text{ if } Sg \in \{++, --\} \\ ? \quad \text{otherwise} \end{cases}$$

reflecting the indeterminacies in conjunction introduction. Once again, it is reasonably straightforward to show that these functions are correct.

In many cases a collection of qualitative expected value arguments can be aggregated under assumptions similar to those suggested for LV, and so we can again define a flattening function. As before we do this in terms of a function $\mathsf{Flat}^{\mathsf{L}}(\cdot)$ which maps from a set of expected value arguments $\mathbf{A}$ for a proposition St from a particular database Δ to the pair of that proposition and some overall measure of validity:

$$\mathsf{Flat}^{\mathsf{L}} : \mathbf{A} \mapsto \langle St, v \rangle$$

where $\mathbf{A}$ is the set of all arguments which are concerned with St, that is:

$$\mathbf{A} = \{(St : G_i : Sg_i) \mid \Delta \vdash_{LEV} (St : G_i : Sg_i)\}$$

and v is the result of a suitable combination of the Sg that takes into account the structure of the set of arguments, that is v is the result of applying a flattening function to the grounds and signs of all the arguments in $\mathbf{A}$:

$$v = \mathsf{flat}^{\mathsf{L}}\Big(\{\langle G_i, Sg_i \rangle \mid (St : G_i : Sg_i) \in \mathbf{A}\}\Big)$$

As ever, the signs Sg_i and the overall validity v can be drawn from the same dictionary or from different dictionaries. A flattening function suitable for the

flat^L	++	+	−	−−
++	++	+	+	$\mathcal{C}$
+	+	+	?	−
−	+	?	−	−
−−	$\mathcal{C}$	−	−	−−

Fig. 12. The flattening function for LEV using the bounded expectation dictionary.

bounded expectation dictionary is given in Figure 12. Once again, if we have expected value arguments based on conflicting values, for instance if we have $(St : G : ++)$ and $(\neg St : G' : ++)$ then such conflicts cannot be resolved within the system and as before are denoted $\mathcal{C}$.

The flattening function for LEV completes the definition of LV and LEV and we can turn to providing an example of their use.

4.4 Example

As an example of the kind of reasoning which LA, LV and LEV can capture, consider the following example adapted from [28]. The following database represents a career choice faced by the second author who needs to decide whether or not to concentrate his efforts on research or teaching:

$$\begin{array}{ll}
(concentrate_on_research : f1 : ++) & \Delta_1 \\
(concentrate_on_teaching : f2 : ++) & \\
(concentrate_on_research \rightarrow good_research : r1 : +) & \\
(good_research \rightarrow job_in_industry : r2 : ++) & \\
(concentrate_on_teaching \rightarrow good_tutor : r3 : ++) & \\
(good_tutor \rightarrow senior_university_job : r4 : ++) &
\end{array}$$

The facts $f1$ and $f2$ represent the possible choices, and the rules $r1$–$r4$ represent a subjective assessment of the relevant causal relations. These are expressed in LA using the bounded generic dictionary, using the probabilistic interpretation. From this information we can build the following arguments in LA by applying Ax and $\rightarrow$-E from Figure 2:

$$\begin{array}{l}
\Delta_1 \vdash_{ACR} (job_in_industry, \{f1, r2, r2\}, +) \\
\Delta_1 \vdash_{ACR} (senior_university_job, \{f2, r3, r4\}, ++)
\end{array}$$

which identify what the outcomes of the different career choices are, and how likely these are to come about; choosing to concentrate on teaching means a senior university job for sure, while concentrating on research means the chance of a job in industry. Now, consider we have the following value assignments in LV:

$$\begin{array}{ll}
(job_in_industry : f3 : +) & \Delta_2 \\
(senior_university_job : f4 : ++) &
\end{array}$$

which represent subjective assessments of the value of the possible outcomes expressed using the bounded cost benefit dictionary. From these we can build the arguments by applying Ax from Figure 7:

$$\Delta_2 \vdash_{VCR} (job\_in\_industry, \{f3\}, ++)$$
$$\Delta_2 \vdash_{VCR} (senior\_university\_job, \{f4\}, ++)$$

which tell us that both a senior university job and a job in industry are judged to be is totally desirable. The two related pairs of arguments can then be combined by applying Ax from Figure 10 :

$$\Delta_1 \cup \Delta_2 \vdash_{LEV} (job\_in\_industry, \{f1, f3, r1, r2\}, +)$$
$$\Delta_1 \cup \Delta_2 \vdash_{LEV} (senior\_university\_job, \{f2, f4, r3, r4\}, ++)$$

These values are expressed in the bounded expectation dictionary. From these arguments it is clear that the option to concentrate on teaching is the best since it will lead to the maximum expected value.

While this is clearly a very straightforward example to formalise, it does show why we feel the argumentation approach has some advantages. The use of the three separate systems makes it possible to separate out the belief elements from the value elements, and identify what reasoning is carried out with both. When belief and value arguments have been combined in LEV it is still clear which elements have been brought to bear. This makes it possible, for instance, to see that the reason that the option to concentrate on research loses out is because of the uncategorical relation $r1$ between concentrating on research and doing good research. This, in turn, gives the approach considerable explanatory power.

4.5 Soundness and completeness

As is the case for arguments about belief in the logic LA, it makes sense to ask what formal guarantees there are for arguments about values and expected values in the logics LV and LEV. The answer to this question is that there are none at the moment, and the investigation of such matters is one of the main foci of our future work on these systems. However, as remarked above, it is reasonably straightforward to obtain at least soundness proofs for both LV and LEV for the dictionaries discussed since all this involves is showing that the combination functions are correct. Furthermore, completeness proofs for systems such as LV and LEV are usually easy to obtain since they follow quite quickly from the inclusion of introduction and elimination rules for each of the connectives used.

4.6 Preferences and commitments

A complete decision theory is generally held to require some means of choosing between alternative actions. Despite the work outlined above the combined system LA/LV/LEV does not have such a mechanism. However, it is possible to extend the idea of arguments about values and expected values to provide one. In particular, we could use expected values to construct a preference ordering over a set of alternative actions as follows:

Condition St is *preferred* to condition St', $pref(St, St')$, if:

$$|\{(St : G_i : Sg_i) \mid \Delta \vdash_{LEV} (St : G_i : Sg_i)\}| \geq \\ |\{(St' : G_j : Sg_j) \mid \Delta \vdash_{LEV} (St' : G_j : Sg_j)\}|$$

In other words, St is preferred to St' if the overall force of all the expected value arguments for it is at least as great as the force of all the expected value arguments for St'. Transitivity of preferences is implicit in this inequality, and it is also possible to take into account the number of opposing arguments.

However we have a problem of potential instability. We could choose to act on a preference, but this preference could be transitory; wait a little longer and we might find that we can construct an argument to the effect that taking the currently preferred action could be disastrous. What is needed is some stronger condition than simply a preference for such and such an action. We would like to be able to prove that the ordering is, in fact, stable or that the benefits of achieving greater stability are outweighed by the costs. Thus we need some closure condition that says, essentially, there are no further arguments that could alter our main preference, a condition which parallels Pollock's [33] idea of a practical warrant for taking an action. Abstractly we can think of this as a "safety argument" of the form:

$$\frac{\begin{array}{l} best(A) : G : {++} \\ safe(A) : cir : {++} \end{array}}{commit(A) : (G, cir) : {++}}$$

where $best(A)$ means that aggregation of the arguments for an action A has greater force than the arguments for any alternative action, and $commit(A)$ represents a non-reversible commitment for executing action A, for example by executing it. Informally such safety arguments might include:

- Demonstrating that there are no sources of information that could lead to arguments which would result in a different best action.
- Demonstrating that the expected costs of not committing to A exceed the expected costs of seeking further information.

However, it is clear, as Pollock points out, that any system which is intended to have practical uses should take seriously the computational problems inherent in checking that no further relevant arguments can be built.

5 Argumentation in practice

While the work on arguments about values and expected values reported in Sections 4.2 and 4.3 is still rather preliminary, this paper being a first attempt at formalising the proposal made in [19], the work on arguments about beliefs has been applied quite widely in projects at the Imperial Cancer Research Fund. The systems to which this model has been applied include a decision support system for general medical practitioners [15], a system for interpreting medical images

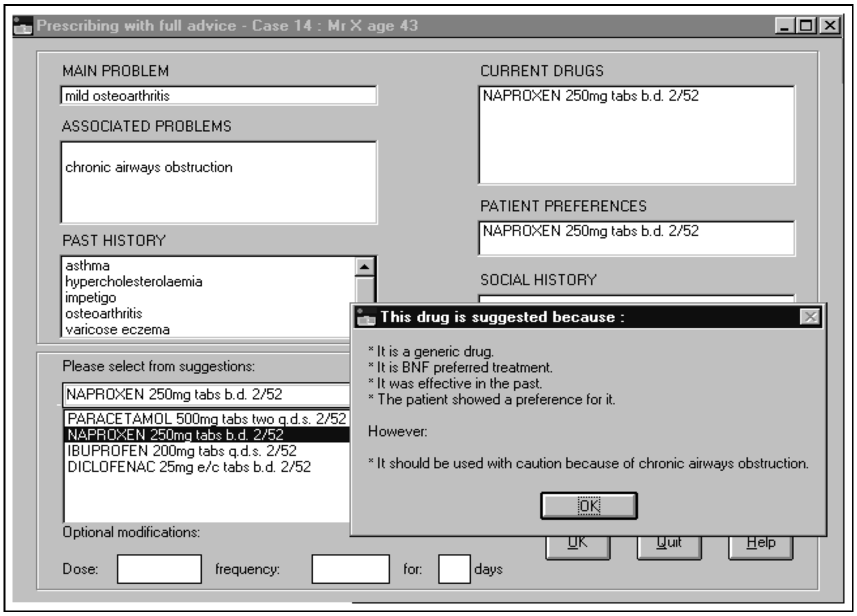

Fig. 13. A example consultation from CAPSULE

[41], and a system to advise on the management of acute asthma. More detail on these and other medical applications may be found in [14]. The model is also the basis for the system of argumentation used to analyse the risk of carcinogenicity of chemical compounds which is described elsewhere in this volume [23].

One recent application built using argumentation is the CAPSULE system which supports general practitioners in drug prescription. The system works in the classic expert system manner. It is equipped with information about which drugs treat which conditions and what constitutes best practice, and it is fed with information about patients. When a specific patient presents with a specific set of symptoms, CAPSULE identifies a range of drugs which are suitable, identifying for each the arguments for and against its use. The doctor can then choose the most appropriate. Figure 13 shows a typical consultation. The patient, who has a history which includes asthma, hypercholesterolaemia, impetigo, and varicose eczema has presented with mild osteoarthritis (a recurring condition). The system has identified a list of possible treatments, one of which is Naproxen. Examining the arguments for it, we find that it is a recommended treatment for four reasons. First, it is a generic drug, meaning it is not the trademarked product of a single drug company. Second, it is the treatment recommended by the BNF (the BNF being the British National Formulary, a list of drugs and the conditions they treat, which is the usual basis for prescribing decisions). Third, it has proved effective in the past (when the patient previously came to the doctor suffering from osteoarthritis). Fourth, the patient actually has a preference for it over other treatments he has tried in the past. There is also an argument against Naproxen—that it should be used with caution because of the associated problem of "chronic airways obstruction" which the patient is known to suffer from.

An evaluation of CAPSULE [43] suggests that this kind of support is extremely useful. A study was carried out in which 42 general practitioners each prescribed for 36 records based on real cases. The doctors were given 3 levels of support, a list of drugs in alphabetical order, a list of preferred drugs (decided upon by the argumentation engine), and the list of preferred drugs along with the arguments for and against. As the level of support increased, the proportion of times that the doctors agreed with a panel of experts rose from 25% to 42%, the proportion of time the doctors ignored a cheaper drug that was as effective as the one they chose fell from 50% to 35%, and the mean score (which measured how closely doctors agreed with the experts) rose from 6 (out of 8) to 6.7.

The version of CAPSULE from which Figure 13 is taken was developed using a system called Pro*forma* [16]. Pro*forma* is a generic technology for building decision support applications. It consists of the Pro*forma* language, a formal specification language in the sense used in software engineering, and a knowledge representation language. The technology also includes a number of software tools, for designing and "enacting" Pro*forma* applications. In particular, these include an editor which makes it possible to rapidly build applications from a set of standard components—plans, decisions, actions and enquiries. The use of argumentation is embedded in Pro*forma*'s decision component. All decisions are reached by building arguments for and against the decision options, and then aggregating these arguments to identify how good options are. Thus in the Pro*forma* version of CAPSULE, the system builds arguments for and against all the relevant drugs (which are precisely those for which arguments may be built) and uses an aggregation function which counts the number of supporting and opposing arguments, subtracts the second from the first, and ranks the decision options using the resulting score. This may appear to be a trivial procedure, but it does appear to be effective.

6 Conclusions and discussion

In order to take, or commit to, a decision we must combine or aggregate arguments in order to establish relative preferences among options. Perhaps surprisingly there is now considerable evidence that such simple decision functions are highly effective for many clinical applications (for instance [5, 24]). The most definitive study to date is that by Pradhan *et al.* who have rigorously assessed the impact of various evidence aggregation methods in medical decision making [34]. This study replicates the earlier findings cited, concluding that the correct qualitative representation of the decision has much more influence on the quality of decision making than the precision of quantitative parameters such as probabilities. We feel that this is strong evidence for the validity of argumentation as a decision making method.

This paper has built on our previous work on using argumentation to reason about beliefs towards making argumentation the basis of a complete decision theory. We identified a number of different types of argument that can participate in making decisions by reasoning about the outcome of possible actions and have suggested some ways in which these arguments may be built and combined. We believe that the framework we have outlined has the potential to integrate the best parts of traditional planning mechanisms and decision theory in the way suggested by Pollock [33] and Wellman and Doyle [45].

Furthermore, the theory seems to be capable of allowing meta-level reasoning about the structure of the decision as well as providing some means for coping with contradictory beliefs and conflicting values and for explicitly including stopping rules and commitment to particular courses of action. In addition to the obvious task of continuing the development of the foundations of this approach, there are a number of areas in which we are working. The first is to refine the set of values and expected values which may be used in order to make the system as expressive as, say, the systems proposed by Pearl [32] and Wilson [46]. The second is to investigate alternative semantics for values and expected values as, for instance, Dubois and Prade [9] have done. The third is to investigate the connections between the model we are proposing and existing means of combining plans and beliefs including the BDI framework [36] and the Domino model [6].

Much remains to be done to provide a secure foundation for this approach to reasoning and decision making but it appears to have potential merit for covering a comparable range of decisions to that addressed by classical decision theory. If this is the case, then the complete theory will provide a basis for implementing sound methods for decision making in the absence of quantitative information and the dynamic construction of the structure of the decision.

Acknowledgments

The first part of this paper has been distilled from the work of many of our colleagues without whose efforts little of the later work would have been possible. They are Simon Ambler, Mike Clarke, Subrata Das, Morten Elvang-Gørranson, Peter Hammond, and Paul Krause. The work that led to this paper has been partially funded by a number of grants, including Esprit Basic Research Actions 3085 DRUMS and 6156 DRUMS 2, ITD/4/1/9053 RED and DTI/SERC 1822 Praxis.

References

1. A. M. Agogino and N. F. Michelena. Qualitative decision analysis. In N. Piera Carreté and M. G. Singh, editors, *Qualitative Reasoning and Decision Technologies*, pages 285–293. CIMNE, Barcelona, Spain, 1993.
2. S. Ambler. A categorical approach to the semantics of argumentation. *Mathematical Structures in Computer Science*, 6:167–188, 1996.

3. S. Ambler and P. Krause. Enriched categories in the semantics of evidential reasoning. Technical Report 153, Advanced Computation Laboratory, Imperial Cancer Research Fund, 1992.
4. J. Bell and Z. Huang. Safety logics. In A. Hunter and S. Parsons, editors, *Applications of Uncertainty Formalisms (this volume)*. Springer Verlag, Berlin, 1998.
5. T. Chard. Qualitative probability versus quantitative probability in clinical diagnosis: a study using a computer simulation. *Medical Decision Making*, 11:38–41, 1991.
6. S. Das, J. Fox, D. Elsdon, and P. Hammond. Decision making and plan management by autonomous agents: theory, implementation and applications. In *Proceedings of the 1st International Conference on Autonomous Agents*, 1997.
7. S. K. Das. How much does an agent believe: an extension of modal epistemic logic. In A. Hunter and S. Parsons, editors, *Applications of Uncertainty Formalisms (this volume)*. Springer Verlag, Berlin, 1998.
8. D. Dubois and H. Prade. *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, NY, 1988.
9. D. Dubois and H. Prade. Possibility theory as a basis for qualitative decision theory. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pages 1924–1930, San Mateo, CA, 1995. Morgan Kaufmann.
10. M. Elvang-Gøransson and A. Hunter. Argumentative logics: reasoning with classically inconsistent information. *Data and Knowledge Engineering*, 16:125–145, 1995.
11. M. Elvang-Gøransson, P. Krause, and J. Fox. Dialectic reasoning with inconsistent information. In *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pages 114–121, San Mateo, CA, 1993. Morgan Kaufmann.
12. J. Fox. Making decisions under the influence of memory. *Psychological Review*, 87:190–211, 1980.
13. J. Fox, D. Barber, and K. D. Bardhan. Alternatives to Bayes? A quantitative comparison with rule-based diagnostic inference. *Methods of Information in Medicine*, 19:210–215, 1980.
14. J. Fox and S. Das. A unified framework for hypothetical and practical reasoning (2): lessons from medical applications. In *Formal and Applied Practical Reasoning*, pages 73–92, Berlin, Germany, 1996. Springer Verlag.
15. J. Fox, A. Glowinski, and M. O'Neil. The Oxford system of medicine: a prototype information system for primary care. In J. Fox, M. Fieschi, and R. Engelbrecht, editors, *AIME 87 European Conference on Artificial Intelligence in Medicine*. Springer Verlag, Berlin, 1987.
16. J. Fox, N. Johns, C Lyons, A. Rahmanzadeh, R. Thomson, and P. Wilson. Pro*forma*: a general technology for clinical decision support systems. *Computer Methods and Programs in Biomedicine*, 54:59–67, 1997.
17. J. Fox, P. Krause, and S. Ambler. Arguments, contradictions and practical reasoning. In *Proceedings of the 10th European Conference on Artificial Intelligence*, pages 623–627, Chichester, UK, 1992. John Wiley & Sons.
18. J. Fox, P. Krause, and M. Elvang-Gøransson. Argumentation as a general framework for uncertain reasoning. In *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pages 428–434, San Mateo, CA., 1993. Morgan Kaufmann.
19. J. Fox and S. Parsons. On using arguments for reasoning about actions and values. In *Proceedings of the AAAI Spring Symposium on Qualitative Preferences in Deliberation and Practical Reasoning*, pages 55–63, 1997.
20. D. Gabbay. *Labelled Deductive Systems*. Oxford University Press, Oxford, UK, 1996.

21. G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic*. Methuen, London, UK, 1968.
22. P. Krause, S. Ambler, M. Elvang-Gøransson, and J. Fox. A logic of argumentation for reasoning under uncertainty. *Computational Intelligence*, 11:113–131, 1995.
23. P. Krause, P. Judson, and M. Patel. Qualitative risk assessment fulfills a need. In A. Hunter and S. Parsons, editors, *Applications of Uncertainty Formalisms (this volume)*. Springer Verlag, Berlin, 1998.
24. M. O'Neil and A. Glowinski. Evaluating and validating very large knowledge-based systems. *Medical Informatics*, 3:237–251, 199o.
25. S. Parsons. Refining reasoning in qualitative probabilistic networks. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, pages 427–434, San Francisco, CA, 1995. Morgan Kaufman.
26. S. Parsons. Comparing normative argumentation to qualitative systems. In *Proceedings of the 6th International Conference on Information Processing and the Management of Uncertainty*, pages 137–142, 1996.
27. S. Parsons. Defining normative systems for qualitative argumentation. In *Formal and Applied Practical Reasoning*, pages 449–465, Berlin, Germany, 1996. Springer Verlag.
28. S. Parsons. Normative argumentation and qualitative probability. In *Qualitative and Quantitative Practical Reasoning*, pages 466–480, Berlin, Germany, 1997. Springer Verlag.
29. S. Parsons. On qualitative probability and order of magnitude reasoning. In *Proceedings of the 10th Florida Artificial Intelligence Research Symposium*, pages 198–203, St Petersburg, FL, 1997. Florida AI Research Society.
30. S. Parsons. *Qualitative approaches to reasoning under uncertainty*. MIT Press, Cambridge, MA, 1998.
31. S. Parsons and J. Fox. Argumentation and decision making: a position paper. In *Formal and Applied Practical Reasoning*, pages 705–709, Berlin, Germany, 1996. Springer Verlag.
32. J. Pearl. From conditional oughts to qualitative decision theory. In *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence*, pages 12–20, San Mateo, CA., 1993. Morgan Kaufmann.
33. J. L. Pollock. New foundations for practical reasoning. *Minds and Machines*, 2:113–144, 1992.
34. M. Pradhan, M. Henrion, G. Provan, B. del Favero, and K. Huang. The sensitivity of belief networks to imprecise probabilities: an experimental investigation. *Artificial Intelligence*, 85:363–397, 1996.
35. H. Raiffa. *Decision Analysis: Introductory Lectures on Choices under Uncertainty*. Addison-Wesley, Reading, MA, 1970.
36. A. Rao and M. P. Georgeff. Modelling rational agents within a BDI-architecture. In *Proceedings of the 2nd International Conference on Knowledge Representation and Reasoning*, pages 473–484, San Mateo, CA, 1991. Morgan Kaufmann.
37. G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.
38. Y. Shoham. Conditional utility, utility independence, and utility networks. In *Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence*, pages 429–436, San Francisco, CA, 1997. Morgan Kaufmann.
39. E. H. Shortliffe. *Computer-Based Medical Consultations: MYCIN*. Elsevier, New York, NY, 1976.

40. Sek-Wah Tan and J. Pearl. Qualitative decision theory. In *Proceedings of the 12th National Conference on Artificial Intelligence*, pages 928–933, Menlo Park, CA, 1994. AAAI Press/MIT Press.
41. P. Taylor, J. Fox, and A. Todd-Pokropek. A model for integrating image processing into decision aids for diagnostic radiology. *Artificial Intelligence in Medicine*, 9:205–225, 1997.
42. S. Toulmin. *The uses of argument*. Cambridge University Press, Cambridge, UK., 1957.
43. R. T. Walton, C. Gierl, P. Yudkin, H. Mistry, M. P. Vessey, and J. Fox. Evaluation of computer support for prescribing (CAPSULE) using simulated cases. *British Medical Journal*, 315:791–794, 1997.
44. M. P. Wellman. *Formulation of tradeoffs in planning under uncertainty*. Pitman, London, UK, 1990.
45. M. P. Wellman and J. Doyle. Preferential semantics for goals. In *Proceedings of the 10th National Conference on Artificial Intelligence*, pages 698–703, Menlo Park, CA, 1991. AAAI Press/MIT Press.
46. N. Wilson. An order of magnitude calculus. In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, pages 548–555, San Francisco, CA., 1995. Morgan Kaufman.