On using degrees of belief in BDI agents

Simon Parsons and Paolo Giorgini*
Department of Electronic Engineering
Queen Mary and Westfield College
University of London
London E1 4NS
United Kingdom

{S.D.Parsons,P.Giorgini}@qmw.ac.uk

Abstract

The past few years have seen a rise in the popularity of the use of mentalistic attitudes such as beliefs, desires and intentions to describe intelligent agents. Many of the models which formalise such attitudes do not admit degrees of belief, desire and intention. We see this as an understandable simplification, but as a simplification which means that the resulting systems cannot take account of much of the useful information which helps to guide human reasoning about the world. This paper starts to develop a more sophisticated system based upon an existing formal model of beliefs desires and intentions.

1 Introduction

In the past few years there has been a lot of attention given to building formal models of autonomous software agents; pieces of software which operate to some extent independently of human intervention and which therefore may be considered to have their own goals, and the ability to determine how to achieve their goals. Many of these formal models are based on the use of mentalistic attitudes such as beliefs, desires and intentions. The beliefs of an agent model what it knows about the world, the desires of an agent model which states of the world the agent finds preferable, and the intentions of an agent model those states of the world that the agent actively tries to bring about. One of the most popular and well-established of these models is the BDI model of Rao and Georgeff [12, 13].

While Rao and Georgeff's model explicitly acknowledges that an agent's model of the world is incomplete, by modelling beliefs as a set of worlds which the agent knows that it might be in, the model makes no attempt to make use of information about how likely a particular possible world is to be the actual world in which the agent operates. Our work is aimed at addressing this issue, which we feel is a weakness of the BDI model, by allowing an agent's beliefs, desires, and intentions to be quantified. In particular this paper considers quantifying an agent's beliefs using Dempster-Shafer theory, which immediately makes it possible for an agent to express its opinion on the reliability of the agents it interacts with, and to revise its beliefs when they become inconsistent. To do this, the paper combines the first author's work on the use of argumentation in BDI agents [11], with the second author's work on belief revision [4]. The question of quantifying desires and intentions is the subject of continuing work.

2 Preliminaries

As mentioned above, our work here is an extension of that in [11] to include degrees of belief. As in [11] we describe our agents using the framework of multi-context systems [8]. We do this because multi-context systems give a neat modular way of defining agents which is then directly executable, not because we are interested in explicitly modelling context. This section briefly recaps the notions of multi-context systems and argumentation as used in [11].

2.1 Multi-context agents

Using the multi-context approach, an agent architecture consists of the following four components (see [10] for a formal definition):

- *Units*: Structural entities representing the main components of the architecture. These are also called *contexts*.
- *Logics*: Declarative languages, each with a set of axioms and a number of rules of inference. Each unit has a single logic associated with it.

*Visiting from Istituto di Informatica, Universitá di Ancona, via Brecce Bianche, 60131, Ancona, Italy.

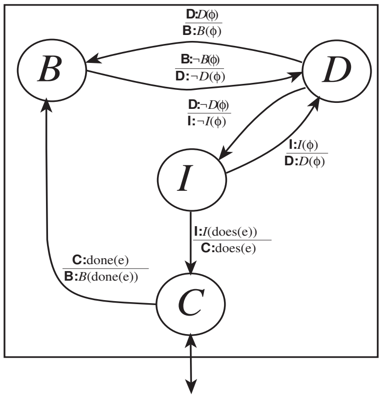

Figure 1: The multi-context representation of a strong realist BDI agent

- *Theories*: Sets of formulae written in the logic associated with a unit.
- *Bridge rules*: Rules of inference which relate formulae in different units.

The way we use these components to model BDI agents is to have separate units for belief B, desires D and intentions I, each with their own logic. The theories in each unit encode the beliefs, desires and intentions of specific agents, and the bridge rules encode the relationships between beliefs, desires and intentions. We also have a unit C which handles communication with other agents. Figure 1 gives a diagrammatic representation of this arrangement. For each of these four units we need to say what the logic used by each unit is. The communication unit uses classical first order logic with the usual axioms and rules of inference. The belief unit also uses first order logic, but with a special predicate B which is used to denote the beliefs of the agent. Under the modal logic interpretation of belief, the belief modality is taken to satisfy the axioms K, D, 4 and 5 [14]. Therefore, to make the belief predicate capture the behaviour of this modality, we need to add the following axioms to the belief unit (adapted from [2]):

$$\begin{array}{ll}
\mathbf{K} & B : B(\varphi \rightarrow \psi) \rightarrow (B(\varphi) \rightarrow B(\psi)) \\
\mathbf{D} & B : B(\varphi) \rightarrow \neg B(\neg\varphi) \\
\mathbf{4} & B : B(\varphi) \rightarrow B(B(\varphi)) \\
\mathbf{5} & B : \neg B(\varphi) \rightarrow B(\neg B(\varphi))
\end{array}$$

The desire and intention units are also based on first order logic, but have the special predicates D and I respectively. The usual treatment of desire and intention modalities is to make these satisfy the K and D axioms [14], and we capture this by adding the relevant axioms. For the desire unit:

$$\begin{array}{ll}
\mathbf{K} & D : D(\varphi \rightarrow \psi) \rightarrow (D(\varphi) \rightarrow D(\psi)) \\
\mathbf{D} & D : D(\varphi) \rightarrow \neg D(\neg\varphi)
\end{array}$$

and for the intention unit:

$$\begin{array}{ll}
\mathbf{K} & I : I(\varphi \rightarrow \psi) \rightarrow (I(\varphi) \rightarrow I(\psi)) \\
\mathbf{D} & I : I(\varphi) \rightarrow \neg I(\neg\varphi)
\end{array}$$

Each unit also contains the *generalisation*, *particularisation*, and *modus ponens* rules of inference. This completes the specification of the logics used by each unit.

The bridge rules are shown as arcs connecting the units. In our approach, bridge rules are used to enforce relations between the various components of the agent architecture. For example the bridge rule between the intention unit and the desire unit is:

$$I : I(\alpha) \Rightarrow D : D(\lceil I(\alpha) \rceil) \qquad (1)$$

meaning that if the agent has an intention α then it desires α[1]. The full set of bridge rules in the diagram are those for the "strong realist" BDI agent discussed in [14] :

$$\begin{array}{rcll}
D : \neg D(\alpha) & \Rightarrow & I : \neg I(\lceil \alpha \rceil) & (2) \\
D : D(\alpha) & \Rightarrow & B : B(\lceil \alpha \rceil) & (3) \\
B : \neg B(\alpha) & \Rightarrow & D : \neg D(\lceil \alpha \rceil) & (4) \\
C : \mathit{done}(e) & \Rightarrow & B : B(\lceil \mathit{done}(e) \rceil) & (5) \\
I : I(\lceil \mathit{does}(e) \rceil) & \Rightarrow & C : \mathit{does}(e) & (6)
\end{array}$$

The meaning of most of these rules is obvious. The two which require some additional explanation are (5) and (6). The first is intended to capture the idea that if the communication unit obtains information that some action has been completed (signified by the term *done*) then the agent adds it to its set of beliefs. The second is intended to express the fact that if the agent has some intention to do something (signified by the term *does*) then this is passed to the communication unit (and via it to other agents).

With these bridge rules, the shell of a strong realist BDI agent is defined in our multi-context framework. To complete the specification of a complete agent it is necessary to fill out the theories of the various units with domain specific information, and it may be necessary to add domain specific bridge rules between units. For an example, see [11].

2.2 Multi-context argumentation

The system of argumentation which we use here is based upon that proposed by Fox and colleagues [6, 9]. As with many systems of argumentation, it works by

[1] Because take B, D and I to be predicates rather than modal operators, when one predicate comes into the scope of another, for instance because of the action of a bridge rule, it needs to be quoted using $\lceil \cdot \rceil$.

constructing a series of logical steps (arguments) for and against propositions of interest and as such may be seen as an extension of classical logic. In classical logic, an argument is a sequence of inferences leading to a true conclusion. In the system of argumentation adopted here, arguments not only prove that propositions are true or false, but also suggest that propositions might be true or false. The strength of such a suggestion is ascertained by examining the propositions used in the relevant arguments.

We fit argumentation into multi-context agents by building arguments using the rules of inference of the various units and the bridge rules between units. The use we make of argumentation is summarised by the following schema:

$$\Gamma \vdash_d (\varphi, G, \alpha)$$

where:

- Γ is the set of formulae available for building arguments,
- $\vdash$ is a suitable consequence relation,
- $d = a_{\{r_1,\ldots,r_n\}}$ means that the formula φ is deduced by agent a from the set of formulae Γ by using the set of inference rules or bridge rules $\{r_1, \ldots, r_n\}$ (when there is no ambiguity the name of the agent will be omitted),
- φ is the proposition for which the argument is made,
- G indicates the set of formulae used to infer φ, $G \subseteq \Gamma$, and
- α is the degree of belief (also called "credibility") associated with φ as a result of the deduction.

This kind of reasoning is similar to that provided by labelled deductive systems [7], but it differs in its use of the labels. Whilst most labelled deductive systems use their labels to control inference, this system of argumentation uses the labels to determine which of its conclusions are most valid.

In the remainder of the paper we drop the 'B :', 'D :' and 'I :' to simplify the notation. With this in mind, we can define an argument in our framework:

Definition 1 *Given an agent a, an* argument *for a formula φ in the language of a is a triple (φ, P, α) where P is a set of grounds for φ and α is the degree of belief in φ suggested by the argument.*

It is the grounds of the argument which relate the formulae being deduced to the set of formulae it is deduced from:

Definition 2 *A set of* grounds *for φ in an agent a is an ordered set $\langle s_1, \ldots, s_n \rangle$ such that:*

1. *$s_n = \Gamma_n \vdash_{d_n} \varphi$;*
2. *every s_i, $i < n$, is either a formula in the theories of a, or $s_i = \Gamma_i \vdash_{d_i} \psi_i$; and*
3. *every p_j in every Γ_i is either a formula in the theories of agent a or ψ_k, $k < i$.*

We call every s_i a step *in the argument.*

For the sake of readability, we often refer to the conclusion of a deductive step with the identifier given to the step. For an example of how arguments are built, see Section 5.

3 A framework for adding degrees

In our previous work we have considered agents whose belief, desire and intention units contain formulae of the form:

$$B(\varphi) \wedge B(\varphi \rightarrow \psi) \rightarrow B(\psi)$$

These have then been used to build arguments as outlined in the previous section. What we want to do is to permit the beliefs, desires and intentions to admit degrees, so that beliefs can have varying degrees of credibility, desires can be ordered, and intentions adopted with varying degrees of resolution.

3.1 Degrees of belief

Since argumentation already allows us to incorporate degrees of belief it is reasonably straightforward to build in this component, and doing so is the subject of the rest of this paper. Degrees of desire and intention are more problematic, and are the subject of continuing work.

Given the machinery already provided by argumentation, the simplest way to build in degrees of belief is to translate every proposition in the belief unit that the agent is initially supplied with (which may contain nested modalities and so be of the form $B(I(\varphi))$) into an argument with an empty set of grounds. Thus $B(I(\varphi))$ becomes the argument:

$$(B(I(\varphi)) : \{\} : \alpha)$$

where α is the associated degree of belief expressed as a mass assignment in Dempster-Shafer theory [16]. Any propositions deduced from this base set will then accumulate grounds as detailed above. In an agent which has been interacting with other agents and making deductions about the world, we can distinguish four different types of proposition by looking at the origin of the propositions. We distinguish the following.

The *basic facts* are the data the agent was originally programmed with. An *observation* is a proposition which describes something the agent has observed about the world in which it is acting. A *communiqué* is a proposition which describes something the agent has received from an another agent. A *deduction* is a proposition that the agent has derived from some other pieces of information (which themselves will have been basic facts, deductions, observations or communiqués). Since the argument attached to each proposition records its origin, the four types of proposition may be distinguished by examining the arguments for them. The reason for distinguishing the types of proposition is that each is handled in a different way.

3.2 Handling communiqués

Consider first the way in which an agent handles an incoming communiqué. This is accepted by the communication unit, and given an argument which indicates which agent it came from and a degree of credibility which reflects the known reliability of that agent. When the communiqué is passed to the belief unit from the communication unit, the agent could be in two different situations.

In the first situation the communiqué is not involved in any conflict with other propositions in the belief unit. In this case, the following procedure is adopted:

1. Calculate the credibility of the new proposition.
2. Propagate the effect of this updating, recalculating the credibility of all the propositions whose arguments either include the new proposition or some consequence of the new proposition.

The credibility is calculated using Dempster-Shafer theory, and the precise way in which we do this depends upon the support for the communiqué. If the communiqué is the same as a proposition that was already in the belief unit, the agent uses both the reliability of the agent which passed it the communiqué and the credibility of the original proposition to calculate the credibility. If the communiqué was not already in the belief framework, the agent can use only the reliability of the agent which passed it the communiqué to calculate the credibility.

In the second situation the communiqué is in conflict with something in the belief unit. In this case we need to revise the agent's beliefs to make them consistent. However this can be done using information about the credibilities of the various beliefs, and the result of the revision also gives information about the reliability of the various agents who have supplied information. The following procedure is followed:

1. Revise the union of the set of beliefs in the belief unit and the new proposition which have been directly observed or communicated. To do this we can use the mechanism proposed in the next section. This mechanism will produce a new credibility degree for each proposition and a new reliability degree for each agent from which communications are received.
2. Pass the new reliability of each communicating agent to the communication unit.

3.3 Handling observations

Essentially same procedure as for communiqués is followed when an agent makes a new observation. The communication unit receives the proposition in question, flags it with a degree of reliability based on the behaviour of the sensor it came from, and passes it to the belief unit. The belief unit then carries out the same procedure as outlined above, but using the reliability of its sensors in place of the reliability of other agents.

3.4 Basic facts

Unlike observations and communiqués, new basic facts do not emerge during an agent's life—by definition they are programmed in when the agent is built. However, they are subject to change, since they are the very propositions which may conflict with observations and communiqués, and so when observations are made and communiqués are received, the basic facts are revised as discussed in the previous two sections.

3.5 Handling deductions

Like basic facts, new deductions are not received as input to the belief unit, but they are revised when observations and communiqués are transmitted to the belief unit. A slightly different procedure is used to revise deductions since they have arguments supporting them and the credibilities of the propositions in the argument are used in order to compute the credibility of the deduction. However, some of these propositions might be intentions or desires, "imported" into the belief unit via bridge rules. For such propositions it is not immediately clear what the credibility should be. For example, if we have the following bridge rule:

$$I : I(\alpha) \Rightarrow B : B(\lceil I(\alpha) \rceil)$$

and if in the intention unit we have $I : I(\alpha)$, then in the belief framework we will have $B : B(\lceil I(\alpha) \rceil)$. Now, what does the credibility of $B : B(\lceil I(\alpha) \rceil)$ depend on? The agent intends α, and this is not doubted. So, if we don't doubt the foundations of the bridge rule, we have to take the proposition as being true, that is with credibility equal to 1. So, if a proposition is supported through the bridge rules only by desires and intentions, its credibility degree will be equal to 1. If, on the other hand its supporting propositions contain some with degrees of credibility other than 1 (because they are based on information from unreliable agents) the overall credibility will be a combination of the credibilities of the unreliable agents. We can again use Dempster-Shafer to carry out the combination.

Another difference with deductions is that even when a deduction is in conflict with an observation or communiqué, the deduction itself is not directly revised. This is because this kind of conflict doesn't depend on the deduction but on the propositions which support it, as may be seen from the following example.

Example 1 Consider we have the following pieces of information:

1. $(\varphi, \{\}, C_\varphi)$
2. $(\varphi \to \psi, \{\}, C_{\varphi\to\psi})$
3. $(\neg\psi, \{\}, C_{\neg\psi})$

from (1) and (2) we have the deduction $(\psi, \langle\{\varphi, \varphi \to \psi\} \vdash_{\text{modus ponens}} \psi\rangle, C_\psi)$ which is in conflict with (3). This conflict depends on (3) and the supporting items (1) and (2). Thus revision must be applied to (1), (2) and (3) rather than the deduction. □

4 Belief revision and updating

Both belief revision and updating allow an agent to cope with a changing world by allowing it to alter its beliefs in response to new, possibly contradictory, information. We can say that:

> If the new information reports a change in the current state of a dynamic world, then the consequent change in the representation of the world is called *updating*.
>
> If the new information reports of new evidence regarding a static world whose representation was approximate, incomplete or erroneous, then the corresponding change is called *revision*.

In this section we will give a suitable mechanism for belief revision and updating in our framework.

4.1 Belief revision

The model for belief revision we adopt is drawn from [4]. Essentially, belief revision consists of redefining the degrees of credibility of propositions in the light of incoming information. The model adopts the *recoverability* principle:

> Any previously believed information item must belong to the current cognitive state if it is consistent with it.

Unlike the case in which incoming information is given priority, this principle makes sure that the chronological sequence of the incoming information has nothing to do with the credibility of that information, and that the changes are not irrevocable.

The propositions we called basic facts, observations and communiqués in the previous section are those items termed "assumptions" below (the term is that used in [4]), and the deductions are the "consequences". We have the following definitions

Definition 3 *A* knowledge base *(KB) is the set of the assumptions introduced from the various sources, and a* knowledge space *(KS) is the set of all beliefs (assumptions + consequences).*

Both the KB and KS grow monotonically since none of their elements are ever erased from memory. Normally both contain *contradictions*.

Definition 4 *A* nogood *is defined as minimal inconsistent subset of a KB. Dually, a* good *is a maximally consistent subset of a KB.*

Thus a nogood is a subset of KB that supports a contradiction and is not a superset of any other nogood. A good is a subset of a KB that is neither a superset of any nogood nor a subset of any other good. Each good has a corresponding *support set*, which is the subset of KS made of all the propositions that are in the good or are consequences of them. These definitions originate from de Kleer's work on assumption-based truth maintenance systems [3]. Procedurally, the method of belief revision consists of four steps:

- S1 Generating the set NG of all the nogoods and the set G of all goods in the KB.
- S2 Defining a credibility ordering over the assumptions in the KB.

S3 Extending this into a credibility ordering over the goods in G.

S4 Selecting the preferred good CG with its corresponding support set SS.

The first step S1 deals with consistency and adopts the set-covering algorithm [15] to find NG and the corresponding G. S2 deals with uncertainty and adopts the Dempster-Shafer theory of evidence [16] to find the credibility of the beliefs and Bayesian conditioning (see [5] for details) to calculate the new reliability of sources. S3 also deals with uncertainty, but at the level of the goods, extending the ordering defined by S2 over the assumptions, into an ordering onto the goods. There are a number of possible methods for doing this [1], including *best-out*, *inclusion-based* and *lexicographic*. An alternative is to order the goods according to the average credibility of their elements. Doing this, however, means that the preferred good may no longer necessarily contain the most credible piece of information. Finally S4 consists of two substeps: selecting a good CG from G (normally, CG is the good with the highest credibility) and selecting from KS the derived sentences that are consequences of the propositions belong to CG. Recapitulating we have:

INPUT:

- New proposition p;
- KB: set of all propositions introduced from the various sources (observations and communiqués); and
- Reliability of all sources.

OUTPUT:

- New credibilities of the propositions in $KB \cup \{p\}$;
- New credibilities of the goods in G;
- Preferred good CG and corresponding support set SS; and
- New reliability of all the sources.

4.2 Belief updating

If the particular application requires updating of beliefs instead of revision, then conceptually there is no difference in the dynamics of the propagation of weights. The main difference between the two procedures is that in updating the incoming information replaces the old. Thus the recoverability principle is substituted by the principle of priority of the incoming information. In order to explain what we exactly mean by updating consider the following example.

Example 2 Suppose the belief unit contains the propositions α and $\alpha \to \beta$. If the new proposition $\neg\beta$ is observed we will have a contradiction between $\alpha, \alpha \to \beta$ and $\neg\beta$ and consequently we will have three different goods:

1. $\{\alpha, \neg\beta\}$
2. $\{\neg\beta, \alpha \to \beta\}$
3. $\{\alpha, \alpha \to \beta\}$

Using belief revision we can choose one of them as the preferred good while updating we can't choose the third because it doesn't contain the new information. □

Thus the only difference between the belief revision and updating is the fourth step S4 of the belief revision procedure. We can define a different step for updating:

S4′ Selecting the preferred good CG which contains the new proposition, with its corresponding support set SS.

5 An example

As an example of the use of the degrees of belief in the multi-context BDI model, let consider the situation in Figure 2. The figure shows the base set of the agent's beliefs above the line and the deductions below it. The agent in question, Nico, knows that Paolo is dead, and also has information from a witness Carl which suggests that Benito shot Paolo, though Nico only judges Carl to be reliable to degree 0.5. From additional information Nico has about shooting and murdering she can conclude that Benito murdered Paolo, though her conclusion is not certain because there is some doubt about Carl's evidence. This conclusion takes the form of the argument:

$$(murderer(paolo, benito) : \langle\{1,2,5\} \vdash_{\mathrm{mp}} murderer(paolo, benito)\rangle : 0.5)$$

where (i) $murderer(paolo, benito)$ is the formulae which is the subject of the argument; (ii) the terms $\{1,2,5\}$[2] are the grounds of the argument which may be used along with modus ponens—signified by the "mp"—to infer $murderer(paolo, benito)$; and (iii) 0.5 is the sign.

If new information that Ana was with Benito at the time of the shooting comes from a second witness Dana, whose reliability is 0.6, then because Nico has

[2] These denote the formulae $dead(paolo)$, $shot(X,Y) \wedge dead(Y) \to murderer(Y,X)$ and $shot(benito, paolo)$.

Index	Argument	Source	Reliability
1	$(dead(paolo) : \{\} : 1)$	-	-
2	$(shot(X,Y) \wedge dead(Y) \rightarrow murderer(Y,X) : \{\} : 1)$	-	-
3	$(was\_with(X,Y) \rightarrow was\_with(Y,X) : \{\} : 1)$	-	-
4	$(was\_with(X,Y) \wedge murderer(Y) \rightarrow suspected(X) : \{\} : 1)$	-	-
5	$(shot(benito, paolo) : \{\} : 0.5)$	carl	0.5
6	$(murderer(paolo, benito) : \langle\{1,2,5\} \vdash_{\mathrm{mp}} murderer(paolo, benito))\rangle : 0.5)$	-	-

Figure 2: The initial state of Nico's belief context.

Index	Argument	Source	Reliability
1	$(dead(paolo) : \{\} : 1)$	-	-
2	$(shot(X,Y) \wedge dead(Y) \rightarrow murderer(Y,X) : \{\} : 1)$	-	-
3	$(was\_with(X,Y) \rightarrow was\_with(Y,X) : \{\} : 1)$	-	-
4	$(was\_with(X,Y) \wedge murderer(Y) \rightarrow suspected(X) : \{\} : 1)$	-	-
5	$(shot(benito, paolo) : \{\} : 0.5)$	carl	0.5
6	$(was\_with(ana, benito) : \{\} : 0.6)$	dana	0.6
7	$(murderer(paolo, benito) : \langle\{1,2,5\} \vdash_{\mathrm{mp}} murderer(paolo, benito))\rangle : 0.5)$	-	-
8	$(suspected(ana) : \langle\{4,6,7\} \vdash_{\mathrm{mp}} suspected(ana)\rangle : 0.3)$	-	-

Figure 3: Nico's belief context after Dana's evidence

some information about co-location and accomplice-hood, Ana becomes a suspect in the killing and Nico's belief context becomes that of Figure 3.

Suppose now that a new information comes from the witness Dana that Benito did not shoot Paolo. This information is not compatible with the Nico's proposition number 5, so the belief revision process calculates new degrees of credibility for her beliefs and new reliabilities for Carl and Dana. After this process Nico's new belief context is that of Figure 4 (where no deductions are shown). If new evidence against Benito emerges, for example an other agent Ewan, whose reliability Nico judges be 0.9, says that Benito did shoot Paolo, the belief context changes again. The belief revision mechanism starts from the reliabilities fixed *a priori* and Nico gets the context of Figure 5. The result of all these revisions is that Nico is fairly sure that Carl and Ewan are reliable and that Benito murdered Paolo. In addition, she believes that Dana is rather unreliable and so does not have much confidence that Ana is a suspect.

6 Summary

This paper has suggested a way of refining the treatment of beliefs in BDI models, in particular those built using multi-context systems as suggested in [11]. We believe that this work brings significant advantages. Firstly because the treatment is based upon the general ideas of argumentation, the approach we take is very general; it would, for instance, be simple to devise an analogous approach which made use of possibility measures rather than measures based on Dempster-Shafer theory. Secondly, the use of degrees of belief, as we have demonstrated, gives a plausible means of carrying out belief revision to handle inconsistent data, something that would be much harder to do in more conventional BDI models. Thirdly, introducing degrees of belief in propositions provides the foundation for using decision theoretic methods within BDI models; currently a topic which has had little attention. However, we acknowledge that this work is rather preliminary. In particular we need to extend the approach to deal with degrees of desire and intention, and to test out the approach in real applications. Both these directions are the topic of ongoing work.

References

[1] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, and H. Prade. Inconsistency management and prioritized syntax-based entailment. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, pages 640–645, 1995.

[2] B. F. Chellas. *Modal Logic: An Introduction.* Cambridge University Press, Cambridge, UK, 1980.

[3] J. de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28:127–162, 1986.

[4] A. Dragoni and P. Giorgini. Belief revision through the belief function formalism in a multi-

Index	Argument	Source	Reliability
1	$(dead(paolo) : \{\} : 1)$	-	-
2	$(shot(X,Y) \wedge dead(Y) \rightarrow murderer(Y,X) : \{\} : 1)$	-	-
3	$(was\_with(X,Y) \rightarrow was - with(Y,X) : \{\} : 1)$	-	-
4	$(was\_with(X,Y) \wedge murderer(Y) \rightarrow suspected(X) : \{\} : 1)$	-	-
5	$(shot(benito, paolo) : \{\} : 0.29)$	carl	0.29
6	$(was\_with(ana, benito) : \{\} : 0.42)$	dana	0.42
7	$(\neg shot(benito, paolo) : \{\} : 0.42)$	dana	0.42

Figure 4: Nico's belief context after Dana's second piece of evidence.

Index	Argument	Source	Reliability
1	$(dead(paolo) : \{\} : 1)$	-	-
2	$(shot(X,Y) \wedge dead(Y) \rightarrow murderer(Y,X) : \{\} : 1)$	-	-
3	$(was\_with(X,Y) \rightarrow was\_with(Y,X) : \{\} : 1)$	-	-
4	$(was\_with(X,Y) \wedge murderer(Y) \rightarrow suspected(X) : \{\} : 1)$	-	-
5	$(shot(benito, paolo) : \{\} : 0.88)$	carl	0.88
6	$(was\_with(ana, benito) : \{\} : 0.06)$	dana	0.06
7	$(\neg shot(benito, paolo) : \{\} : 0.06)$	dana	0.06
8	$(shot(benito, paolo) : \{\} : 0.88)$	ewan	0.88
9	$(murderer(paolo, benito) : \langle\{1,2,5\} \vdash_{\mathrm{mp}} murderer(paolo, benito),$ $\{1,2,8\} \vdash_{\mathrm{mp}} murderer(paolo, benito)\rangle : 0.88)$	- -	-
10	$(suspected(ana) : \{4,7,9\} : 0.06)$	-	-

Figure 5: Nico's belief context after Ewan's evidence.

agent environment. In *Proceedings of the 3rd International Workshop on Agent Theories, Architectures and Languages*, 1996.

[5] A. Dragoni and P. Giorgini. Learning agents' reliability through Bayesian conditioning: a simulation study. In *Learning in DAI Systems*, pages 151–167, 1997.

[6] 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, 1992.

[7] D. Gabbay. *Labelled Deductive Systems*. Oxford University Press, Oxford, UK, 1996.

[8] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: How we can do without modal logics). *Artificial Intelligence*, 65:29–70, 1994.

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

[10] P. Noriega and C. Sierra. Towards layered dialogical agents. In *Proceedings of the 3rd International Workshop on Agents Theories, Architectures and Languages*, pages 157–171, 1996.

[11] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 1998, (to appear).

[12] A. Rao and M. Georgeff. BDI agents: From theory to practice. In *Proceedings of the 1st International Conference on Multi-Agent Systems*, pages 312–319, 1995.

[13] A. S. Rao and M. P. Georgeff. Modeling Rational Agents within a BDI-Architecture. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, 1991.

[14] A. S. Rao and M. P. Georgeff. Formal Models and Decision Procedures for Multi-Agent Systems. Technical Note 61, Australian Artificial Intelligence Institute, 1995.

[15] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 53, 1987.

[16] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.