

# A Belief Augmented Frame Computational Trust Model

Colin Keng-Yan TAN

Department of Computer Science,

School of Computing,

National University of Singapore

ctank@comp.nus.edu.sg

## Abstract

We present a novel trust model based on BAF-Logic, a system of reasoning that was originally developed for Belief Augmented Frames (BAF), to perform inexact reasoning over knowledge represented as Minsky frames and augmented with twin belief values that measure an Agent's degree of belief for and against a proposition. By applying BAF-Logic to trust modeling, we are not only able to model trust based on statistical measures, but also with propositional logic, thus enabling an Agent to evaluate another Agent's trustworthiness not only based on experience and reputation, but also based on logical arguments for and against trusting that other Agent. We present an extended example demonstrating how this model may be applied, followed by a discussion, and finally we conclude this paper with suggestions for further work.

## Introduction

In a society of agents one agent may provide information to another agent that is correct, misleading, or completely incorrect. As such, upon receiving the information, the receiving agent must assess the reliability of the source agent before incorporating this information into its knowledge base.

In this paper, the aim of a computational trust model is to provide a formal specification on how a receiving agent is to work out the trustworthiness of the source.

A number of models have been proposed in recent years. (Wang and Vassileva 2003) presents a trust model based on Bayesian networks. This model links various aspects of trust together as nodes in a Bayesian network, and applies the Bayesian rule to obtain an overall trust value. The authors present a case study for a file sharing service where the various aspects of trust in a file provider are reflected by components like download speed, file quality and types of files available. A weighted scheme is used to give more priority to some aspects and less to others. A noteworthy feature of their model is that agents can "gossip" with each other by comparing their Bayesian networks. (Ramchurn et al 2003) apply fuzzy sets and rules to evaluate the trustworthiness of contractors fulfilling obligations that may be explicitly stated in the contract, or implicitly assumed in the environment. The reputation of contractors

is taken as an aggregation of opinions of other agents with respect to the contractor over a particular issue. (Yu and Singh 2002) apply Dempster-Shafer Theory to computing trust (Dempster 1967), (Shafer 1976). This allows their model to state that an agent has no known reputation, as opposed to having a bad reputation, which may be implied by classical statistical approaches. They define two types of beliefs in the trustworthiness of the agent; a *local belief* that defines what one agent thinks of another agent based on actual encounters, and a *total belief* that combines the local belief together with beliefs of other agents in the society, gathered through a TrustNet, a network of agents giving opinions on other agents. (Josang 99) introduces Subjective Logic, a system of logic for evaluating trust in authentication chains. An agent's opinion about the trustworthiness of another agent is modeled as a triple  $\{b, d, u\}$ , where  $b$  is the degree of belief,  $d$  is the degree of disbelief, and  $u$  is the uncertainty (ignorance) that the agent holds, and  $b + d + u = 1$ . He defines five operations on opinions, and defines an algebra for computing certifying and authenticating public keys in a public-key infrastructure.

In this paper we present a novel trust model based on Belief Augmented Frames. Belief Augmented Frames are introduced in (Tan and Lua 2003a) and (Tan and Lua 2003b). The receiving agent A (called the "Assessor Agent") assesses the source S (called the "Subject Agent") based on its own experiences of the reliability of the S, on various prejudices that A holds, and on S's reputation amongst other agents, called "Accessory Agents". Our approach is novel in that it employs belief masses and a belief-based reasoning system called "Belief Augmented Frame Logic" or BAF-Logic, which is proven in (Tan 2003) to be logically complete and sound.

## An Introduction to Belief Augmented Frames

In belief models that possibility of an event occurring is modeled as a range of values, rather than as a single point probability. This range allows us to express ignorance, which standard statistical measures do not accommodate. An example of the limitation of statistical measures is given in (Adams 1985), where doctors, who predicted with

a certainty of  $x\%$  that a patient was suffering from a particular illness, were reluctant to predict with a certainty of  $(100 - x)\%$  that the patient was not suffering from that illness.

This apparent contradiction is a reflection of classical statistic's inability to cater to ignorance.

The Assessor Agent  $A$  faces a similar situation in assessing the trustworthiness of a Subject Agent  $S$ .  $A$  may be willing to assign a particular probability that  $S$  is trustworthy, but may be reluctant to assign a probability of  $(1 - S)$  that  $S$  is untrustworthy. This is simply because  $A$  may not know the trustworthiness of  $S$ . Thus beliefs form a logical model for assessing trustworthiness.

Various models of beliefs have been proposed, including the seminal Dempster-Shafer model mentioned earlier. Smets generalized the model in (Smets 2000) to form the Transferable Belief Model. Picard (Picard 2000) proposes the Probabilistic Argumentation System, which combines propositional logic with probability measures to perform reasoning.

### Belief Augmented Frames

In classical AI a frame represents an object in the world, and slots within the frame indicate the possible relations that this object can have with other objects. A value (or set of values) in a slot indicates the other objects that are related to this object through the relation represented by the slot. The existence of a slot-value pair indicates a relation; the absence indicates that there is no relation.

In BAFs each slot-value pair is augmented by a pair of belief masses  $\phi_{rel}^T$  and  $\phi_{rel}^F$ .  $\phi_{rel}^T$  is the degree of belief in the claim that the relationship  $rel$  exists, while  $\phi_{rel}^F$  is the degree of belief that the relationship does not exist. Both  $\phi_{rel}^T$  and  $\phi_{rel}^F$  are bound by:

$$0 \leq \phi_{rel}^T, \phi_{rel}^F \leq 1 \quad (1)$$

$$\text{In general, } \phi_{rel}^T + \phi_{rel}^F \neq 1 \quad (2)$$

Equation (2) states that  $\phi_{rel}^T$  and  $\phi_{rel}^F$  may not necessarily sum to 1. This frees us from the classical statistical assumption that  $\phi_{rel}^F = 1 - \phi_{rel}^T$  and allows us to model ignorance. It is also possible that  $\phi_{rel}^T + \phi_{rel}^F > 1$ . See our discussion of ignorance for more on this.

Both  $\phi_{rel}^T$  and  $\phi_{rel}^F$  may be derived from various independent sources, or may be computed by using a system of logic called "BAF-Logic" which will be presented in the next section. Effectively this allows us to model the belief in a problem as a set of arguments for the belief, and a set of arguments against it.

While  $\phi_{rel}^T$  and  $\phi_{rel}^F$  represent the degree of belief for and against a claim, the overall truth is given by the *Degree of Inclination DI*:

$$DI_{rel} = \phi_{rel}^T - \phi_{rel}^F \quad (3)$$

$$-1 \leq DI_{rel} \leq 1 \quad (4)$$

$DI_{rel}$  measures the overall degree of truth of the relationship  $rel$ , with  $-1$  representing falsehood,  $1$  representing truth, and values in between representing various degrees of truth and falsehood. As an example, we could take  $-0.25$  to mean "possibly false",  $-0.5$  to mean "probably false", etc.

The Utility Function  $U_{rel}$  is defined as:

$$U_{rel} = \frac{DI_{rel} + 1}{2} \quad (5)$$

The Utility Function maps the Degree of Inclination to a  $[0,1]$  range. If we normalize the Utility Functions for all relations so that they sum to 1, we can use these normalized values as statistical measures representing the probability of a relation being true.

The plausibility  $pl_{rel}$  is the upper bound that  $\phi_{rel}^T$  can take, and it is defined by:

$$pl_{rel} = 1 - \phi_{rel}^F \quad (6)$$

If  $\phi_{rel}^T > pl_{rel}$ , then either the data generating  $\phi_{rel}^T$  is overly optimistic ( $\phi_{rel}^T$  is too large), or overly pessimistic ( $\phi_{rel}^F$  is too large, resulting in  $pl_{rel}$  being too small). In either case the condition  $\phi_{rel}^T > pl_{rel}$  indicates that the data generating  $\phi_{rel}^T$  and  $\phi_{rel}^F$  is conflicting.

The ignorance in our system is given by  $ig_{rel}$ , and is defined as:

$$ig_{rel} = pl_{rel} - \phi_{rel}^T \quad (7)$$

Note that  $ig_{rel}$  is a negative number when  $\phi_{rel}^T > pl_{rel}$ . As discussed earlier this is indicative that the data supporting and refuting  $rel$  is conflicting. In such cases  $\phi_{rel}^T + \phi_{rel}^F > 1$ .

There are many other features of BAFs like inheritance of relationships, generalization of concepts and daemons that are beyond the scope of this paper. The interested reader is referred to (Tan 2003).

### Belief Augmented Frame Logic

Belief Augmented Frame Logic (BAF-Logic) is a system designed to reason over the  $\phi_{rel}^T$  and  $\phi_{rel}^F$  values in the frame.

Given two relations  $P$  and  $Q$ , we define the conjunction  $P \wedge Q$  as:

$$\varphi_{P \wedge Q}^T = \min(\varphi_P^T, \varphi_Q^T) \quad (8)$$

$$\varphi_{P \wedge Q}^F = \max(\varphi_P^F, \varphi_Q^F) \quad (9)$$

This definition is based on the intuitive idea that the strength of  $P \wedge Q$  being true rests on the strength of the weakest proposition  $P$  or  $Q$ . Likewise, if either  $P$  or  $Q$  were false, then  $P \wedge Q$  would be false, and we can base our degree of belief in  $P \wedge Q$  being false on the strongest proposition that either  $P$  or  $Q$  is false.

We define the disjunction between  $P$  and  $Q$  as:

$$\varphi_{P \vee Q}^T = \max(\varphi_P^T, \varphi_Q^T) \quad (10)$$

$$\varphi_{P \vee Q}^F = \min(\varphi_P^F, \varphi_Q^F) \quad (11)$$

This is again based on the intuition that for  $P \vee Q$  to be true, either  $P$  is true or  $Q$  is true, and we can place our confidence that  $P \vee Q$  is true on the strength of the strongest proposition. Likewise for  $P \vee Q$  to be false, both  $P$  and  $Q$  must be false. Similar to our definition of conjunction, we base our confidence that  $P \vee Q$  is false on the weakest proposition that  $P$  or  $Q$  is false.

Finally, we define the logical NOT operation as:

$$\varphi_{\neg P}^T = \varphi_P^F \quad (12)$$

$$\varphi_{\neg P}^F = \varphi_P^T \quad (13)$$

This means that the degree that we believe that our data support  $\neg P$  is equal to the degree that it refutes  $P$ . Likewise the degree that our data refutes  $\neg P$  is equal to the degree that it supports  $P$ .

### Updating Beliefs

Without loss of generality, at any time  $t$  we can timestamp the belief values of  $P$ . For brevity we take  $P_t$  to collectively mean  $(\varphi_{P,t}^T, \varphi_{P,t}^F)$ , i.e. the pair of supporting/refuting belief masses for  $P$  at time  $t$ . Likewise we define  $P_*$  to be  $(\varphi_{P,*}^T, \varphi_{P,*}^F)$ , which is the derived belief masses  $\varphi_{P,*}^T$  and  $\varphi_{P,*}^F$  obtained from some source. We then define  $P_t$  to be:

$$P_t = P_{t-1} \otimes_{\alpha} P_* \quad (14)$$

For brevity, where equations for both  $\varphi_P^T$  and  $\varphi_P^F$  are identical, we will use the notation  $\varphi_P^X$ , where  $X = \{T, F\}$ .

Here  $\otimes$  is the belief revision operator, and this is defined as:

$$\varphi_{P,t-1}^X \otimes_{\alpha} \varphi_{P,*}^X = \alpha \varphi_{P,*}^X + (1 - \alpha) \varphi_{P,t-1}^X \quad (15)$$

The weight  $\alpha$  controls how much importance is placed on the derived belief masses in  $P_*$ , and how much importance is placed on the previous belief masses.

We define two default belief masses  $\varphi_{DEF}^T$  and  $\varphi_{DEF}^F$  to be:

$$\varphi_{DEF}^X = 0 \quad (16)$$

This means that by default, if we do not know the supporting and refuting belief masses for a relation, we assume that these masses are both 0. Following equations 3, 5, 6 and 7 we obtain:

$$DI_{DEF} = 0.0 \quad (17)$$

$$U_{DEF} = 0.5 \quad (18)$$

$$Pl_{DEF} = 1.0 \quad (19)$$

$$Ig_{DEF} = 1.0 \quad (20)$$

Thus by choosing default values of 0 for both supporting and refuting belief masses, we get a  $DI$  representing ignorance, a Utility Function that is 50% true and 50% untrue, a plausibility that represents that there is no reason why the relationship cannot be true, and complete ignorance about the truth of the relationship.

We therefore choose  $\varphi_{DEF}^T$  and  $\varphi_{DEF}^F$  as the initial masses for  $P_0$ .

$$\varphi_{P,0}^X = \varphi_{DEF}^X \quad (21)$$

## Modeling Beliefs with Belief Augmented Frames

We can now describe how  $A$  models her belief in the trustworthiness of  $S$ .

### Modeling the Trust

Within  $A$ 's mind she models her trust of  $S$  as a *trust* relation between herself and  $S$ , with belief masses  $\varphi_{trust}^T$  and  $\varphi_{trust}^F$ . This is shown in Figure 1.

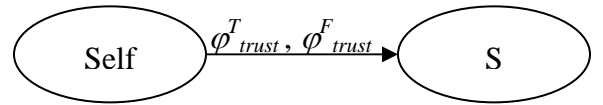


Figure 1. Modeling Trust Between A and S

The frame labeled “Self” represents  $A$  herself and the frame labeled  $S$  represents the Subject Agent  $S$ , all in  $A$ 's mind. The *trust* belief masses are composed of  $S$ 's personal opinion of  $A$ 's reliability, and what Accessory Agents in the environment say about  $S$ . I.e.  $S$ 's reputation in the society.  $A$  weighs these two components with a weight  $\beta$ :

$$\varphi_{trust}^X = \beta \varphi_{opinion}^X + (1 - \beta) \varphi_{reputation}^X \quad (22)$$

**Deriving Opinion Values.**  $A$ 's opinion of  $S$  is made of two components; one component is her personal experience with  $S$ 's trustworthiness, the other component is her prejudices for and against certain traits that  $S$  may or may not have (we term this her "world view"). Again she weighs these traits with a weight  $\chi$ :

$$\varphi_{opinion}^X = \chi \varphi_{experience}^X + (1 - \chi) \varphi_{wv}^X \quad (23)$$

The *experience* belief masses may come from various sources. For example, they may come from the number of times that what  $S$  has told  $A$  turns out to be true, over the total number of things he has told her. Alternatively, any of the statistical trust measures surveyed earlier may also be used to derive  $\varphi_{experience}^T$  and  $\varphi_{experience}^F$  by computing  $\varphi_{experience}^T$  and setting  $\varphi_{experience}^F = 1 - \varphi_{experience}^T$ .

The belief masses for her worldview  $wv$  will come from propositional rules that are inherent in  $A$ . For example, given certain facts  $P, Q, R$  and  $V, W, Y$  and  $Z$ ,  $A$ 's world view might be modeled as:

$$trust(A, S) : -P \wedge Q \vee \neg R \quad (24)$$

$$\neg trust(A, S) : \neg V \wedge W \vee (Y \wedge \neg Z) \quad (25)$$

I.e.  $A$  trusts  $S$  if  $P$  and  $Q$  are true, or if  $R$  is not true, etc. This can be modeled using a modified application of BAF Logic:

$$\varphi_{wv}^T = \max(\min(\varphi_P^T, \varphi_Q^T), \varphi_R^F) \quad (26)$$

$$\varphi_{wv}^F = \max(\min(\varphi_V^T, \varphi_W^T), \min(\varphi_Y^T, \varphi_Z^F)) \quad (27)$$

$A$  might also rate the importance of the facts used in the rules. This can be modeled as weights  $\omega_P, \omega_Q$  etc. Let  $\varphi_P^X, = \omega_P \varphi_P^X$  for some fact  $P$ . Equations 26 and 27 then become:

$$\varphi_{wv}^T = \max(\min(\varphi_P^{T'}, \varphi_Q^{T'}), \varphi_R^{F'}) \quad (28)$$

$$\varphi_{wv}^F = \max(\min(\varphi_V^{T'}, \varphi_W^{T'}), \min(\varphi_Y^{T'}, \varphi_Z^{F'})) \quad (29)$$

**Deriving Reputation Values.**  $S$ 's reputation is derived from the *trust* belief masses worked out by Accessory Agents  $C_i$  in the society. In addition,  $A$  would also have various *trust* relations with each  $C_i$ , each with their own belief masses. Let  $\alpha_i$  be the Utility value (see equation 5) of the *trust* belief masses between  $A$  and  $C_i$ . We first initialize  $\varphi_{reputation,0}^X = \varphi_{DEF}^X, X = \{T, F\}$ . We call this pair of initial reputation values  $rep_{S,0}$ . The final reputation belief masses are then derived by composing all the belief masses returned by each of the  $n$  Accessory Agent  $C_i$  ( $i=0..n-1$ ) Let  $trust_i$  represent the  $(\varphi_{rep,i}^T, \varphi_{rep,i}^F)$  pair returned (or reported) by an Accessory Agent  $C_i$ . Then:

$$rep_{s,n-1} = rep_{s,0} \otimes_{\alpha_0} trust_0 \otimes_{\alpha_1} trust_1, \dots \quad (30)$$

**Deriving Weight Values.** Equations 22 to 29 employ weights to weigh the importance of issues relating to how trustworthy  $A$  thinks that  $S$  is.

For this we propose 6 grades of importance: Disregarded, Unimportant, Slightly Unimportant, Slightly Important, Important, Crucial. We employ a 6-valued Lukasiewicz Logic, where the  $i$ th grade of importance  $imp_i$  is given the weight:

$$imp_i = \frac{i}{N-1} = \frac{i}{5} \quad (31)$$

We can then express the grades of importance in a fuzzy set (Zadeh 1965):

$$set_{imp} = \{Disregarded/0.0, Unimportant/0.2, Slightly Unimportant/0.4, Slightly Important/0.6, Important/0.8, Crucial/1.0\} \quad (32)$$

## An Example Application

### Computing the Trust Model

Alice works in an office with Bob, Charlie, Eve and Simon. Simon has just made a claim to Alice, and Alice must assess how truthful Simon is. Alice has a unique world view. In particular she distrusts someone if he wears sunglasses all the time, or if he has a long straggly beard. Unfortunately this is both true of Simon. However Alice trusts a person if he has a PhD, or if he dresses well. This time, fortunately, both are also true of Simon. Alice also decides to take Simon's reputation amongst Bob, Charlie, and Eve into consideration. Alice estimates that about 70% of what Simon tells her turns out to be true.

After gathering the facts, Alice decides that she will rank each factor in the following way:

Factor	Importance	Weight
Personal Opinion	Important	$\beta = 0.8$
Personal Experience	Important	$\chi = 0.8$
Wears Sunglasses	Slightly Unimportant	$\omega_{sunnies} = 0.4$
Has beard	Slightly Important	$\omega_{beard} = 0.6$
Has PhD	Crucial	$\omega_{PhD} = 1.0$
Dresses Well	Slightly Important	$\omega_{dress} = 0.6$

**Table 1. Alice's Importance Rankings**

We can now work out the following equations:

$$\varphi_{trust}^X = 0.8 \varphi_{opinion}^X + 0.2 \varphi_{reputation}^X \quad (33)$$

$$\varphi_{opinion}^X = 0.8 \varphi_{experience}^X + 0.2 \varphi_{wv}^X \quad (34)$$

$$\varphi_{wv}^T = \max(\varphi_{PhD}^T, \varphi_{dress}^T) \quad (35)$$

$$\varphi_{wv}^F = \max(\varphi_{sunnies}^T, \varphi_{beard}^T) \quad (36)$$

$$\varphi_{PhD}^T = \varphi_{PhD}^T, \varphi_{dress}^T = 0.6\varphi_{dress}^T, \quad (37)$$

$$\varphi_{sunnies}^T = 0.4\varphi_{sunnies}^T, \varphi_{beard}^T = 0.6\varphi_{beard}^T$$

**Computing Alice's Personal Opinion of Simon.** Since Alice knows for a fact that Simon has a PhD, that he dresses well, that he wears sunglasses and that he has a beard, we can set  $\varphi_{PhD}^T = \varphi_{dress}^T = \varphi_{sunnies}^T = \varphi_{beard}^T = 1.0$ , and  $\varphi_{PhD}^F = \varphi_{dress}^F = \varphi_{sunnies}^F = \varphi_{beard}^F = 0.0$ . We will then have  $\varphi_{PhD}^T = 1.0$ ,  $\varphi_{dress}^T = 0.6$ ,  $\varphi_{sunnies}^T = 0.4$ ,  $\varphi_{beard}^T = 0.6$ , while at the same time  $\varphi_{PhD}^F = \varphi_{dress}^F = \varphi_{sunnies}^F = \varphi_{beard}^F = 0.0$ . From this we can derive:

$$\varphi_{wv}^T = \max(1.0, 0.6) = 1.0 \quad (38)$$

$$\varphi_{wv}^F = \max(0.4, 0.6) = 0.6 \quad (39)$$

Alice knows that Simon is right about 70% of the time. Hence Alice derives:

$$\varphi_{experience}^T = 0.7 \quad (40)$$

$$\varphi_{experience}^F = 1.0 - \varphi_{experience}^T = 0.3 \quad (41)$$

We can now compute Alice's personal opinion of Simon's trustworthiness:

$$\begin{aligned} \varphi_{opinion}^T &= 0.8 * 0.7 + 0.2 * 1.0 \\ &= 0.76 \end{aligned} \quad (42)$$

$$\begin{aligned} \varphi_{opinion}^F &= 0.8 * 0.2 + 0.2 * 0.6 \\ &= 0.28 \end{aligned} \quad (43)$$

**Computing Simon's Reputation.** Table 2 shows Alice's trust belief masses for Bob, Charlie and Eve.

Accessory Agent	$\varphi_{trust,i}^T$	$\varphi_{trust,i}^F$	$\alpha_i = \varphi_{trust,i}$
i=Bob	1.0	0.0	1.0
i=Charlie	0.8	0.1	0.85
i=Eve	0.1	0.9	0.1

**Table 2. Alice's Trust Belief Masses for Bob, Charlie and Eve**

Table 3 shows the trust belief masses for Simon as reported by Bob, Charlie and Eve.

Accessory Agent	$\varphi_{rep,i}^T$	$\varphi_{rep,i}^F$
i=Bob	0.8	0.1
i=Charlie	0.7	0.2
i=Eve	1.0	0.0

**Table 3. Trust Belief Masses as Reported by Bob, Charlie and Eve**

Starting with an initial  $\varphi_{reputation}^X = \varphi_{DEF}^X$ , Table 4 shows the progression of Simon's reputation has Alice takes into account Bob's view, followed by Charlie's, and finally followed by Eve's.

Accessory Agent	$\alpha_i$	$\varphi_{reputation}^T$	$\varphi_{reputation}^F$
Initial	N/A	0.0	0.0
i=Bob	1.0	0.8	0.1
i=Charlie	0.85	0.715	0.185
i=Eve reports	0.1	0.745	0.167

**Table 4. Reputation Values taking Bob, Charlie and Eve's Reported Trust Belief Masses**

After incorporating Eve's reported trust belief masses, we obtain a final reputation value of  $\varphi_{reputation}^T = 0.745$ ,  $\varphi_{reputation}^F = 0.167$ . We finally compute Alice's trust belief values for Simon:

$$\begin{aligned} \varphi_{trust}^T &= 0.8 * 0.76 + 0.2 * 0.745 \\ &= 0.757 \end{aligned} \quad (44)$$

$$\begin{aligned} \varphi_{trust}^F &= 0.8 * 0.28 + 0.2 * 0.167 \\ &= 0.257 \end{aligned} \quad (45)$$

We summarize the key components of Alice's trust belief masses for Simon in Table 5, listing the  $DI$ ,  $Ig$  and  $U$  values.

Component	$\varphi^T$	$\varphi^F$	$DI$	$U$	$Pl$	$Ig$
World View	1.0	0.6	0.4	0.7	0.4	-0.6
Personal Experience	0.7	0.3	0.4	0.7	0.7	0.0
Personal Opinion	0.76	0.28	0.48	0.74	0.72	-0.04
Reputation	0.75	0.17	0.58	0.79	0.83	0.09
Trust	0.76	0.26	0.5	0.75	0.74	-0.01

**Table 5. Summary of Key Component Values**

## Discussion

Alice's World View was influenced heavily by the fact that she places essential emphasis on Simon's PhD. This dominated the belief that Simon was trustworthy. On the other hand, her belief mass that Simon was not trustworthy was dominated by his beard, which Alice placed more emphasis on when evaluating Simon.

Alice's evaluation of Simon was somewhat conflicting, and this is reflected in a negative ignorance value caused

by her overemphasis on Simon's PhD. In any case, in forming her personal opinion of Simon, Alice placed much more emphasis on her past contacts with Simon, and thus her experience of Simon's trustworthiness dominated her final opinion score of (0.76, 0.28). The conflicting opinions in Alice's world view carried over into her opinion, resulting in a slightly negative ignorance value.

Alice forms her reputation belief masses for Simon by asking Bob, Charlie and Eve, in that order. She has varying degrees of trust for each of them, and since she derived these trust belief masses in the same way that she is deriving Simon's belief masses, we see a practical example of how these masses can be employed to evaluate opinions of other agents. Starting with a completely ignorant position on Simon's reputation, she updates her belief masses as each Accessory Agent reports.

Alice then combines both her opinion of Simon and Simon's reputation in the community to derive the final trust belief masses for Simon. She can now use  $U_{trust}=0.75$  to evaluate Simon's claim.

## Conclusion and Further Work

In this paper we presented a novel computational trust model based on Belief Augmented Frames and BAF-Logic. In our model the Assessing Agent *A* can not only incorporate past experience of the Subject Agent *S* and his reputation in the community, but we can also incorporate biases and prejudices (world view) of the *A* into the computation. Though not demonstrated in the Case Study, it is also possible to incorporate rumors from Accessory Agents which, combined with *A*'s world view, can influence *A*'s assessment of *S*'s trustworthiness. All this makes our approach very powerful and expressive.

As a next step we will assess this model against other computational trust models, to evaluate its effectiveness and the implications of incorporating *A*'s world view and rumors from other Accessory Agents. We will also explore more sophisticated ways that the Assessing Agent can weigh its various issues, and how to incorporate other belief models.

## References

Yao, W., and Vassileva, J., 2003. Trust and Reputation Models in Peer-to-Peer Networks, In *Proceedings of the Third International Conference on Peer-to-Peer Computing (P2P'03)*, 150-159. Lkoping, Sweden.

Ramchurn, S. D., Jennings, N. R., Sierra, C., and Godo, L., 2003. A Computational Trust Model for Multi-Agent Interactions based on Confidence and Reputation, In *Proceedings of the 6<sup>th</sup> International Workshop on Deception, Fraud and Trust in Agent Societies*, 69-75. Melbourne, Australia.

Yu ,B., Singh, M. P., 2002. An Evidential Model of Distributed Reputation Management, In *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 294-301. Bologna, Italy.

Josang, A., 1999. An Algebra for Assessing Trust in Authentication Chains, In *Proceedings of the 1999 Network and Distributed Systems Security Symposium*, San Diego, California, USA. Page numbers not available.

Tan, C. K. Y., Lua, K. T., 2003b. Discourse Understanding with Discourse Representation Theory and Belief Augmented Frames, In *Proceedings of the 2<sup>nd</sup> International Conference on Computational Intelligence, Robotics and Autonomous Systems*, Singapore. Page numbers not available.

Tan, C. K. Y., Lua, K. T., 2003a. Belief Augmented Frames for Knowledge Representation in Spoken Dialog Systems, In *Proceedings of the 1<sup>st</sup> International Indian Conference on Artificial Intelligence*, Hyderabad, India. Page numbers not available.

Tan, C. K. Y., 2003. Belief Augmented Frames. Ph.D. diss., Dept. of Computer Science, School of Computing, National University of Singapore.

Dempster, A. P., 1967. Upper and Lower Probabilities Induced by Multivalued Mapping. *Annals of Mathematical Statistics*, 38 :325-339.

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

Adams, J. T., 1985. Probabilistic Reasoning and Certainty Factors. *Rule Based Expert Systems*, 263-271. Addison – Wesley.

Smets, P., 2000. Belief Functions and the Transferable Belief Model, <http://ippserv.rug.ac.be/documentation/belief/belief.pdf>

Picard, J., 2000. Probabilistic Argumentation Systems Applied to Information Retrieval. Ph.D. diss, Universite de Neuchatel, Suisse.

Zadeh L., 1965. Fuzzy Sets, *Information and Control*, 338-353.