Possible Histories: A way to model Context-Aware Preferences

Arthur van Bunningen   Maarten Fokkinga

December 22, 2006

Abstract

Nowadays more and more information becomes available in digital form. To be able to guide users through this wealth of information, a possibility is to only provide the user with relevant information, where relevancy is determined by the preferences of the user. To determine the precise relation between relevancy and preferences, we somehow need to formalize both concepts. This paper proposes a way to formalize the preferences of a user by grounding them in possible histories of the user. We explore this technique and its relations to other possible models.

1 Introduction

We are dealing with systems that, upon an implicit or explicit request by a user, produce an output that takes into account the preference of the user. The preferences may depend on the context the user is in, hence the adjective context-aware. Our goal is to model the context-aware preferences a user has in a mathematical way. The mathematical objects that represent preferences are called agents. Past behavior, the current situation, explicit preferences, etc. can all place constraints on the agent. Different agents could model the same person and have different or similar predictions, but we could also imagine groups of persons modeled by a single agent. Therefore, it should be possible to *combine* agents into a resulting agent which predicts in accordance with the agents of which it is a combination. If these agents cannot agree on a integrated prediction, the resulting agent shall not predict anything at all. Since it is not expected that an agent can predict exactly what the user prefers, nor that the user even has these preferences explicitly, we expect the agent to be able to present its beliefs about the preferences of the user by giving multiple options together with a measure of confidence.

In this work we will try to define these agents using probabilities, focussing on how to merge different agents and how to express context-aware preferences; preferences which depend on the situation the user is in. We start our report with the postulates that form the basis for the rest of our discussion in Section 2. We then present two approaches on defining these agents using Dempster-Shafer in Section 3, from which we conclude that we need support for reasoning about ratios between preferences. Therefore we introduce in Section 4 our new model based on possible histories which address this constraint. In Section 5 we explore the relation this model has to probability theory. We conclude our report and summarize open issues in Section 6.

2 Basics

Before we try to model the user we first introduce the notation that we will use throughout this report.

First we define *Pref* as the set of of *oPtions* that the user has. The options that can be described with these properties are subsets of *Pref*. For example, *Pref* = {Jazz, Eighties} indicates that the properties of the options of the user are listening to jazz music and listening to eighties music. Given this *Pref*, listening to jazz music from the eighties is described by the set {Jazz, Eighties} and listening to nothing at all by means of the empty set.

Furthermore, we define *Sit* as the set of properties of *Situations* that the user can be in. Similar to the options, the situations that can be described with these properties are subsets of *Sit*. For example, if we take *Sit* = {Coffeeroom, Happy} we can indicate that the user can be happy and/or/nor in the coffeeroom.

As a running example throughout the report we will use a user that has preferences on jazz and/or/nor eighties music while being happy and/or/nor in the coffeeroom. Therefore, from now on, we will represent Jazz music as J, music of the eighties as E, being in the coffeeroom as C, and being happy as H.

3 Agents based on Dempster-Shafer theory

3.1 Introduction to Dempster-Shafer

Since we want to be able to combine agents having (maybe different) evidences about the preferences of a user, the use of Dempster-Shafer theory (DS) [1] seems a natural choice since it is concerned with the combination of evidences to calculate the probability of an event. In this section we will make two attempts to model the agents and their combination using DS, we will therefore begin by introducing the basic concepts of DS.

First of all, the *frame of discernment* (FoD) defines the hypotheses about which the agent can have beliefs:

Definition 1. A *frame of discernment* is an exhaustive set of mutually exclusive hypotheses about some domain.

Second, the *basic probability assignment* expresses the relative confidence in subsets of the FoD:

Definition 2. Let set D be a frame of discernment, then a function $m : 2^D \rightarrow [0..1]$ is called a *basic probability assignment*, abbreviated as bpa, whenever $m(\emptyset) = 0$ and $\sum_{S \subseteq D} m(S) = 1$

The quantity $m(S)$ expresses a relative confidence in exactly S. The total confidence in S, which we call belief, is the sum of the probability assignments committed to all subsets of S .

Finally, to combine two bpa's m_1 and m_2 into a new bpa one can use *Dempster's* *rule of combination*, which defines the combined bpa $m_{1,2}$ as follows:

$$m_{1,2}(\emptyset) = 0$$
$$m_{1,2}(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B) m_2(C)$$
$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C)$$

Here K is a measure of the amount of conflict, and the bpa $m_{1,2}$ is only defined if $K \neq 1$.

DS defines belief, plausibility and ignorance as measures for confidence; the belief of a set defines the total confidence in this set, and plausibility an upper bound on the probability of this set:

$$Bel(A) = \sum_{B \subseteq A} m(B)$$
$$Pl(A) = 1 - Bel(\neg A)$$

3.2 Agents on situations and options

In our first attempt we want our agents to speak over both situations and options. We will thefore define the FoD of the agent as $D = \mathcal{P}(Sit \cup Pref)$. This is possible because there is a bijective function f which maps $\mathcal{P}(Sit \cup Pref)$ on $\mathcal{P}(Sit) \times \mathcal{P}(Pref)$. Now we define an agent as a bpa over the frame of discernment where a bpa assignment over a single item of the FoD, $bpa(\{s, p\}) = v$, indicates that we have a relative confidence of size v in a preference for option p in situation s. A bpa assignment over a set, $bpa(S) = v$, means that we have a relative confidence of size v in the combinations of preferences for options and situations in set S.

For example, suppose $Sit = \{C\}$ and $Pref = \{J\}$ and hence $D = \{\{J, C\}, \{C\}, \{J\}, \{\}\}$. If an agent m has a relative confidence of v that a user prefers jazz and is in the coffeeroom, we can represent this as $m(\{\{J, C\}\}) = v$. Furthermore, if we have an agent m which has $m(\{\{J, C\}, \{E, C\}\}) = v$, it means that this agent has a relative confidence of v that the user either has a preference for jazz and is in the coffeeroom or has a preference for eighties and is in the coffeeroom. Moreover, the belief function for an agent m can be used to express the confidence in a set of options; for example $Bel(\{\{J, C\}, \{E, C\}\}) = v = \sum_{x \subseteq \{\{J,C\},\{E,C\}\}} m(x)$ means that the total confidence of agent m that the user has a preference for jazz or eighties and is in the coffeeroom, is v.

Problems arise when we want to represent conditional probabilities, such as, 'the confidence that the user has a preference for jazz music *when* she is in the coffeeroom equals 0.8'. Such property cannot be represented by assigning absolute values to outputs of m. We can, however, represent it as a constraint on m stating the ratio between $m(\{\{J, C\}\})$ and $m(\{\{C\}\})$ as follows:

$$m(\{\{J, C\}\}) : m(\{\{C\}\}) = 8 : 2$$

We will continue on this path in Section 4, but first lets see what we can do without ratios.

3.3 Agents on options only

Although in current DS theory there are no operations on statements about ratios, in this section we want to show that DS can help us if we focus on the preferences and forget, for one moment, the situations in which these preferences hold. In this case the agent is, similar to the previous section, defined as the bpa over a FoD. Since we, however, only look at options we define the FoD of the agent as $D = \mathcal{P}(Pref)$.

A bpa assignment over a single item of the FoD $bpa(\{p\}) = v$, indicates that we have a relative confidence of size v in a preference for option p and a bpa assignment over a set $bpa(S) = v$, means that we have a relative confidence of v in the preferences for options in set S.

For example, suppose $Pref = \{J, E\}$ and hence the frame of discernment is $\{\{J\}, \{J, E\}, \{E\}, \{\}\}$. If we have an agent m_1 with $m_1(\{\{J\}, \{J, E\}\}) = 0.8$ this is interpreted that, with a relative confidence of 0.8, the user represented by this agent prefers either jazz music which is not from the eighties or jazz music from the eighties. Since $\sum_{S \subseteq D} m(S)$ should be 1, we have to say something about the remaining 20% of the cases, if we want to fix m_1 completely. Suppose that in the remaining cases we don't know anything at all about the preferred option, then we say $m_1(\{\{J\}, \{J, E\}, \{E\}, \{\}\}) = 0.2$.

Now, suppose we have an agent m_2 with $m_2(\{\{J, E\}, \{E\}\}) = 0.7$, representing a user that likes music from the eighties with a relative confidence of 0.7, and we also assume that this agent does not know anything about the other 30%; $m_2(\{\{J\}, \{J, E\}, \{E\}, \{\}\}) = 0.3$. If both agents are valid in the same situation, we can combine the two agents using Dempster's rule of combination. Applying this rule gives (since K is zero)::

$$
\begin{aligned}
m_{1,2}(\{\{J\}, \{J, E\}, \{E\}, \{\}\}) &= \frac{1}{1-K} * (0.2 * 0.3) = 0.06 \\
m_{1,2}(\{\{J, E\}, \{E\}\}) &= \frac{1}{1-K} * (0.2 * 0.7) = 0.14 \\
m_{1,2}(\{\{J\}, \{J, E\}\}) &= \frac{1}{1-K} * (0.8 * 0.3) = 0.24 \\
m_{1,2}(\{\{J, E\}\}) &= \frac{1}{1-K} * (0.8 * 0.7) = 0.56
\end{aligned}
$$

Suppose we want to compare the four options: jazz music from the eighties, jazz music not from the eighties, and playing nothing at all, we can write down the belief and plausibility for each option as follows:

Hypothesis set	Bel	Pl
$\{\{J, E\}\}$	0.56	1
$\{\{J\}\}$	0	0.3
$\{\{E\}\}$	0	0.2
$\{\}$	0	0.06

This indicates that the new agent represents a total confidence of 0.56 in the user preferring jazz music from the eighties. It has no confidence for other options, but places upper bounds on their probability expressing that a preference for jazz music is more plausible than a preference for eighties music. Based on this information a system can decide for example, given the music available, which music to play.

	$\emptyset$	$\{C\}$	$\{C, J\}$	$\ldots$	$\{C, H\}$	$\{C, H, J\}$	$\ldots$	$\{C, H, J, E\}$
f		0.1		$\ldots$	0.4		$\ldots$	0.3
f'		0.8		$\ldots$	0.0		$\ldots$	0.1
f''								

Table 1: Example agent

4 Agents as sets of functions

From the previous section we may conclude that for non-context-aware preferences, Dempster-Shafer seems a suitable technique to model and combine beliefs about these preferences. However, to try to include both situations and preferences in the same model, and meanwhile support the ratios for representing choices of users in situations as mentioned in Section 3.2, we will in this section try an alternative method of defining the agents in terms of sets of functions on the frame of discernment, distributing a probability mass among this set. We motivate this by assuming that the user prefers the same properties of options in new situations as he did in situations with the same properties in his history. This history we will call a *perfect history* (perfect for purposes of determining the preferences of a user in new situations), and we want to model this history. We first provide a formal definition and operators for combining agents and extending their domain. We end the section by giving an example of how to derive information from an agent by introducing some constraints.

Since in this section and in the rest of the report we often discuss the domain of the agents, we will, from now on, represent the domain (and not the FoD) as D. The domain is in our case $Sit \cup Pref$.

Definition 3. An agent M is defined as $M \subseteq \{f : \mathcal{P}D \rightarrow 0..1 | \sum_x f(x) = 1\}$, where $D = Sit \cup Pref$.

Such a set of functions indicates that the history of the user modeled by this agent is described by either one of these functions. Where $f(x) = v$, means that in $(v * 100)\%$ of the cases, x was the case. For example, for the domain $\{C, H, J, E\}$, the history of the user modeled in Table 4 is either f, or f', or f''. The motivation behind providing *possible histories* is that we might be unsure about the whole history, for example because knowledge about the history came from different sources. It is possible though that we know certain ratio's between certain options in the past in a certain situation, for example resulting from datamining sensor information at a specific location. These ratios can be represented using these agents.

For example, to represent a user who 80% of the time listened to jazz music when he was in the coffeeroom, using an agent M_1, we can write:

$$M_1 = \{f : \mathcal{P}D_1 \rightarrow 0..1 | \sum_x f(x) = 1 \wedge \sum_{x \ni C, J} f(x) : \sum_{x \ni C, \not\ni J} f(x) = 8 : 2\}$$

with, for example, $D_1 = \{J, C\}$.

A specific function f from M_1 is:

	{J}	{}
{C}	0.4	0.1
{}	0.5	0.0

In this table, a value v at row x and column y means that $f(x \cup y) = r$. As another example, consider a user who 70% of the time listened to music of the eighties when he was happy, represented by agent M_2, with $D_2 = \{E, H\}$:

$$M_2 = \{f : \mathcal{P}D_2 \to 0..1 | \sum_x f(x) = 1 \wedge \sum_{x \ni E,H} f(x) : \sum_{x \ni H, \not\ni E} f(x) = 7 : 3\}$$

4.1 Extending the domain of an agent with hypothesis about new statements

To be able to combine agents with different domains, we have to be able to extend the domain of an agent.

For this purpose we will define the *domain introduction operator* ◂ as follows:

$$\begin{aligned}
&M_{old} \blacktriangleleft D_{ext} = M_{new} \implies \\
&\quad \exists F : M_{new} \overset{\text{surjective}}{\to} M_{old}, \text{ such that} \\
&\quad\quad \forall_{f \in M_{old}; g \in M_{new}} \bullet Fg = f \Leftrightarrow g \in G \\
&\quad\quad\quad ,\text{where } G = \{g | \forall_{X \subseteq D_{old}} \left(f(X) = \sum_{Y \subseteq D_{new} | Y \cap D_{old} = X} g(Y) \right) \} \\
&\quad , \text{ where } D_{new} = D_{old} \cup D_{ext}
\end{aligned}$$

In words: If agent M_{new} is an extension of M_{old} with domain D_{ext}, this means that M_{new} consists of, for each function f in M_{old}, the set of all possible combinations of functions that, if we sum their output for input values which intersected with the domain of M_{old} give a certain input value x the result is equal to $f(x)$ in the old agent.

The intuition behind this is that the new agent does not have any constraints on the added domain but keeps the constraints on the old domain.

Since we are mainly interested in ratios between different outcomes, we want to prove that ratios in a certain agent also exist in all extensions of the agent:

$$\begin{aligned}
&M_{old} \blacktriangleleft D_{ext} = M_{new} \implies \\
&\quad \forall_{P,Q \subseteq D_{old}} ((\forall_{f \in M_{old}} f(P) : f(Q) = a : b) \implies \\
&\quad \forall_{g \in M_{new}} \sum_{P' \subseteq D_{new} | P' \cap D_{old} = P} g(P') : \sum_{Q' \subseteq D_{new} | Q' \cap D_{old} = Q} g(Q') = a : b)
\end{aligned}$$

Because all functions in M_{new} can be mapped on a function in M_{old}, if a ratio is present in all functions of M_{old}, it will also be present in all functions of M_{new}. Therefore, it suffices to prove that the ratios of each individual function f in M_{old} are also present in the functions g of M_{new} that map on f.

$$\begin{aligned}
&\forall_{P,Q \subseteq D_{old}} (f(P) : f(Q) = a : b \implies \\
&\forall_{g | g \in M_{new} \wedge Fg = f} \sum_{P' \subseteq D_{new} | P' \cap D_{old} = P} g(P') : \sum_{Q' \subseteq D_{new} | Q' \cap D_{old} = Q} g(Q') = a : b)
\end{aligned}$$

Since for each of these functions g, that map on f, we know that $Fg = f$, we can use the requirements of the mapping function F to show that:

$$\forall_{X \subseteq D_{old}} \left(f(X) = \sum_{Y \subseteq D_{new} | Y \cap D_{old} = X} g(Y) \right)$$

This means that (even) the values for a and b are the same for f and the sum of its counterparts in g, and hence, also the ratio between a and b.

We presume that, next to these ratios present in the original agent, no new constraints will introduced in the extended agent. For example, if we want to extend the domain of agent M_1 with the domain of agent M_2 the domain becomes $D'_1 = D_1 \cup D_2$ and we presume that the agent becomes:

$$M'_1 = \{f : \mathcal{P}(D'_1) \to 0..1 | \sum_x f(x) = 1 \wedge \sum_{x \ni C,J} f(x) : \sum_{x \ni C, \not\ni J} f(x) = 8 : 2\}$$

Note that, because in agent M_1 we represented the values over which we have to do the summation as set inclusion, we don't have to change the constraint.

4.2 Combining agents

We will use a combination operator, $\boxplus$, to combine two agents. For this we first extend the domain of the two agents to a common domain as shown in the previous section, after which we intersect the resulting agents:

Definition 4. For two agents M_1 and M_2 the *combination operator* $\boxplus$ is defined as $(M_1 \blacktriangleleft D_2) \cap (M_2 \blacktriangleleft D_1)$

The motivation behind this way of combining agents is that both agents represent a set of possible histories which satisfy their knowledge. Assuming both agents are correct, the real history of the user must be in both sets of possible histories so we can take the intersection of them.

As a result, if the agents have conflicting constraints on their functions the resulting agent has an empty set of functions. For example, if we combine agent M_1 from Section 4 with the following agent:

$$M_{conflict} = \{f : \mathcal{P}D_1 \to 0..1 | \sum_x f(x) = 1 \wedge \sum_{x \ni C,J} f(x) : \sum_{x \ni C, \not\ni J} f(x) = 9 : 1\}$$

the resulting agent is $\emptyset$ and we have no knowledge about the history of the user.

Example If we want to combine agents M_1 and M_2 in the previous example to a new agent $M_{1,2}$, $D_{1,2}$ becomes $\{J, C, E, H\}$ and the new agent $M_{1,2}$ becomes:

$$\begin{aligned} M_{1,2} = \{f : \mathcal{P}D_{1,2} \to 0..1 | \sum_x f(x) = 1 \wedge \\ \sum_{x \ni C,J} f(x) : \sum_{x \ni C, \not\ni J} f(x) = 8 : 2 \\ \wedge \sum_{x \ni E,H} f(x) : \sum_{x \ni H, \not\ni E} f(x) = 7 : 3\} \end{aligned}$$

4.3 Deducing information from agents

In this section we try to show how to deduce usefull information out of an agent in a specific situation, by introducing some constraints. In this case, we are interested in the type of music to play if the user is happy and in the coffeeroom, based on agent

$M_{1,2}$, which is the combination from M_1 and M_2 (as introduced before). To repeat, the constraints on the agent were:

$$\sum_x f(x) = 1 \tag{1}$$

$$\sum_{x \ni C,J} f(x) : \sum_{x \ni C, \not\ni J} f(x) = 8 : 2 \tag{2}$$

$$\sum_{x \ni E,H} f(x) : \sum_{x \ni H, \not\ni E} f(x) = 7 : 3 \tag{3}$$

The assumption we make is that constraints of the constraints of agents M_1 and M_2 are independent of the situations of the other; if the user is in the coffeeroom, being happy doesn't influence the like or dislike for Jazz and neither being in the coffeeroom influences the like or dislike of eighties music, given that the user is happy. Or, represented as ratios, whether the user is happy or not, does not influence the ratio between jazz and not jazz, when the user is in the coffeeroom and whether or not the user is in the coffeeroom, does not influence the ratio between eighties and not eighties, when the user is happy. A reason for this assumption could be that there is simply no information about any dependency between the two constraints and that it is our best guess.

This can be translated to adding the following constraints to $M_{1,2}$, which we will call $M'_{1,2}$:

$$\sum_{x \ni J,C,H} f(x) : \sum_{x \ni C,H, x \not\ni J} f(x) = \sum_{x \ni J,C, x \not\ni H} f(x) : \sum_{x \ni C, x \not\ni J,H} f(x) \tag{4}$$

$$= \sum_{x \ni J,C} f(x) : \sum_{x \ni C, x \not\ni J} f(x) \tag{5}$$

$$\sum_{x \ni E,C,H} f(x) : \sum_{x \ni C,H, x \not\ni E} f(x) = \sum_{x \ni E,H, x \not\ni C} f(x) : \sum_{x \ni H, x \not\ni E,C} f(x) \tag{6}$$

$$= \sum_{x \ni E,H} f(x) : \sum_{x \ni H, x \not\ni E} f(x) \tag{7}$$

Here (5) follows from (4), and (7) follows from (6) by simple arithmetic only.

Suppose we want to choose between four options of music with the properties: $\{J, E\}, \{J\}, \{E\}, \{\}$ and we want to know the relative probabilities that agent $M'_{1,2}$ has that these combinations of properties were chosen during the history of the user:

$$\sum_{x \ni J,E,C,H} f(x) \tag{8}$$

$$\sum_{x \ni J,C,H, x \not\ni E} f(x) \tag{9}$$

$$\sum_{x \ni E,C,H, x \not\ni J} f(x) \tag{10}$$

$$\sum_{x \ni C,H, x \not\ni J,E} f(x) \tag{11}$$

.

We therefore use the constraints of agent $M'_{1,2}$ on the options. Based on equation (5) and (2) we know that the agent has the constraint:

$$\sum_{x \ni J,C,H} f(x) : \sum_{x \ni C,H, x \not\ni J} f(x) = 8 : 2 \tag{12}$$

And similarly, from equation (7) and (3), we know our agent has the constraint:

$$\sum_{x \ni E,C,H} f(x) : \sum_{x \ni C,H,x \not\ni E} f(x) \qquad = 7:3 \tag{13}$$

From equation(12) and (13), it follows that:

$$\begin{aligned} (eq.8 + eq.9) : (eq.10 + eq.11) &\qquad = 8:2 \\ (eq.8 + eq.10) : (eq.9 + eq.11) &\qquad = 7:3 \end{aligned}$$

Because we wanted to have relative probabilities we can distribute a mass of one among the options. This gives, if we adhere to the constraints:

$$\begin{aligned} \sum_{x \ni J,E,C,H} f(x) &= 0.56 + x \\ \sum_{x \ni J,C,H,x \not\ni E} f(x) &= 0.24 - x \\ \sum_{x \ni E,C,H,x \not\ni J} f(x) &= 0.14 - x \\ \sum_{x \ni C,H,x \not\ni J,E} f(x) &= 0.06 + x \end{aligned}$$

With x such that the all values are greater than zero. We can use these constraints to determine what music to play when the user is happy and in the coffeeroom.

5 In perspective: Relation to probability theory

In this section we will explain how our agent relates to probability theory. First of all, our agent could be mapped to probability theory as a disjunction of probability distributions. For an agent M, with domain D this results in:

$$\bigvee_{f \in M} \Big(\bigwedge_{x \subseteq D} p(\wedge_{q \in x} q \wedge_{r \in D \setminus x} \neg r) = f(x) \Big)$$

This follows almost immediately from the definition of an agent, where the histories translate to probability distributions on the domain. For example, the agent in Table 4, can be translated to the following distribution:

$$\begin{aligned} &(p(C, \neg H, \neg J, \neg E) = 0.1 \wedge ... \wedge p(C, H, \neg J, \neg E) = 0.4 \wedge ... \wedge p(C, H, J, E) = 0.3) \\ &\vee (p(C, \neg H, \neg J, \neg E) = 0.8 \wedge ... \wedge p(C, H, \neg J, \neg E) = 0.0 \wedge ... \wedge p(C, H, J, E) = 0.1) \\ &\vee ... \end{aligned}$$

Similarly we can go back from constraints on the probability space to an agent. We thereby distinguish between completely known probability distributions and "uncertain" distributions. In case the complete distribution is known, we can represent the constraint with an agent of one function:

$$M = \{f\}, \text{where} \forall_{x \subseteq D} f(x) = p(\wedge_{q \in x} q \wedge_{r \in D \setminus x} \neg r)$$

In case we do not completely know the probability distribution, the information we know from the distribution is translated to constraints on all the functions of the agent.

This means that all possible probability distributions are represented as a single function (which can lead to an infinite amount of functions). Suppose Q represents the set probability distributions that comply with a certain constraint, this constraint is modeled using the following agent M:

$$M = \{f | \exists_{p\in Q} \forall_{x\subseteq D} f(x) = p(\wedge_{q\in x} q \wedge_{r\in D\setminus x} \neg r)\}$$

A special constraint on probabilities is the conditional probability which can be translated to a ratio, for all $A, B \subseteq D$ with $A \cap B = \emptyset$:

$$P(A|B) = x \Leftrightarrow M = \{f | \sum_{A,B\subseteq x} f(x) : \sum_{B\subseteq x, A\cap x=\emptyset} = x : (1-x)\}$$

For example:

$$\begin{aligned} D &= \{J, E, C, H\} \\ A &= \{J, E\} \\ B &= \{C, H\} \\ P(J, E|C, H) = 0.9 &\Leftrightarrow M = \{f | \sum_{x\ni J,E,C,H} f(x) : \sum_{x\ni C,H, x\not\ni J,E} = 0.9 : 0.1\} \end{aligned}$$

Example using conditional probabilities To show the similarity between our approach and probability theory, we model the agents with constraints from Section 4.3 using probability theory with conditional probabilities. We still assume that the user prefers the same properties of options in new situations as he did in situations before with the same properties. In this case our knowledge that the user listened 80% of the time to jazz music when he was in the coffeeroom and listened 70% of the time to music of the eighties when he was happy, is represented by two conditional probabilities:

$$\begin{aligned} P(J|C) &= 0.8 \\ P(E|H) &= 0.7 \end{aligned}$$

The assumption that the constraints of agents M_1 and M_2 were independent of the situations of the other translates to:

$$\begin{aligned} P(J|C, H) &= P(J|C) \\ P(E|H, C) &= P(E|H) \end{aligned}$$

Now, if we know that the user is happy and in the coffeeroom, we know that $P(C, H) = 1$ which means that from the agent it follows that:

$$\begin{aligned} P(J, C, H) &= P(J|C, H)P(C, H) && = 0.8 \\ P(E, C, H) &= P(E|C, H)P(C, H) && = 0.7 \end{aligned}$$

From probability theory it follows that:

$$P(J, C, H) = P(J, C, H, E) + P(J, C, H, \neg E) \quad = 0.8 \qquad (14)$$

$$P(E, C, H) = P(E, C, H, J) + P(E, C, H, \neg J) \quad = 0.7 \qquad (15)$$

Suppose we have to choose again between our four options of music with the properties: $\{J, E\}, \{J\}, \{E\}, \{\}$. We can distribute the probabilities of equations (14) and (15) over the four different options. This results in the same probabilities as our agent based approach (as the reader may check):

$$P(J, E, C, H) = 0.56 + x$$
$$P(J, \neg E, C, H) = 0.24 - x$$
$$P(\neg J, E, C, H) = 0.14 - x$$
$$P(\neg J, \neg E, C, H) = 0.06 + x$$

With x, again, such that the all values are greater than zero.

6 Conclusions and Future Work

In this report we discussed a way to predict the preferences the user has by modeling the user as an agent; relating properties of situations to properties of options the user prefers. We first looked at these agents as belief probability assignments in Dempster Shafer theory (DS), from which we concluded that DS could be usefull but we needed support for reasoning about ratios between outputs of the belief probability assignment function, to reason about conditional probabilities.

Therefore we presented an alternative by defining an agents as a set of functions, motivated by the notion of *possible histories*. We explored the relation to probability theory; formally and with an example. From this, we conclude that with the current definition, we cannot express more things than in ordinary probability theory nor did we, for the example, arrive at easier ways of deducing information from the agents.

However, we expect that the contribution of our work is in the description of the "prediction" of (preferred) situations by looking at different indicators, and modeling them and their integration as possible histories. We see applications for this method in the combination/integration of rules from rule mining systems. Another closely related direction of future research is to assign different strengths to the constraints on the agents. This could for example result from the integration of different agents that had different confidence in their possible histories. Related to rule mining systems, we could imagine that by assigning these strengths, we could address both *confidence* and *support* of rules when using rules for prediction.

References

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