

Modelling Uncertainty in Agent Programming

Johan Kwisthout and Mehdi Dastani

ICS, Utrecht University

j.kwisthout@student.kun.nl , 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, using simple support functions as a representation for the agent's beliefs, the computational complexity of Dempster's Rule of Combination can be controlled. In particular, the certainty value of a proposition can be deduced 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, using 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 [4, 9, 11, 3, 2]. 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, which 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 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 beliefs, 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 the next section we propose a mapping between this theory, and agent beliefs. In section four, we deal with implementational issues, and show that computational complexity can be controlled given certain restrictions on the belief representation. In section five we show how the agent's belief base can be updated and queried. Finally, in section six we conclude our 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), probability 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 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 a mass value to this

particular subset of Ω (say 0.6) 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:

$$m_1(X) = \begin{cases} 0.6 & \text{if } X = \{\text{Peter, John, Paul}\} \\ 0.4 & \text{if } X = \Omega \\ 0.0 & \text{otherwise} \end{cases}$$

Note that no value whatsoever is assigned to *subsets* of $\{\text{Peter, John, Paul}\}$. 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:

$$m_2(X) = \begin{cases} 0.9 & \text{if } X = \{\text{John, Mary}\} \\ 0.1 & \text{if } X = \Omega \\ 0.0 & \text{otherwise} \end{cases}$$

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

$$m_1 \oplus m_2(X) = \frac{\sum_{Y \cap Z = X} m_1(Y) \cdot m_2(Z)}{\sum_{Y \cap Z \neq \emptyset} m_1(Y) \cdot m_2(Z)} \quad \text{and}$$

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

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:

$$m_1 \oplus m_2(X) = \begin{cases} 0.06 & \text{if } X = \{\text{Peter, John, Paul}\} \\ 0.36 & \text{if } X = \{\text{John, Mary}\} \\ 0.54 & \text{if } X = \{\text{John}\} \\ 0.04 & \text{if } X = \Omega \\ 0.00 & \text{otherwise} \end{cases}$$

A *simple support function* is a special case of a mass function, where the evidence only supports A , and no non-zero mass value is assigned to any subset of Ω other than A :

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:

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

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 the mass function as follows:

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

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 $Pl(X) = 1 - Bel(\Omega \setminus X)$. 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 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 n -by- m 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. Assume that, at a given moment during the execution of the program, the agent has the following belief base in a 2-by-2 grid-world:

$BB_1 : safe(1)$

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 denoting 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

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

can be represented by a mass function. Such a mass function would assign 1 to the subset of Ω which contains all hypotheses that are true with respect to the belief base. In general, if a belief base is a certain belief formula φ , then it could be represented by a mass function $m_\varphi(X)$, where

$$m_\varphi(X) = \begin{cases} 1.0 & \text{if } X \subseteq \Omega \text{ \& models}(X, \varphi) \text{ \& } \forall Y \subseteq \Omega (\text{models}(Y, \varphi) \Rightarrow Y \subseteq X) \\ 0.0 & \text{otherwise} \end{cases}$$

In this definition, the relation $\text{models}(X, \varphi)$ 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 that are models of φ . In the sequel, we use $X \models^\Omega \varphi$ as a shorthand for the complex condition $X \subseteq \Omega \text{ \& models}(X, \varphi) \text{ \& } \forall Y \subseteq \Omega (\text{models}(Y, \varphi) \Rightarrow Y \subseteq X)$. Using this shorthand notation, the mass function that represents the belief base φ can be rewritten as:

$$m_\varphi(X) = \begin{cases} 1.0 & \text{if } X \models^\Omega \varphi \\ 0.0 & \text{otherwise} \end{cases}$$

The belief base BB_1 can then be represented by :

$$m_{\text{safe}(1)}(X) = \begin{cases} 1.0 & \text{if } X \models^\Omega \text{safe}(1) \\ 0.0 & \text{otherwise} \end{cases}$$

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 $\text{safe}(2)$ to the belief base, with the following mass function:

$$m_{\text{safe}(2)}(X) = \begin{cases} 1.0 & \text{if } X \models^\Omega \text{safe}(2) \\ 0.0 & \text{otherwise} \end{cases}$$

We can combine both pieces of evidence using Dempster's rule of combination. Since the only non-empty intersection of sets defined by either $m_{\text{safe}(1)}$ or $m_{\text{safe}(2)}$ is the subset $\{X \mid X \models^\Omega \text{safe}(1) \wedge \text{safe}(2)\}$, the resulting mass func-

tion $m_1 = m_{safe(1)} \oplus m_{safe(2)}$ is defined as follows¹:

$$m_1(X) = \begin{cases} 1.0 & \text{if } X \models^\Omega \text{safe}(1) \wedge \text{safe}(2) \\ 0.0 & \text{otherwise} \end{cases}$$

Note that X 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 certain evidence, that a location *probably* contains a bomb, so that the belief formula $bomb(3)$, with a probability value of 0.7, is 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, and define it as follows:

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 $\{ \text{safe}(1): 1.0, \text{safe}(2): 1.0, \text{bomb}(3): 0.7 \}$. Of course, we could represent $\text{bomb}(3): 0.7$ as a mass function, as we did with beliefs $\text{safe}(1)$ and $\text{safe}(2)$. 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 :

$$m_{bomb(3)}(X) = \begin{cases} 0.7 & \text{if } X \models^\Omega \text{bomb}(3) \\ 0.3 & \text{if } X = \Omega \\ 0.0 & \text{otherwise} \end{cases}$$

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

$$m_2 = m_1 \oplus m_{bomb(3)}(X) = \begin{cases} 0.7 & \text{if } X \models^\Omega \text{safe}(1) \wedge \text{safe}(2) \wedge \text{bomb}(3) \\ 0.3 & \text{if } X \models^\Omega \text{safe}(1) \wedge \text{safe}(2) \\ 0.0 & \text{otherwise} \end{cases}$$

The mass function m_2 represents our updated belief base. Note that $X \models^\Omega \text{safe}(1) \wedge \text{safe}(2)$ is exactly the set $\{H_1, H_2, H_3, H_4\}$, and $X \models^\Omega \text{safe}(1) \wedge \text{safe}(2) \wedge \text{bomb}(3)$ is the set $\{H_3, H_4\}$.

3.2 Deleting beliefs

We can also delete belief formulae from our belief base. For example, we could conclude $\neg \text{bomb}(3)$ during the execution of our program. We will model deletion of a formula as the addition of its negation. This corresponds to the maximal

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 constitute the basic belief formula.

set of hypotheses according to which there is *no* bomb on location 3:

$$m_{\neg bomb(3)}(X) = \begin{cases} 1.0 & \text{if } X \models^{\Omega} \neg bomb(3) \\ 0.0 & \text{otherwise} \end{cases}$$

Combining m_2 and $m_{\neg bomb(3)}$ leads to the following mass function:

$$m_3 = m_2 \oplus m_{\neg bomb(3)}(X) = \begin{cases} 1.0 & \text{if } X \models^{\Omega} save(1) \wedge \neg bomb(3) \\ 0.0 & \text{otherwise} \end{cases}$$

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 $\neg bomb(3)$: 0.3. We would represent this formula as:

$$m_{\neg bomb(3)}(X) = \begin{cases} 0.3 & \text{if } X \models^{\Omega} \neg bomb(3) \\ 0.7 & \text{if } X = \Omega \\ 0.0 & \text{otherwise} \end{cases}$$

Combining this alternative mass function $m_{\neg bomb(3)}$ and m_2 leads to the following mass function:

$$m_4 = m_2 \oplus m_{\neg bomb(3)}(X) = \begin{cases} 0.59 & \text{if } X \models^{\Omega} save(1) \wedge \neg bomb(3) \\ 0.41 & \text{if } X \models^{\Omega} save(1) \wedge bomb(3) \\ 0.00 & \text{otherwise} \end{cases}$$

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. This is illustrated in table 2. We can illustrate composite beliefs with an example. Consider the following two formulae in the already mentioned grid-world:

1. $safe(2) \wedge (safe(3) \vee safe(4))$
2. $safe(1) \vee (\neg safe(2) \wedge safe(3))$

These formulae correspond to the the following sets, respectively:

1. the set $\{1, 2, 3, 9, 10, 11\}$
2. the set $\{1, 2, 3, 4, 5, 6, 7, 8, 13, 14\}$

If formula 1 has a probability of p , and formula 2 has a probability of q , then these formula could be represented by basic belief formulae as follows:

belief formula	Dempster-Shafer
safe(X)	set of possible worlds (hypotheses) in which safe(X) is true
safe(X) \vee safe(Y)	union of the set of possible worlds (hypotheses) in which safe(X) is true and the set in which safe(Y) is true
safe(X) \wedge safe(Y)	intersection of the set of possible worlds (hypotheses) in which safe(X) is true and the set in which safe(Y) is true
\neg safe(X)	complement of the set of possible worlds (hypotheses) in which safe(X) is true

Table 2. Logical connectives and Dempster-Shafer sets of hypotheses

$$m_1(X) = \begin{cases} p & \text{if } \models^\Omega \text{safe}(2) \wedge (\text{safe}(3) \vee \text{safe}(4)) \\ 1 - p & \text{otherwise} \end{cases} \quad \text{and}$$

$$m_2(X) = \begin{cases} q & \text{if } \models^\Omega \text{safe}(1) \vee (\neg \text{safe}(2) \wedge \text{safe}(3)) \\ 1 - q & \text{otherwise} \end{cases}$$

Obviously, the conjunction of these two statements is:

$$\text{safe}(1) \wedge \text{safe}(2) \wedge (\text{safe}(3) \vee \text{safe}(4))$$

And from the table follows, that this result corresponds to the set $\{1, 2, 3\}$, which is the intersection of $\{1, 2, 3, 9, 10, 11\}$ and $\{1, 2, 3, 4, 5, 6, 7, 8, 13, 14\}$. Therefore, using Dempster's Rule of Combination leads to the following mass function:

$$m_1 \oplus m_2(X) = \begin{cases} p \cdot q & \text{if } X \models^\Omega \text{safe}(1) \wedge \text{safe}(2) \wedge (\text{safe}(3) \vee \text{safe}(4)) \\ p \cdot (1 - q) & \text{if } X \models^\Omega \text{safe}(2) \wedge (\text{safe}(3) \vee \text{safe}(4)) \\ (1 - p) \cdot q & \text{if } X \models^\Omega \text{safe}(1) \vee (\neg \text{safe}(2) \wedge \text{safe}(3)) \\ (1 - p) \cdot (1 - q) & \text{if } X = \Omega \\ 0.00 & \text{otherwise} \end{cases}$$

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 [8]):

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:

$$m_1(X) = \begin{cases} 0.99 & \text{if } X = \text{'killer is Peter'} \\ 0.01 & \text{if } X = \text{'killer is Paul'} \\ 0.00 & \text{otherwise} \end{cases}$$

$$m_2(X) = \begin{cases} 0.99 & \text{if } X = \text{'killer is Mary' } \\ 0.01 & \text{if } X = \text{'killer is Paul' } \\ 0.00 & \text{otherwise} \end{cases}$$

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.

$$m_1 \oplus m_2(X) \begin{cases} 1.00 & \text{if } X = \text{'killer is Paul' } \\ 0.00 & \text{otherwise} \end{cases}$$

Sentz [13] describes a number of alternatives for this 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:

$$m_1 \oplus m_2(X) \begin{cases} 0.0001 & \text{if } X = \text{'killer is Paul' } \\ 0.9999 & \text{if } X = \Omega \\ 0.0000 & \text{otherwise} \end{cases} \quad (\text{Yager}), \text{ respectively}$$

$$m_1 \oplus m_2(X) \begin{cases} 0.0001 & \text{if } X = \text{'killer is Paul' } \\ 0.9999 & \text{if } X = \emptyset \\ 0.0000 & \text{otherwise} \end{cases} \quad (\text{Smets})$$

Jøsang [8] 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:

$$m_1 \oplus m_2(X) \begin{cases} 0.495 & X = \text{'killer is Peter' } \\ 0.495 & X = \text{'killer is Mary' } \\ 0.010 & X = \text{'killer is Paul' } \\ 0.000 & X = \emptyset \end{cases}$$

We can summarize these approaches (Dempster, Yager/Smets and Jøsang) using the 'murder case' example, as shown in table 3:

Notice, 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 $\{\text{Paul, Peter, Mary}\}$.

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	0.00	0.00	0.00	0.495
Paul	0.01	0.01	1.00	0.0001	0.01
Mary	0.00	0.99	0.00	0.00	0.495
\emptyset/Ω	0.00	0.00	0.00	0.9999	0.00

Table 3. Attribution of mass to inconsistent combinations

If, on the other hand, our frame would be $\Omega = \{\text{Paul, Peter, Mary, Peter or Mary}\}$, and we would map the phrase 'killer is Peter' to the subset $\{\text{Peter, Peter or Mary}\}$ and the phrase 'killer is Mary' to the subset $\{\text{Mary, Peter or Mary}\}$, 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 our 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 during the execution of the program. This implies that the relation between frame of discernment and beliefs is a many-to-many mapping.

This problem can, however, be solved. According to some agent programming languages such as 3APL [4], 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 nonempty subset of hypotheses in Ω . In

other words, each conjunction of basic belief formulae has a nonempty 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:

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

The combination of belief formulae $b_1 \dots b_n$ can thus be represented with a mass function $m_k = m_1 \oplus \dots \oplus m_n$, related to beliefs b_1, \dots, b_n , where the number of subsets of Ω that are used to define m_k and have a mass value > 0 is equal to 2^n . When a belief formula m_n is combined with an already existing combination $m_1 \oplus \dots \oplus m_{n-1}$, the resulting mass function $m_1 \oplus \dots \oplus m_n$ is defined by the non-empty intersections of all subsets of Ω in m_n , with all subsets of Ω in $m_1 \oplus \dots \oplus m_{n-1}$. Since we use simple support functions to represent our beliefs, 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 Orponen [10] showed, Dempsters's Rule of Combination is #p-complete. Wilson [19] has provided a number of algorithms to overcome this problem, and Barnett [1] has shown, that the calculation of the combination is linear if only singleton subsets are used (which effectively reduces the Dempster-Shafer theory to standard probability calculation). But, as we will see, there is no need to update - or, for that matter, even calculate - the entire mass function, because we only need to combine simple support functions as presented in definition 2.

Of course, an empty belief base has a trivial mass function of:

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

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

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

Adding $b_2 : p_2$ leads to:

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

Note that these consecutive mass combinations can be generalized to a situation with n basic belief formulae, and that the combined mass function will grow exponentially. Fortunately, there is no need to calculate the entire mass function. If we need the mass value for a certain subset X of Ω , 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 we use simple support functions to represent basic belief formulae, and if we define Ω to be such that each conjunction of basic belief formulae that the agent can hold during its execution maps to a nonempty 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.

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 . Let the domain $D(m_\varphi)$ of a simple support function m_φ be defined as the set $\{X \mid X \models^\Omega \varphi \text{ or } X = \Omega\}$. This domain consists of the only two subsets of the frame of discernment that have a non-zero probability. We introduce the concept *M-completeness* to denote the mentioned condition on Ω , and define it as follows:

Definition 5. (*M-complete*). Let M be a set of mass functions and Ω be a set of hypotheses, then Ω will be called *M-complete* if and only if $\forall m_\phi, m_\psi \in M_\Omega(\{X \mid X \models^\Omega \phi\} \in D(m_\phi) \ \& \ \{X \mid X \models^\Omega \psi\} \in D(m_\psi) \Rightarrow \{X \mid X \models^\Omega \phi \wedge \psi\} \in \Omega)$

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 $Y \cap Z = X$.

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 from the domain of m_ϕ and m_ψ , i.e. $D(m_\phi) = \{M_Y, \Omega\}$ where $M_Y \models^\Omega \phi$ and $D(m_\psi) = \{M_Z, \Omega\}$ where $M_Z \models^\Omega \psi$. In other words, Y ranges over $D(m_\phi)$ and Z ranges over $D(m_\psi)$. 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 Dempster's combination rule. Given that Y and Z range over $D(m_\phi)$ and $D(m_\psi)$ respectively, it is clear that for each $X \subseteq \Omega$ we can have at most one $Y \in D(m_\phi)$ for which $m_\phi(Y) \neq 0$ and at most one $Z \in D(m_\psi)$ 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 $X \models^\Omega \phi$, then $Y = X$ and $Z = \Omega$.
 If $X \models^\Omega \psi$, then $Y = \Omega$ and $Z = X$.
 If $X \models^\Omega \psi \wedge \phi$, then $Y \models^\Omega \phi$ and $Z \models^\Omega \psi$.

For all other $X \subseteq \Omega$, there are no Y and Z . From our definition of Ω follows, that for these Y and Z we have $Y \cap Z \neq \emptyset$, and since we use simple support structures $\sum_{Y \cap Z \neq \emptyset} m_\phi(Y) \cdot m_\psi(Z) = 1$. 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 $Y \cap Z = X$. \square

From the associativity of Dempster's Rule of Combination[13] follows, that Y_1, \dots, Y_n in the formula $Y_1 \cap \dots \cap Y_n = X$ can be determined in a similar way. For example, in a belief base consisting of three basic belief formulae ϕ , ψ , and χ , then where $X \models^\Omega \psi \wedge \phi$, $Y_\phi \models^\Omega \phi$, $Y_\psi \models^\Omega \psi$ 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:*

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

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

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

Given the mass functions $m_{\phi_1}, \dots, m_{\phi_n} \in M_\Omega$ and for any $X \subseteq \Omega$ and $1 \leq i \leq n$, there exists at most one $Y_i \subseteq \Omega$ ranging over the domain $D(m_{\phi_i})$ such that $Y_1 \cap \dots \cap Y_n = X$ 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 $S_X^+ = \{\phi \mid Y_i \models^\Omega \phi \ \& \ \phi \in S_\Omega\}$ 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 $S_X^- = S_\Omega \setminus S_X^+$ proves the theorem. \square

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 $m_{b_1} \oplus m_{b_2} \oplus m'_{b_1} = m_{b_1} \oplus m'_{b_1} \oplus m_{b_2} = (m_{b_1} \oplus m'_{b_1}) \oplus m_{b_2}$. From Dempster's rule of combination follows, that the new probability of b_1 is $p_1 + p_3 - p_1 \cdot p_3$.

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 as follows. As discussed in section 3, the belief base Γ can be represented by a mass function m_Γ that assigns a mass value to each subset of the frame of discernment. As the belief formula φ can be represented as a subset of the frame of discernment, we can consider the above test as computing the lower and upper limits on the probability that the mass function m_Γ assigns to the subset of Ω that represents the belief formula φ , i.e. we can compute the probability that m_Γ assigns to $X \subseteq \Omega$ where $X \models^\Omega \varphi$. Since $Bel(X) = \sum_{Y \subseteq X} m(Y)$ and we have defined the

shorthand $X \models^\Omega \varphi$ as $(X \subseteq \Omega \ \& \ \text{models}(X, \varphi) \ \& \ \forall Y \subseteq \Omega (\text{models}(Y, \varphi) \Rightarrow Y \subseteq X))$, we can define $Bel(X \mid X \models^\Omega \varphi)$ as the sum of all $m(Y)$, where $(Y \subseteq \Omega \ \& \ \text{models}(Y, \varphi))$. Likewise, $Pl(X \mid X \models^\Omega \varphi)$ can be defined as the sum of all $m(Y)$, where $(Y \subseteq \Omega \ \& \ \neg \text{models}(Y, \neg \varphi))$ ⁴. As an example, let's consider the belief base $\{b_1 : 0.7, b_2 : 0.3, b_3 : 0.6\}$, and test whether $b_1 \vee (b_2 \wedge b_3)$ can be deducted based on a Closed World Assumption (i.e. $Bel(X) = Pl(X)$):

⁴ If the deduction of beliefs is based on the Closed World Assumption, $\Gamma \models_{CWA} \varphi \Leftrightarrow \Gamma \not\models_{CWA} \neg \varphi$, and therefore $Bel(X \mid X \models^\Omega \varphi) = Pl(X \mid X \models^\Omega \varphi)$. Consequently, the calculated probability of a certain proposition is a single value, rather than an interval.

To calculate $Bel(X \mid X \models^\Omega b_1 \vee (b_2 \wedge b_3))$ we must add all $m(X)$ that satisfy $(X \subseteq \Omega \ \& \ models(X, \varphi))$, and the only non-zero $m(X)$ that satisfy this condition are $\{X \mid X \models^\Omega b_1\}$, $\{X \mid X \models^\Omega b_1 \wedge b_2\}$, $\{X \mid X \models^\Omega b_1 \wedge b_3\}$, $\{X \mid X \models^\Omega b_1 \wedge b_2 \wedge b_3\}$, and $\{X \mid X \models^\Omega b_2 \wedge b_3\}$. We calculate these mass elements using the formula presented earlier. Note that we only need to calculate five mass elements, instead of the entire mass function, which has eight elements:

$$\begin{aligned} m(X \mid X \models^\Omega b_1) &= 0.196 \\ m(X \mid X \models^\Omega b_1 \wedge b_2) &= 0.084 \\ m(X \mid X \models^\Omega b_1 \wedge b_3) &= 0.294 \\ m(X \mid X \models^\Omega b_1 \wedge b_2 \wedge b_3) &= 0.126 \\ m(X \mid X \models^\Omega b_2 \wedge b_3) &= 0.054 \end{aligned}$$

$Bel(X \mid X \models^\Omega b_1 \vee (b_2 \wedge b_3))$ equals $0.196 + 0.084 + 0.294 + 0.126 + 0.054 = 0.754$. This means the probability of the proposition $b_1 \vee (b_2 \wedge b_3)$ can be derived from our belief base with a certainty value 0.754.

6 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, [6] proposed the system AX_{MEAS} , [7] introduced the P_FD system, and [16] further refined this system to P_FKD45 . Some of these logics are suggested to be a good candidate to be used as an underlying system for agent programming (see for example [5]). 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. Nevertheless, the combinatorial explosion of the combination rule, the issue of inconsistency, and the logical validity of the combination rule (see for example [18]) 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. 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. L. Braubach, A. Pokahr, and W. Lamersdorf. Jadex: A short overview. In *Main Conference Net.ObjectDays 2004*, pages 195–207, september 2004.
3. P. Busetta, R. Ronnquist, A. Hodgson, and A. Lucas. Jack intelligent agents - components for intelligent agents in java. Technical report, 1999.
4. 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.
5. 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.
6. R. Fagin and J.Y. Halpern. Reasoning about knowledge and probability. *Journal of the ACM*, 41:340–367, 1994.
7. M. Fattorosi-Barnaba and G. Amati. Studio logica. 46:383–393, 1987.
8. A. Jøsang. The consensus operator for combining beliefs. *Artificial Intelligence Journal*, 142(1-2):157–170, 2002.
9. 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 (LABS-02), April 22-25, Toulouse, France*, pages 45–59, 2002.
10. P. Orponen. Dempster’s rule of combination is # p-complete. *Artificial Intelligence*, 44:245–253, 1990.
11. L. Padgham and M. Winikoff. *Developing Intelligent Agent Systems: A Practical Guide*. John Wiley and Sons, June 2004.
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.