Modelling Uncertainty in Agent Programming

Johan Kwisthout* and Mehdi Dastani

ICS, Utrecht University
{johank, mehdi}@cs.uu.nl

Abstract. Existing cognitive agent programming languages that are based on the BDI model employ logical representation and reasoning for implementing the beliefs of agents. In these programming languages, the beliefs are assumed to be certain, i.e. an implemented agent can believe a proposition or not. These programming languages fail to capture the underlying uncertainty of the agent's beliefs which is essential for many real world agent applications. We introduce Dempster-Shafer theory as a convenient method to model uncertainty in agent's beliefs. We show that the computational complexity of Dempster's Rule of Combination can be controlled. In particular, the certainty value of a proposition can be deduced in linear time from the beliefs of agents, without having to calculate the combination of Dempster-Shafer mass functions.

1 Introduction

In multi-agent systems, individual agents are assumed to be situated in some environment and are capable of autonomous actions in the environment in order to achieve their objectives [20]. An autonomous agent interacts with its environment, based on its information and objectives, both of which are updated with the information acquired through interaction. In order to develop multi-agent systems, many programming languages have been proposed to implement individual agents, their environments, and interactions [5, 10, 4, 3]. These languages provide programming constructs to enable the implementation of agents that can reason about their information and objectives and update them according to their interactions.

Unfortunately, although Rao and Georgeff [12] already uses beliefs - rather than knowledge operators - due to the agent's lack of knowledge about the state of the world, many of the proposed programming languages assume that the information and objectives of agents are certain. This is obviously an unrealistic assumption for many real world applications. In such applications, either the environment of the agents involves uncertainty or the uncertainty is introduced to agents through imperfect sensory information. Past research dealing with the application of existing programming languages such as 3APL for robot control [17] showed, that sensory input is not always accurate, and that external actions

* The work of this author was partially supported by the Netherlands Organisation for Scientific Research NWO

have unpredictable outcomes: the environment in which the agent operates is both inaccessible and indeterministic. This seriously devalues the practical use of such agent programming languages for real applications like mobile robot control. Therefore, we believe that individual agents need to be able to reason and update their states with uncertain information, and that agent-oriented programming languages should facilitate these functionalities.

In this paper, we focus on cognitive agents which can be described and implemented in terms of cognitive concepts such as beliefs and goals. We consider programming languages that provide programming constructs to implement agent's beliefs, to reason about beliefs, and update beliefs. In order to allow the implementation of cognitive agents that can work with uncertain information, we investigate the possible use of *Dempster-Shafer theory* to incorporate uncertainty in BDI-type agent programming languages. We discuss how uncertain beliefs can be represented and reasoned with, and how they can be updated with uncertain information.

The structure of this paper is as follows. First we will introduce the most important relevant concepts in Dempster-Shafer theory. In section 3 we propose a mapping between this theory and agent beliefs. In section 4 we deal with implementational issues, and show that computational complexity can be controlled given certain restrictions on the belief representation. In section 5 we show how the agent's belief base can be updated and queried, while the computational complexity of these operations are discussed in section 6. Finally, in section 7 we conclude the paper.

2 Dempster-Shafer theory

The concept *uncertainty* is closely related to probability theory. We differentiate between the notion of *chance* and *probability*: a chance represents an objective, statistical likeliness of an event (such as throwing a six with a dice), while probability represents the likeliness of an event given certain subjective knowledge (for example, the probability of six, given that we know the number is even). Probabilistic reasoning deals with the question how *evidence* influences our *belief* in a certain hypothesis H. We define the probability of H, denoted as $P(H)$, as a real number between 0 and 1, with $P(H) = 0$ meaning H is definitely false, and $P(H) = 1$ meaning H is definitely true. A value between 0 and 1 is a measure for the probability of H.

The theory of Dempster and Shafer [14] can be seen as a generalisation of probability theory. In this theory, a *frame of discernment* Ω is defined as the set of all hypotheses in a certain domain. On the power set 2^{Ω}, a *mass function* $m(X)$ is defined for every $X \subseteq \Omega$, with $m(X) \geq 0$ and $\sum_{X \subseteq \Omega} m(X) = 1$. If there is no information available with respect to Ω, $m(\Omega) = 1$, and $m(X) = 0$ for every subset of Ω. For example, in a murder case Ω is a list of suspects, {Peter, John, Paul, Mary, Cindy}. If the investigator has no further information, the mass function associated with Ω will assign 1 to Ω and 0 to all real subsets of Ω. If there is evidence found regarding certain subsets of Ω, for example a slightly unreliable witness claims the killer was probably a male, we assign an appropriate mass value (say 0.6) to this particular subset of Ω and - since we have no further information and $\sum_{X \subseteq \Omega} m(X) = 1$ by definition - we assign a mass value of 0.4 to Ω. The mass function in this case would be:

$$m_1(X) = \begin{cases} 0.6 \text{ if } X = \text{\{Peter, John, Paul\}} \\ 0.4 \text{ if } X = \Omega \\ 0 \quad \text{otherwise} \end{cases}$$

Note that no value whatsoever is assigned to *subsets* of {Peter, John, Paul}. If we receive further evidence, for example that the killer was most likely (say with a probability of 0.9) left-handed, and both John and Mary are left-handed, then we might have another mass function like:

$$m_2(X) = \begin{cases} 0.9 \text{ if } X = \text{\{John, Mary\}} \\ 0.1 \text{ if } X = \Omega \\ 0 \quad \text{otherwise} \end{cases}$$

Dempster's Rule of Combination is a method to combine both pieces of evidence into one combined mass function. This function for the combination of $m_1 \oplus m_2$ is defined as:

Definition 1. *(Dempster's Rule of Combination [14]). Let $X, Y, Z \subseteq \Omega$. Then the following holds:*

$$m_1 \oplus m_2(X) = \frac{\sum\limits_{Y \cap Z = X} m_1(Y) \cdot m_2(Z)}{\sum\limits_{Y \cap Z \neq \emptyset} m_1(Y) \cdot m_2(Z)} \qquad \textit{and}$$

$$m_1 \oplus m_2(\emptyset) = 0$$

Dempster's Rule of Combination is commutative and associative, as shown in [13]. In our example, combining both pieces of evidence would lead to the following mass function:

$$m_1 \oplus m_2(X) = \begin{cases} 0.06 \text{ if } X = \text{\{Peter, John, Paul\}} \\ 0.36 \text{ if } X = \text{\{John, Mary\}} \\ 0.54 \text{ if } X = \text{\{John\}} \\ 0.04 \text{ if } X = \Omega \\ 0 \quad \text{otherwise} \end{cases}$$

Given a certain mass function, the subsets of Ω that have a mass value greater than zero are called *focal elements*, and we will denote the set of focal elements of a given mass function φ as the *core* of that mass function. A *simple support function* is a special case of a mass function, where the evidence only supports a certain subset A of Ω, and zero mass is assigned to all subsets of Ω other than A, i.e., the core of a simple support function is $\{A, \Omega\}$:

Definition 2. *(simple support function [14]). Let $X \subseteq \Omega$ and A be an evidence with probability s. Then, the simple support function related to A is specified as follows:*

$$m(X) = \begin{cases} s & \textit{if } X = A \\ \textit{1-s} & \textit{if } X = \Omega \\ 0 & \textit{otherwise} \end{cases}$$

On a mass function, two other functions are defined, namely a *belief function* $Bel(X)$ and a *plausibility function* $Pl(X)$.

Definition 3. *(belief and plausibility function [14]). Let $X, Y \subseteq \Omega$, then the belief and plausibility functions can be defined in terms of a certain mass function m as follows:*

$$Bel(X) = \sum_{Y \subseteq X} m(Y) \qquad \textit{and} \qquad Pl(X) = \sum_{X \cap Y \neq \emptyset} m(Y)$$

Informally, the belief and plausibility functions can be seen as a lower respectively upper limit on the probability of the set of hypotheses X. Note, that $Pl(X) = 1 - Bel(\Omega \backslash X)$. The difference between $Bel(X)$ and $Pl(X)$ can be regarded as the *ignorance* with respect to X.

3 Mapping agent beliefs to Dempster-Shafer sets

Can the theory of Dempster and Shafer be applied to the beliefs of an agent, if they are represented by proposition-logical formulae in an agent programming language? To investigate this question, suppose we have an agent-based program that operates in the context of a *2-by-2* grid-world where bombs can appear in certain positions in the grid and an agent can partially perceive the environment and move around. The agent tries to sense the bombs surrounding him, thus locating all bombs and safe squares in his environment. We assume that the agent's belief base is a set of logical formulae that the agent believes to hold. This belief base can therefore be understood as the conjunction of all formulae from the set, i.e. it can be represented as a conjunctive formula. Assume that, at a given moment during the execution of the program, the agent has the formula safe(1) in its belief base (say BB_1).

This indicates that the agent believes that square 1 is a safe location. How can we relate this belief in terms of the Dempster-Shafer theory? The frame of discernment Ω can be understood as the set of all models of the grid-world, as shown in table 1. In a 2-by-2 grid-world there are 16 models, ranging from 'all squares are safe' to 'all squares contain bombs'. We can relate the agent's current beliefs to a subset of hypotheses from Ω, where *each hypothesis is considered as a model of that belief.*

For example, if we define the hypotheses as in table 1 then the belief formula $safe(1)$ is a representation of the set $\{H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8\}$ of hypotheses, which is exactly the set of all models of the belief base BB_1. If we define a mass-function $m_{safe(1)}$ according to this belief base, we would assign 1 to this set, and 0 to Ω (and to all other subsets of Ω). In fact, each belief base can be represented by a mass function. Such a mass function would assign 1 to

Hyp.	1	2	3	4	Hyp.	1	2	3	4
1	Safe	Safe	Safe	Safe	9	Bomb	Safe	Safe	Safe
2	Safe	Safe	Safe	Bomb	10	Bomb	Safe	Safe	Bomb
3	Safe	Safe	Bomb	Safe	11	Bomb	Safe	Bomb	Safe
4	Safe	Safe	Bomb	Bomb	12	Bomb	Safe	Bomb	Bomb
5	Safe	Bomb	Safe	Safe	13	Bomb	Bomb	Safe	Safe
6	Safe	Bomb	Safe	Bomb	14	Bomb	Bomb	Safe	Bomb
7	Safe	Bomb	Bomb	Safe	15	Bomb	Bomb	Bomb	Safe
8	Safe	Bomb	Bomb	Bomb	16	Bomb	Bomb	Bomb	Bomb

Table 1. Bomb location and associated hypothesis

the particular subset of Ω that contains all hypotheses that are true with respect to the belief base, or in other words: the maximal subset of hypotheses in Ω that are models of safe(1). Notice that the set $\{H_1, H_2, H_3, H_4, H_5, H_6\}$ consists of models of safe(1) as well, but it is not the maximal subset with this property. If a belief base is a certain belief formula φ, then it could be represented by a simple support function $m_\varphi(X)$ that supports only the maximal set of hypotheses in Ω that are models of φ. This can be formalised as follows:

$$m_\varphi(X) = \begin{cases} 1 \text{ if } X \subseteq \Omega \ \ \& \text{ models}(X, \varphi) \ \ \& \ \ \forall Y \subseteq \Omega \ (\text{models}(Y, \varphi) \Rightarrow Y \subseteq X) \\ 0 \text{ otherwise} \end{cases}$$

In this definition, the relation *models(X, φ)* is defined as $\forall M \in X \ \ M \models \varphi$, where M is a model and $\models$ is the propositional satisfaction relation. The condition of the if-clause indicates that X is the maximum set of hypotheses in Ω that are models of φ. In the sequel we use the function $\max_\Omega(\varphi)$ to denote this set, and define it as follows: $\max_\Omega(\varphi) = X \iff X \subseteq \Omega \ \& \text{ models}(X, \varphi) \ \& \ \forall Y \subseteq \Omega \ (\text{models}(Y, \varphi) \Rightarrow Y \subseteq X)$. Using this auxillary function, the mass function that represents the belief base BB_1 can then be rewritten as:

$$m_{safe(1)}(X) = \begin{cases} 1 \text{ if } X = \max_\Omega(\text{safe}(1)) \\ 0 \text{ otherwise} \end{cases}$$

3.1 Adding beliefs

If we add another belief formula to the belief base, the resulting belief base can be represented by the combination of the mass function of both belief formulae. Suppose we add *safe(2)* to the belief base, with the following mass function:

$$m_{safe(2)}(X) = \begin{cases} 1 \text{ if } X = \max_\Omega(\text{safe}(2)) \\ 0 \text{ otherwise} \end{cases}$$

We can combine both pieces of evidence using Dempster's Rule of Combination. Since the only non-empty intersection of sets defined by either $m_{safe(1)}$ or $m_{safe(2)}$ is the set $\max_\Omega(\text{safe}(1) \wedge \text{safe}(2))$, the resulting mass function $m_1 =$

$m_{safe(1)} \oplus m_{safe(2)}$ is defined as follows[1]:

$$m_1(X) = \begin{cases} 1 \text{ if } X = \max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2)) \\ 0 \text{ otherwise} \end{cases}$$

Note that $\max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2))$ corresponds to the subset $\{H_1, H_2, H_3, H_4\}$ of hypotheses. Apart from these beliefs, which can be either true or false, we could imagine a situation where a belief is uncertain. We might conclude, on the basis of specific evidence, that a location *probably* contains a bomb; such a belief formula (say *bomb(3)* with a probability value of 0.7), should be added to the belief base. In order to incorporate such cases, we introduce the concept of a *basic belief formula* to represent uncertain belief formulae.

Definition 4. *(Basic belief formula). Let φ be a belief formula and $p \in [0..1]$. Then the pair $\varphi : p$, which indicates that φ holds with probability p, will be called a basic belief formula*[2].

With these basic belief formulae, the above mentioned belief base can be represented as { safe(1): 1.0, safe(2): 1.0, bomb(3): 0.7 }. Of course, we could represent *bomb(3): 0.7* as a mass function, as we did with beliefs safe(1) and safe(2). This mass function would assign a probability value of 0.7 to the set of hypotheses, all having a bomb on location 3, and (because we have no further information) a probability value of 0.3 tot Ω:

$$m_{bomb(3)}(X) = \begin{cases} 0.7 \text{ if } X = \max_{\Omega}(\text{bomb}(3)) \\ 0.3 \text{ if } X = \Omega \\ 0 \quad \text{otherwise} \end{cases}$$

If we would combine m_1 and $m_{bomb(3)}$ using Dempster's Rule of Combination, we would get the following mass function:

$$m_2 = m_1 \oplus m_{bomb(3)}(X) = \begin{cases} 0.7 \text{ if } X = \max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2) \wedge \text{bomb}(3)) \\ 0.3 \text{ if } X = \max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2)) \\ 0 \quad \text{otherwise} \end{cases}$$

This combined mass function m_2 represents our updated belief base. Note, that the set $\max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2))$ is exactly $\{H_1, H_2, H_3, H_4\}$, and the set $\max_{\Omega}(\text{safe}(1) \wedge \text{safe}(2) \wedge \text{bomb}(3))$ is equal to $\{H_3, H_4\}$.

3.2 Deleting beliefs

We can also delete belief formulae from our belief base. For example, we could conclude $\neg$*bomb(3)* during the execution of our program. We will model deletion of a formula as the addition of its negation. This corresponds to the maximal set of hypotheses according to which there is *no* bomb on location 3:

[1] We will use simple indices for the combined mass functions to improve readability.

[2] The term *basic belief formula* should not be confused with an *atomic belief formula*. Note, that the probability assigned to a basic belief formula cannot be (further) distributed to the atomic formulae that may constitute the basic belief formula.

$$m_{\neg bomb(3)}(X) = \begin{cases} 1 \text{ if } X = \max_{\Omega}(\neg\text{bomb}(3)) \\ 0 \text{ otherwise} \end{cases}$$

Combining $m_{safe(1)}$ and $m_{\neg bomb(3)}$ leads to the following mass function:

$$m_3 = m_{save(1)} \oplus m_{\neg bomb(3)}(X) = \begin{cases} 1 \text{ if } X = \max_{\Omega}(\text{save}(1) \wedge \neg\text{bomb}(3)) \\ 0 \text{ otherwise} \end{cases}$$

Of course, we could also conclude that a certain belief becomes less probable instead of impossible. In that case, the negation of the formula under consideration will be added with a certainty value, for example *¬bomb(3): 0.3*. We would represent this formula as:

$$m_{\neg bomb(3)}(X) = \begin{cases} 0.3 \text{ if } X = \max_{\Omega}(\neg\text{bomb}(3)) \\ 0.7 \text{ if } X = \Omega \\ 0 \quad \text{otherwise} \end{cases}$$

Combining this alternative mass function $m_{\neg bomb(3)}$ and $m_{safe(1)}$ leads to the following mass function:

$$m_4 = m_{safe(1)} \oplus m_{\neg bomb(3)}(X) = \begin{cases} 0.59 \text{ if } X = \max_{\Omega}(\text{save}(1) \wedge \neg\text{bomb}(3)) \\ 0.41 \text{ if } X = \max_{\Omega}(\text{save}(1) \wedge \text{bomb}(3)) \\ 0 \quad \text{otherwise} \end{cases}$$

3.3 Composite beliefs

Until now we have only used atomic formula in our examples. However, we can also model disjunctions, conjunctions and negation of beliefs as sets of hypotheses by mapping disjunction, conjunction and negation of beliefs, to respectively unions, intersections, and complements of sets of hypotheses. We can illustrate such composite beliefs with an example. Consider the following two formulae in the already mentioned grid-world:

φ: safe(2) $\wedge$ (safe(3) $\vee$ safe(4)), and ψ: safe(1) $\vee$ ($\neg$safe(2) $\wedge$ safe(3)).

These formulae correspond to the the sets $\{H_1, H_2, H_3, H_9, H_{10}, H_{11}\}$ and $\{H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8, H_{13}, H_{14}\}$, respectively. If φ has a probability of p, and ψ has a probability of q, then these formula could be represented by basic belief formulae as follows:

$$m_{\varphi}(X) = \begin{cases} p \quad\quad \text{if } X = \max_{\Omega}(\text{safe}(2) \wedge (\text{safe}(3) \vee \text{safe}(4))) \\ 1 - p \text{ otherwise} \end{cases}$$

$$m_{\psi}(X) = \begin{cases} q \quad\quad \text{if } X = \max_{\Omega}(\text{safe}(1) \vee (\neg\text{safe}(2) \wedge \text{safe}(3))) \\ 1 - q \text{ otherwise} \end{cases}$$

Obviously, the conjunction $\varphi \wedge \psi$ equals safe(1) $\wedge$ safe(2) $\wedge$ (safe(3) $\vee$ safe(4)), and from table 1 follows, that this result corresponds to the set $\{H_1, H_2, H_3\}$, which is the intersection of $\{H_1, H_2, H_3, H_9, H_{10}, H_{11}\}$ and $\{H_1, H_2, H_3, H_4, H_5, H_6, H_7, H_8, H_{13}, H_{14}\}$.

3.4 Inconsistency problem

The issue of inconsistency in Dempster's Rule of Combination deserves further attention. In the original rule, as defined in definition 1, combinations that lead to an empty set have a mass probability of zero, and the other combinations are scaled to make sure all mass probabilities add to one. This leads to unexpected results when two mass functions with a high degree of conflict are combined. This can be demonstrated with an often-used example (e.g. in [9]). In a murder case there are three suspects: Peter, Paul and Mary. There are two witnesses, who both give highly inconsistent testimonies, which can be represented with the following mass functions:

$$m_1(X) = \begin{cases} 0.99 \text{ if } X = \text{'killer is Peter'} \\ 0.01 \text{ if } X = \text{'killer is Paul'} \\ 0 \quad \text{otherwise} \end{cases} \quad m_2(X) = \begin{cases} 0.99 \text{ if } X = \text{'killer is Mary'} \\ 0.01 \text{ if } X = \text{'killer is Paul'} \\ 0 \quad \text{otherwise} \end{cases}$$

Combining these two mass functions leads to a certain belief that Paul is the killer, although there is hardly any support in either of the witnesses' testimonies.

$$m_1 \oplus m_2(X) \begin{cases} 1 \text{ if } X = \text{'killer is Paul'} \\ 0 \text{ otherwise} \end{cases}$$

Sentz [13] describes a number of alternatives for the Rule of Combination. The most prominent (according to Sentz) is Yager's *modified Dempster's Rule*[21]. Ultimately, this rule attributes the probability of combinations, which lead to the empty set, to Ω [3]. A similar approach is demonstrated by Smets [15], which states that in the case of inconsistent mass functions, the closed world assumption (the assumption that one of the three suspects is the murderer) is not valid. The probability of empty sets should be attributed to $\emptyset$, as a sort of 'unknown third'. This would lead to a mass function of:

$$m_1 \oplus m_2(X) \begin{cases} 0.0001 \text{ if } X = \text{'killer is Paul'} \\ 0.9999 \text{ if } X = \Omega \text{ (Yager), resp., } X = \emptyset \text{ (Smets)} \\ 0 \qquad \text{otherwise} \end{cases}$$

Jøsang [9] poses an alternative, namely the *consensus operator*, which attributes the *means* of the probabilities of two inconsistent beliefs to the combination, rather than their multiplication:

$$m_1 \oplus m_2(X) \begin{cases} 0.495 \; X = \text{'killer is Peter'} \\ 0.495 \; X = \text{'killer is Mary'} \\ 0.01 \quad X = \text{'killer is Paul'} \\ 0 \qquad X = \emptyset \end{cases}$$

[3] To be more exact, Yager differentiates between *ground probabilities q(X)* and *basic probabilities m(X)*. The empty set can have a $q(\emptyset) \geq 0$. When combining, these ground probabilities are used and the mass is attributed *after* the combination, where $m(X) = q(X)$ for $X \neq \emptyset$, and $m(\Omega) = q(\Omega) + q(\emptyset)$.

Suspect	W_1	W_2	Dempster	Yager/Smets	Jøsang
Peter	0.99	0	0	0	0.495
Paul	0.01	0.01	1	0.0001	0.01
Mary	0	0.99	0	0	0.495
$\emptyset$ or Ω	0	0	0	0.9999	0

Table 2. Attribution of mass to inconsistent combinations

We can summarise these approaches (Dempster, Yager/Smets and Jøsang) using the 'murder case' example, as shown in table 2. Note, that the issue of inconsistency directly relates to the choice of the frame of discernment Ω. In this example, we restrict our frame of discernment to be the set of three *mutually exclusive* hypotheses, namely {Paul, Peter, Mary}. If, on the other hand, our frame would be Ω = {Paul, Peter, Mary, Peter or Mary}, and we would map the phrase 'killer is Peter' to the subset {Peter, Peter or Mary} and the phrase 'killer is Mary' to the subset {Mary, Peter or Mary}, then there would be no inconsistency at all. We deal with the choice of our frame of discernment in the next section.

3.5 The frame of discernment

Until now, we have mapped agent beliefs to a given set of 16 hypotheses in a 2-by-2 grid-world. Unfortunately, the frame of discernment that corresponds to a given agent program is unknown, and, just as important, there is no unique frame of discernment in such a program. We might just as well add a totally irrelevant hypothesis H_{17}, stating 'All squares contain apples'. We do not know if a certain hypothesis, say H_{16}, can become true at all during the execution of the program. This implies that the relation between the frame of discernment and the agent's beliefs is a many-to-many mapping.

This problem can, however, be solved. According to some agent programming languages such as 3APL [5], the number of beliefs an agent can hold during execution of the program is finite. For example, in the programming language 3APL only *basic actions* can update the belief base. The update corresponding to a basic action is specified as the post-condition of the basic action which is determined by the programmer before running the program. Therefore, in a given 3APL program all possible beliefs are given either by the initial belief base or by the specification of the basic actions. For this type of agent programs we can construct a *theoretical* frame of discernment that includes a set of hypotheses such that each belief that the agent can hold during its execution can be mapped to a subset of the frame of discernment. Shafer states [14, p.281], that in general the frame of discernment cannot be determined beforehand (i.e. without knowing which evidence might be relevant), and that we tend to enlarge it as more evidence becomes available. But, on the other side, if Ω is too large, holding too much irrelevant hypotheses, the probability of any hypothesis is unreasonably small. By stating that the frame of discernment should be large enough to hold all relevant hypotheses with respect to the program under consideration, Ω will be neither too small nor too large.

In this paper, we demand that Ω should be such that for each belief that an agent can hold during its execution (i.e. each combination of the basic belief formulae) there must be at least one non-empty subset of hypotheses in Ω. This corresponds to the demand, that the belief base of an agent should be consistent and remain so during its execution. This is the case with, e.g., 3APL agents. In other words, each conjunction of basic belief formulae has a non-empty subset of hypotheses from Ω that are models of the conjunction of basic belief formulae.

4 Mass calculation

As we have seen, any given belief base can be represented with a mass function. Generally, a belief formula $b_i$with probability p_i divides the set of all hypotheses Ω into:

$$m_{b_i}(\mathrm{X}) = \begin{cases} p_i & \text{if } X = \max_\Omega(b_i) \\ 1 - p_i & \text{if } X = \Omega \\ 0 & \text{otherwise} \end{cases}$$

The combination of belief formulae $b_1, \ldots, b_n$ can thus be represented with a mass function $m_k = m_1 \oplus \ldots \oplus m_n$, related to beliefs $b_1, \ldots, b_n$, where the number of subsets of Ω that are used to define m_k and have a mass value greater than zero, is equal to 2^n. When a belief formula m_n is combined with an already existing combination $m_1 \oplus \ldots \oplus m_{n-1}$, the resulting mass function $m_1 \oplus \ldots \oplus m_n$ is defined by the non-empty intersections of all subsets of Ω in m_n, with all subsets of Ω in $m_1 \oplus \ldots \oplus m_{n-1}$. Since we use simple support functions to represent our beliefs - with focal elements Ω and $\max_\Omega(\varphi)$ - the number of resulting subsets is doubled with each added belief formula.

Because n belief formulae lead to a mass function of 2^n combinations, keeping a mass function in memory and updating it when the belief base changes will lead to a combinatorial explosion in both processing time and memory requirements, as the following scenario will show. Suppose we start with an empty belief base, which has the following trivial mass function:

$$\mathrm{m}(X) = \begin{cases} 1 \text{ if } X = \Omega \\ 0 \text{ otherwise} \end{cases}$$

If we add basic belief formula $b_1 : p_1$, we compute the following mass function:

$$m_{b_1}(X) = \begin{cases} p_1 & \text{if } X = \max_\Omega(b_1) \\ 1 - p_1 & \text{if } X = \Omega \\ 0 & \text{otherwise} \end{cases}$$

Adding $b_2 : p_2$ leads to[4]:

[4] Note that this general scheme is based on the assumption that $b_1 \wedge b_2$ is not equivalent to b_1 or to b_2, i.e. it is assumed that the belief formulae are such that their pairwise intersections are not equivalent with one of the constituents. When this assumption

$$m_{b_1} \oplus m_{b_2}(X) = m(X) = \begin{cases} p_1 \cdot (1-p_2) & \text{if } X = \max_\Omega(b_1) \\ p_2 \cdot (1-p_1) & \text{if } X = \max_\Omega(b_2) \\ p_1 \cdot p_2 & \text{if } X = \max_\Omega(b_1 \wedge b_2) \\ (1-p_1)\cdot(1-p_2) & \text{if } X = \Omega \\ 0 & \text{otherwise} \end{cases}$$

Note that these consecutive mass combinations can be generalised to a situation with n basic belief formulae, and that the combined mass function will grow exponentially[5]. Fortunately, there is no need to calculate the entire mass function. If we need the mass value for a certain set $X \subseteq \Omega$, we can calculate it using the probabilities in the belief base without the need to calculate the entire mass function. Before formulating and proving this theorem, we will first show that we can simplify Dempster's Rule of Combination if 1) we use simple support functions to represent basic belief formulae, and 2) we define Ω to be such that each conjunction of basic belief formulae that the agent can hold during its execution maps to a non-empty subset of Ω, as we discussed in section 3.5. The two demands are intuitive: the first demand states, that the evidence that is represented by each basic belief formula only supports one subset of Ω, and the second guarantees that the conjunction of basic belief formulae has a model, i.e. updating the belief base always results in a consistent belief base.

To facilitate further considerations, we define p_φ as the probability assigned to a certain belief formula φ, and define m_φ to be a simple support function associated with φ. The core of m_φ is denoted as $C(m_\varphi)$. We introduce the concept *M-completeness* to denote the mentioned condition on Ω, and define it as follows:

Definition 5. *(M-complete). Let Ω be a set of hypotheses and let M be a set of mass functions. Then Ω will be called M-complete if and only if $\forall m_\phi, m_\psi \in M$ $(max_\Omega(\phi) \in C(m_\phi) \;\;\&\;\; max_\Omega(\psi) \in C(m_\psi)) \Rightarrow max_\Omega(\phi \wedge \psi) \subseteq \Omega$*

With this notion of M-completeness, we can formulate and prove how Dempster's Rule of Combination can be simplified.

Theorem 1. *Let S_Ω be the set of all basic belief formulae, and M be the set of all mass functions associated with basic belief formulae from S_Ω. Let Ω be M-complete, and let ϕ and ψ be two non-equivalent basic belief formulae (i.e. $\neg(\phi \equiv \psi)$). Then $\sum_{Y \cap Z \neq \emptyset} m_\phi(Y) \cdot m_\psi(Z) = 1$, and for each $X \subseteq \Omega$ there are at most one $Y \subseteq \Omega$ and one $Z \subseteq \Omega$ relevant for the Dempster's combination rule such that this rule can be simplified to $m_\phi \oplus m_\psi(X) = m_\phi(Y) \cdot m_\psi(Z)$ where $Y \cap Z = X$.*

would not hold, some of the clauses would be identical, and there would be less clauses in the remaining scheme.

[5] In fact, Orponen [11] showed that Dempster's Rule of Combination is #P-complete. Wilson [19] has provided a number of approximation algorithms to overcome this problem, and Barnett [1, 2] has shown, that the calculation of the combination is linear if only singleton subsets are used or if the subsets are atomic with respect to the evidence. The latter restriction is problematic since we are not aware of the actual hypotheses that constitute the frame of discernment.

Proof. Since ϕ and ψ are two basic belief formulae, the only $Y, Z \subseteq \Omega$ for which $m_\phi(Y) \neq 0$ and $m_\psi(Z) \neq 0$ are focal elements of m_ϕ and m_ψ, i.e. $C(m_\phi) = \{M_Y, \Omega\}$ where $M_Y = \max_\Omega(\phi)$ and $C(m_\psi) = \{M_Z, \Omega\}$ where $M_Z = \max_\Omega \psi$. In other words, Y ranges over $C(m_\phi)$ and Z ranges over $C(m_\psi)$. For all other subsets Y and Z from Ω, we have $m_\phi(Y) = 0$ and $m_\psi(Z) = 0$ such that $m_\phi(Y) \cdot m_\psi(Z) = 0$ which does not influence the summation $\sum_{Y \cap Z = X} m_\phi(Y) \cdot m_\psi(Z)$ in the numerator of the Combination Rule. Given that Y and Z range over $C(m_\phi)$ and $C(m_\psi)$ respectively, it is clear that for each $X \subseteq \Omega$ we can have at most one $Y \in C(m_\phi)$ for which $m_\phi(Y) \neq 0$ and at most one $Z \in C(m_\psi)$ for which $m_\psi(Z) \neq 0$ such that $Y \cap Z = X$. More specifically, for any $X \subseteq \Omega$, the subsets Y and Z can be determined as follows:

If $X = \Omega$ then $Y = \Omega$ and $Z = \Omega$.
If $X = \max_\Omega(\phi)$, then $Y = X$ and $Z = \Omega$.
If $X = \max_\Omega(\psi)$, then $Y = \Omega$ and $Z = X$.
If $X = \max_\Omega(\phi \wedge \psi)$, then $Y = \max_\Omega(\phi)$ and $Z = \max_\Omega(\psi)$.

Since m_ϕ and m_ψ are simple support functions, $m_\phi(\max_\Omega(\phi)) = p_\phi, m_\phi(\Omega) = 1 - p_\phi, m_\psi(\max_\Omega(\psi)) = p_\psi$, and $m_\psi(\Omega) = 1 - p_\psi$. Then, we have $\sum_{Y \cap Z \neq \emptyset} m_\phi(Y) \cdot m_\psi(Z) = p_\phi \cdot p_\psi + p_\phi \cdot (1 - p_\psi) + p_\psi \cdot (1 - p_\phi) + (1 - p_\phi) \cdot (1 - p_\psi) = 1$. This proves that the denominator does not influence the result of the Combination Rule. Therefore, Dempster's Rule of Combination can be simplified to $m_\phi \oplus m_\psi(X) = m_\phi(Y) \cdot m_\psi(Z)$ where $Y \cap Z = X$. □

This result can be generalised for consecutive combinations. From the associativity of Dempster's Rule of Combination[13] follows, that $Y_1, \ldots, Y_n$ in the formula $Y_1 \cap \ldots \cap Y_n = X$ can be determined in a similar way. For example, in a belief base consisting of three basic belief formulae ϕ, ψ, and χ, the set $X = \max_\Omega(\psi \wedge \psi)$ can be determined as intersection of $Y_\phi = \max_\Omega(\phi)$, $Y_\psi = \max_\Omega(\psi)$ and $Y_\chi = \Omega$. In the second theorem we formulate a straightforward method to calculate the mass of any combination of basic belief formula, and prove that this calculation leads to the same result as the simplified Rule of Combination.

Theorem 2. *Let S_Ω be the set of all basic belief formulae, and M be the set of all mass functions associated with basic belief formulae from S_Ω. Let Ω be M-complete. For each subset $X \subseteq \Omega$, there exists an unique bi-partition of the set of basic belief formulae S_Ω, say S_X^+ and S_X^-, such that the general case of Dempster's combination rule can be simplified as follows:*

$$\bigoplus_{i=1\ldots n} m_i(X) = \prod_{\varphi \in S_X^+} p_\varphi \cdot \prod_{\varphi \in S_X^-} (1 - p_\varphi)$$

Proof. Let Y and Z be subsets of Ω. Based on theorem 1, Dempster's Rule of Combination can be reduced to $m_1 \oplus m_2(X) = m_1(Y) \cdot m_2(Z)$, where $Y \cap Z = X$. The mass function that is formed by n consecutive combinations, is then equal to

$$\forall X, Y_1, \ldots, Y_n \subseteq \Omega \;:\; \bigoplus_{i=1\ldots n} m_i(X) = \prod_{i=1\ldots n} m_i(Y_i), \quad \text{where} \bigcap_{i=1\ldots n} Y_i = X$$

Given the mass functions $m_{\phi_1}, \ldots, m_{\phi_n} \in M$ and for any $X \subseteq \Omega$ and $1 \leq i \leq n$, there exists at most one $Y_i \subseteq \Omega$ ranging over the core $C(m_{\phi_i})$ such that $Y_1 \cap \ldots \cap Y_n = X$ and $m_{\phi_i}(Y_i) \neq 0$ (for the same reason as in theorem 1). According to the definition of the simple support function, $m_{\phi_i}(Y_i)$ can be either p_{ϕ_i} or $1 - p_{\phi_i}$ for $1 \leq i \leq n$. Let then $S_X^+ = \{\phi \mid Y_i = \max_\Omega(\phi) \;\&\; \phi \in S_\Omega\}$ which is the set of all basic belief formula for which the corresponding mass function assigns p_{ϕ_i} to the subset Y_i (rather than $1 - p_{\phi_i}$). Using $S_X^- = S_\Omega \backslash S_X^+$ proves the theorem. □

5 Updating and querying the belief base

In order to incorporate new information (e.g. by observation or communication) in their beliefs, agents need to update their belief base. Since both an existing belief base (consisting of basic belief formulae) and a new basic belief formulae can both be represented by mass functions, we can add this new basic belief formula to the belief base which in its turn can be represented by a mass function. This would yield the same result as combining each single support function associated with the basic belief formulae, as we proved in theorem 2. However, if the belief base already contains this belief formula, we can update it using the associative nature of Dempster's Rule of Combination. For example, suppose a belief base, which consists of two basic belief formulae $b_1 : p_1$ and $b_2 : p_2$, is updated with the basic belief formula $b_1 : p_3$. The new probability of b_1 can be calculated since $m_{b_1} \oplus m_{b_2} \oplus m'_{b_1} = m_{b_1} \oplus m'_{b_1} \oplus m_{b_2} = (m_{b_1} \oplus m'_{b_1}) \oplus m_{b_2}$. Combining m_{b_1} and m'_{b_1} leads to a single support function with a mass value of $p_1 + p_3 - p_1 \cdot p_3$ for the set $X = \max_\Omega(b_1)$, therefore we can update the probability of b_1 in the belief base to $p_1 + p_3 - p_1 \cdot p_3$.

Furthermore, we can test (query) if a proposition φ can be derived from a belief base Γ. In section 2, we discussed the belief and plausibility functions (defined in terms of a certain mass function) that return the total mass assigned to models of φ and the total mass that is *not* assigned to models of the *negation* of φ. Using this functions, we can test if φ can be derived from Γ within a certain probability interval $[L, U]$ (denoted as $\Gamma \models_{[L,U]} \varphi$). This can be done by checking if $Bel(\max_\Omega(\varphi)) \geq L$ and $Pl(\max_\Omega(\varphi)) \leq U$, since Bel and Pl indicate the lower and upper bound probability, respectively. Note that Bel and Pl are defined in terms of a mass function (definition 3). When we consider the belief base Γ, the mass function used in Bel and Pl will represent the belief base. The mass function is then denoted by m_Γ. As discussed in section 3, the mass function m_Γ assigns a mass value to each subset of the frame of discernment related to the belief base Γ. Therefore, the test if φ can be deducted from Γ within $[L, U]$ can be formulated as follows: $\Gamma \models_{[L,U]} \varphi \iff Bel(\max_\Omega(\varphi)) \;\&\; Pl(\max_\Omega(\varphi)) \iff \sum_{Y \subseteq \max_\Omega(\varphi)} m_\Gamma(Y) \geq L \;\;\&\;\; \sum_{Y \cap \max_\Omega(\varphi) \neq \emptyset} m_\Gamma(Y) \leq U$

6 Complexity of belief queries

We can calculate $Bel(\max_{\Omega}(\varphi))$ by adding all values for $m(X)$, where $X \subseteq \Omega$ and models(X, φ). Moreover, $Pl(\max_{\Omega}(\varphi))$ can be calculated in a similar way by adding all values for $m(X)$, where $X \subseteq \Omega$ and $\neg$models$(X, \neg\varphi)$). Without restrictions on the basic belief formulae in the belief base, we need to iterate over all focal elements of m, which suggests an exponential computational complexity for the determination whether $\Gamma \models_{[L,U]} \varphi$. However, if deduction is based on the Closed World Assumption, then $Bel(X) = Pl(X)$, and, furthermore, if we restrict the logical formulae that constitute the basic belief formulae in the belief base to be prolog facts (i.e. atoms) then the computational complexity of $Bel(\max_{\Omega}(\varphi))$ is linear in the length of φ.

Theorem 3. *Let $Bel(X)$ and $Pl(X)$ be Dempster-Shafer belief and plausibility functions, respectively, defined on a certain mass function m. If the deduction $X \models \varphi$ is based on the Closed World Assumption, then we have $Bel(X) = Pl(X)$.*

Proof. In the Closed World Assumption, we can test the belief in a certain formula φ by calculating $Bel(\max_{\Omega}(\varphi))$, and the belief in the negation of this formula by calculating $Bel(\Omega \backslash \max_{\Omega}(\varphi))$. Since $Bel(X)$ is defined as $\sum_{Y \subseteq X} m(Y)$, then $Bel(\Omega \backslash X) = \sum_{Y \not\subseteq X} m(Y)$ in the CWA. Since $\sum_{Y \subseteq X} m(Y) + \sum_{Y \not\subseteq X} m(Y) = \sum_{X \subseteq \Omega} m(X) = 1$, it follows that $Bel(\Omega \backslash X) = 1 - Bel(X)$. By definition[6], $Pl(X) = 1 - Bel(\Omega \backslash X)$: the plausibility of a set of hypotheses is 1 minus the belief in the complement (with respect to Ω) of this set. But $Bel(\Omega \backslash X) = 1 - Bel(X)$, and therefore $Pl(X) = Bel(X)$ under the Closed World Assumption. □

The belief in a certain formula φ can be calculated straightforward by using standard probability calculus in linear time with respect to the number of atoms and connectives in φ.

Theorem 4. *Let the belief base be a set of atomic formulae, to which a probability value is assigned. If $\phi : p_\phi$ and $\psi : p_\psi$ are basic belief formulae in the belief base, then $Bel(max_{\Omega}(\phi))$, $Bel(max_{\Omega}(\neg\phi))$, $Bel(max_{\Omega}(\phi \wedge \psi))$, and $Bel(max_{\Omega}(\phi \vee \psi))$ can be calculated as follows.*

- $Bel(max_{\Omega}(\phi)) = p_\phi$
- $Bel(max_{\Omega}(\neg\phi)) = Bel(\Omega \backslash max_{\Omega}(\phi)) = 1 - p_\phi$
- $Bel(max_{\Omega}(\phi \wedge \psi)) = p_\phi \cdot p_\psi$
- $Bel(max_{\Omega}(\phi \vee \psi)) = p_\phi + p_\psi - p_\phi \cdot p_\psi$

Proof. The proof of the first clause is as follows: $Bel(\max_{\Omega}(\varphi))$ is defined as the sum of all mass values that are defined for a certain subset of hypotheses in Ω that are models of φ. If the belief base consists of exactly the basic belief formula $\phi : p_\phi$, then the first clause is proved by the definition of mass function, i.e.,

[6] See section 2

$$m_\phi(X) = \begin{cases} p_\phi & \text{if } X = \max_\Omega(\phi) \\ 1 - p_\phi & \text{if } X = \Omega \\ 0 & \text{otherwise} \end{cases}$$

If the belief base is, however, a combination of two mass functions, say m_ϕ and m_χ, then $\max_\Omega(\phi)$ and $\max_\Omega(\phi \wedge \chi)$ are the only focal elements of $m_\phi \oplus m_\chi$ that consist of models of ϕ. Their mass values are then respectively $p_\phi \cdot p_\chi$ and $p_\phi \cdot 1 - p_\chi$ that sum up to p_ϕ. The second clause follows by definition from the first, using $Bel(\max_\Omega(\phi)) + Bel(\max_\Omega(\neg\phi)) = 1$. The proof of the third clause follows the lines of the proof of the first clause. When combining mass functions related to m_ϕ and m_ψ, the models of $\phi \wedge \psi$ are constructed by intersecting the maximal set of hypotheses in Ω that are models of ϕ and ψ, and from Dempster's Rule of Combination follows that the mass of this intersection equals $p_1 \cdot p_2$. The fourth clause follows from the second and third clause. □

7 Conclusion and further work

A lot of research has been conducted on the topic of reasoning with uncertainty. Many approaches are based on extending epistemic logic with probability. For example, [7] proposed the system AX_{MEAS}, [8] introduced the $P_F D$ system, and [16] further refined this system to $P_F KD45$. Some of these logics are suggested to be a good candidate to be used as an underlying system for agent programming (see for example [6]). However, next to the epistemic logic approach, alternative notions of uncertainty are suggested, like the Certainty Factor model used in MYCIN, Bayesian (or causal) networks, and the Dempster-Shafer theory of evidence. Particularly appealing in the latter is the ability to model *ignorance* as well as uncertainty, the presence of a combination rule to combine evidence, and the concept of hypotheses which can be easily related to models of logical formulae. Nevertheless, the compuational complexity, the issue of inconsistency, and the logical validity of the combination rule (see for example [18] for a discussion) are serious disadvantages of this theory for the practical application of this theory to agent programming.

We have investigated a possible mapping of Dempster-Shafer sets to belief formulae, which are represented by logical formulae, in agent programming languages. We have shown that, with restrictions on the mass functions and on the frame of discernment, Dempster-Shafer theory is a convenient way to model uncertainty in agent beliefs, and these disadvantages can be overcome. Because we do not need to keep a combined mass function of n beliefs in memory and update it with each belief update (but compute the mass value of a particular subset of Ω based on the beliefs in the belief base) there is no combinational explosion. Currently, we are working on an implementation of uncertain beliefs in the agent programming language 3APL. Further research will be conducted on the consequences of uncertain beliefs to agent deliberation.

References

1. J.A. Barnett. Computational methods for a mathematical theory of evidence. In *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pages 868–875, 1981.
2. J.A. Barnett. Calculating Dempster-Shafer plausibility. *IEEE transactions on pattern analysis and machine intelligence*, 13:599–603, 1991.
3. L. Braubach, A. Pokahr, and W. Lamersdorf. Jadex: A short overview. In *Main Conference Net.ObjectDays 2004*, pages 195–207, september 2004.
4. P. Busetta, R. Ronnquist, A. Hodgson, and A. Lucas. Jack intelligent agents - components for intelligent agents in java. Technical report, 1999.
5. M. Dastani, B. van Riemsdijk, F. Dignum, and J.-J. Meyer. A programming language for cognitive agents : Goal directed 3APL. In *Proceedings of the First Workshop on Programming Multiagent Systems (ProMAS03)*, 2003.
6. N. de C. Ferreira, M. Fisher, and W. van der Hoek. A simple logic for reasoning about uncertainty. In *Proceedings of the ESSLLI'04 Student Session*, pages 61–71, 2004.
7. R. Fagin and J.Y. Halpern. Reasoning about knowledge and probablility. *Journal of the ACM*, 41:340–367, 1994.
8. M. Fattorosi-Barnaba and G. Amati. Studio logica. 46:383–393, 1987.
9. A. Jøsang. The consensus operator for combining beliefs. *Artificial Intelligence Journal*, 142(1-2):157–170, 2002.
10. A.F. Moreira and R.H. Bordini. An operational semantics for a BDI agent-oriented programming language. In J.-J. C. Meyer and M. J. Wooldridge, editors, *Proceedings of the Workshop on Logics for Agent-Based Systems*, pages 45–59, 2002.
11. P. Orponen. Dempster's rule of combination is # P-complete. *Artificial Intelligence*, 44:245–253, 1990.
12. A.J. Rao and M.P. Georgeff. Modelling rational agents within a BDI-architecture. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning.* Morgan Kaufmann Publishers, San Mateo, CA, 1991.
13. K. Sentz. *Combination of Evidence in Dempster-Shafer Theory.* PhD thesis, Binghamton University, 2002.
14. G. Shafer. *A mathematical theory of evidence.* Princeton Univ. Press, Princeton, NJ, 1976.
15. P. Smets. The combination of evidence in the transferable belief model. *IEEE Pattern Analysis and Machine Intelligence*, 12:447–458, 1990.
16. W. van der Hoek. Some considerations on the logic PFD. *Journal of Applied Non Classical Logics*, 7:287–307, 1997.
17. M. Verbeek. 3APL as programming language for cognitive robotics. Master's thesis, Utrecht University, 2003.
18. F. Voorbraak. *As Far as I know - Epistemic Logic and Uncertainty.* PhD thesis, 1993.
19. N. Wilson. Algorithms for dempster-shafer theory. In D.M. Gabbay and P. Smets, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 5: Algorithms, pages 421–475. Kluwer Academic Publishers, 2000.
20. M. Wooldridge. Intelligent agents. In G. Weiss, editor, *Multiagent Systems.* The MIT Press, 1999.
21. R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41:93–137, 1987.