

A Formalization of Trust Alignment

Andrew Koster^a Jordi Sabater-Mir^a Marco Schorlemmer^a

^a *Artificial Intelligence Research Institute, CSIC*

Abstract. We present a mathematical framework for communicating about trust in terms of interactions. We argue that sharing an ontology about trust is not enough and that interactions are the building blocks that all trust- and reputation models use to form their evaluations. Thus, a way of talking about these interactions is essential to gossiping in open heterogeneous environments. We give an overview of the formal framework we propose for aligning trust, based on the gossip agents send.

1. Introduction

In complex, distributed systems, such as multi-agent systems, the artificial entities have to cooperate, negotiate, compete, etc. amongst themselves. Thus the social aspect of these systems plays a crucial role in their functioning. One of the issues in such a social system is the question of whom to trust and how to find this out. There are several systems already in development that model trust and reputation [1], ranging from a straightforward listing of evaluations (such as eBay's [2] reputation system), to complex cognitive models (such as RePage [3]). We anticipate that in an open multi-agent system, there will be a large diversity of models in concurrent use by different agents, depending on the wishes of the programmer and the user. However, even if there is consensus on some model, this is still only a consensus on the computational representation. In a heterogeneous environment it is inevitable that, if the trust model an agent uses is based on cognitive principles [7,8], the way different agents interpret their environment will still lead to differences in trust. We will show how, despite agreeing on the ontological underpinnings of the concepts, there is the need to align trust so as to enable reliable *gossip*. With gossip we refer to all communication about trust.

We will emphasize the need to align trust further by considering a simple example of a multi-agent system with three agents.

Example. *Alice wants to know if Dave would be a good keynote speaker for the conference she is organizing. However, she does not know enough about him. She asks Bob. Bob has never collaborated with Dave directly, but they work at the same institute and play squash together. Through these interactions, Bob has trust in Dave and tells this to Alice.*

Lets analyse Bob's model. He does not know Dave professionally and bases his trust in Dave on personal interactions. This is a perfectly valid model, but

Alice’s model works differently: she only takes academic accomplishments into account. She should therefore disregard Bob’s gossip, because it is based on, what she considers, *unreliable* information. We emphasize that we differentiate between the trust she has in Bob and the reliability of the information he sends her. Her trust in Bob is grounded in her trust and reputation model. However, what we want to find out is whether gossip Bob sends can be interpreted reliably in Alice’s model.

The main question we address in this article is: *what information would be useful for agents to assess the reliability of gossip and what methods can be used for this assessment?*

In the next section we discuss related work on the issue and consider our views on the nature of trust and how we can ascertain the way other agents compute it.

2. Related Work and Our Approach

We are not the only ones to consider the communication between agents about trust as a problem and some work has been done in defining common ontologies for trust [4,5], however in practice these ontologies do not have the support of many of the different trust methodologies in development. Even if support were added for all systems and a common ontology emerged, we could still not use it to communicate effectively. Trust is an inherently personal phenomenon and has subjective components which cannot be captured in an ontology. An adaptable approach that takes the different agents’ points of view into account is needed.

Abdul-Rahman and Hailes’ reputation model [6] approaches the problem from another direction, by defining the trust evaluations based on the actual communications. The interpretation of gossip is based on previous interactions with the same sender. The problem with this, however, is that it is incomplete: firstly it assumes all other agents in the system use the same model, which in a heterogeneous environment will hardly ever be the case. Secondly, it uses a heuristic based on prior experiences, called the semantic distance, to “bias” received messages. The semantic distance is an average of all previous experiences. They do not differentiate between recommendations of different agents, which are based on different types of interactions.

We propose to enrich the model of communication by considering it separate from the actual trust model. By doing this, we can allow for different trust models. We note, however, that while trust is modeled in disparate ways, all definitions do agree on the fact that trust is a social phenomenon. Just as any social phenomenon, it arises from the complex relationships between the agents in the environment and, without losing generality, we say these relationships are based on any number of interactions between the agents. These interactions can have many different forms, such as playing squash with someone, buying a bicycle on eBay or telling Alice that Dave is a trustworthy keynote speaker. Note that not all interactions are perceived equally by all participants. Due to having different goals, agents may observe different things, or even more obviously: by having a different vantage point. Simply by having more (or different) information avail-

able, agents may perceive the interaction itself differently. In addition, interactions may be accompanied by some kind of social evaluation of the interaction. These can range from an emotional response, such as outrage at being cheated in a trade, to a rational analysis. Thus, we see that how an agent experiences an interaction is unique and personal. This only adds to the problem we are considering. To be able to align, there needs to be some common ground from which to start the alignment, but any agent’s experience of an interaction is subjective, and thus not shared. We call this personal interpretation of the interaction an *observation*. We say an agent’s observations allow it to evaluate trust.

Now that we have discussed what interactions mean to a single agent, we will return to the focus of communicating about trust. One interaction may be observed by any number of agents, each making different observations, which support different trust evaluations of different targets performing different roles. However, to communicate about trust evaluations, the agents need to have a starting point: some basic building blocks they implicitly agree they share. We note that the interactions provide precisely such a starting point. While all the agents’ observations are different, they do share one specific thing: *the interaction itself*. We therefore argue that to find a reliable alignment between two agents they can align based on these interactions.

Our approach uses these shared interactions as building blocks to align the agents’ trust models, based on the gossip they send each other. The gossip specifies certain interactions, which each agent observes differently. These observations form the support for an agent’s trust evaluation. If another agent communicates this trust evaluation, the interpretation should be based on the underlying interactions. An alignment of the trust models gives a way of doing this by gossiping about the agents’ trust evaluations and the observations (and thus interactions) they base these on.

3. Framework

Before we consider possible solutions we need a clear definition of the problem we are considering. Firstly we are considering agents with heterogeneous trust models, but we have no clear description of what a trust model is in the first place. Furthermore, to align the agents need to communicate. For this we need to define a language. And finally, the agents need to have some method of forming an alignment based on the statements in this language. Throughout this section we will illustrate the main definitions with an example. We will first give the basic scenario, which is a modified version of the example in the introduction:

Alice is organizing a conference and needs to invite a keynote speaker. She assigns the task of finding this person to her personal computational agent. It must contact the other agents in the system. The agent’s first choice is Bob. It sends Bob’s agent a message with an invitation to the conference, but he can’t make it. Instead his agent recommends Zack. However, Alice and Bob’s agents have never aligned their models and therefore Alice’s agent doesn’t know how to assess the reliability of this gossip. It asks Bob’s agent to start the alignment process. There are various other people who both Bob and Alice

have interacted with. Their agents contain knowledge about this and they gossip about these to form the alignment.

3.1. A formal representation of trust models

As argued in Section 2, interactions form the building blocks for talking about trust. While these interactions have a wealth of properties, we will start with the bare minimum we need to know to start the trust alignment. We define an interaction just as the set of observing agents. In practice we will use a more descriptive interpretation of what an interaction is. This we will introduce in Section 3.2.

We will denote with \mathfrak{I} the set of all interactions in the environment and with Ag the set of all agents. $\mathfrak{I}_{|A} \equiv \{\langle id, Ag \rangle \in \mathfrak{I} \mid A \in Ag\}$ the set of interactions observed by agent A ; id is a unique identifier for this interaction. This agent’s observations of these interactions form a separate set.

Definition 3.1 (Observation). An agent A ’s observation is given by the function $observe_A : \mathfrak{I}_{|A} \rightarrow \mathfrak{D}_A$, which associates each interaction $i \in \mathfrak{I}_{|A}$ with an observation $o \in \mathfrak{D}_A$. The set \mathfrak{D}_A is the entire set of observations of agent A and is in the form of the agent’s internal representation.

Example: Alice’s agent (Observations). *Alice’s agent stores the observations in its belief base. An example observation of an interaction where Bob and Dave play racquetball together.*

The observe-function simply maps the interaction itself (BD) onto a belief: $observe_{Alice}(BD) = (play_racquetball(Bob, Dave))$.

Because we will generally work with sets of observations, it is useful to extend the function $observe_A : \mathfrak{I}_{|A} \rightarrow \mathfrak{D}_A$ to $Observe_A : \mathcal{P}(\mathfrak{I}_{|A}) \rightarrow \mathcal{P}(\mathfrak{D}_A)$ such that for $I \subseteq \mathfrak{I}_{|A}$: $Observe_A(I) = \{o \in \mathfrak{D}_A \mid \exists i \in I : observe_A(i) = o\}$.

An agent A ’s observations *support* some trust evaluation. This is the essence of the trust model. As we argued in Section 2 there are many different computational trust models, but all compute trust evaluations, based on observations of interactions. We say these observations support the trust evaluation. These trust evaluations are statements in a language \mathcal{L}_{Trust} . The agents share the syntax of this language, but each agent can have its own semantics, defined by its trust model. While this shared syntax is not strictly necessary it makes the framework more comprehensible. Because the semantics will differ anyway, whether an agent calls reputation “reputation” or “jackhammer” doesn’t really matter from a conceptual point of view, each agent will have their own semantics and will have to distinguish between the syntax used by each agent anyway. It makes things easier on a computational level, however, if we can define a language in which agents agree on the syntax. even though they need to align the semantics. It also makes it easier to understand for humans, who will ultimately be directing agents equipped with the system. Thus this assumption may be just as important from the perspective of making the framework “explicable”.

\mathcal{L}_{Trust} is a standard predicate language, with one restriction: because trust is always about some target agent (all trust predicates have an object), we will only consider those predicates which give an evaluation of exactly one such target T :

$$\mathcal{L}_{Trust}[T] \equiv \{\varphi \in \mathcal{L}_{Trust} \mid \text{all predicates in } \varphi \text{ have target } T\}$$

We can now give a very abstract model of trust, which we can use as a basis for communication. We ground this framework in a mathematical model of information flow, introduced by Barwise & Seligman [9]. This is a very general model of how information flows and has been shown to be a good foundation for alignment [10] We use this same framework to formalize our earlier assertion that a trust model gives an evaluation of target agents *supported* by the agent’s observations.

Definition 3.2 (Trust model). A trust model \mathcal{M}_A of agent A is given as the tuple $\langle \mathcal{P}(\mathfrak{D}_A), \text{Evals}_A, \models_A \rangle$, with the following definitions:

- $\mathcal{P}(\mathfrak{D}_A)$ the power set of the observations of agent A .
- $\text{Evals}_A \subseteq \mathcal{L}_{Trust}$ the trust evaluations of agent A .
- \models_A a binary relation: $\models_A \subseteq \mathcal{P}(\mathfrak{D}_A) \times \text{Evals}_A$, such that if $O \subseteq \mathfrak{D}_A$ (in other words, $O \in \mathcal{P}(\mathfrak{D}_A)$) and $\varphi \in \text{Evals}_A$, then $O \models_A \varphi$ represents that for agent A , the trust evaluation φ is supported by O .

Example: Alice’s agent (Alice’s trust model). *Alice’s beliefbase contains the following observations: $\{\text{play_racquetball}(\text{Bob}, \text{Dave}), \text{co_author}(\text{Alice}, \text{Dave})\}$ and \mathcal{L}_{Trust} is the set of predicates formed by the predicate $\text{trustworthy}(T)$ and its negation, with $T \in \{\text{Alice}, \text{Bob}, \text{Dave}\}$. Some examples of support relations her trust model can have:*

$$\begin{aligned} \{\text{co_author}(\text{Alice}, \text{Dave})\} &\models_{\text{Alice}} \text{trustworthy}(\text{Dave}) \\ \{\text{play_racquetball}(\text{Bob}, \text{Dave})\} &\models_{\text{Alice}} \neg \text{trustworthy}(\text{Dave}) \end{aligned}$$

This framework models the entire space of potential observations and their supported trust evaluations. However, in practice only one of these evaluations will be the actual evaluation of the target, namely the one supported by the observations in the actual state of the agent. However, to assess gossip it is important to base this on a larger amount of data than just the real trust evaluation.

We will now use this formal model to define a channel, following the mathematical model put forth in [9]. The trust alignment between two agents is grounded in this channel, which models the relation of two different trust models.

3.2. Formalizing gossip

Now that we have described trust models algebraically, we can focus on the formal model of how to assess each others’ gossip. To do so, the agents should form an alignment and to do this, the agents A and B establish some set of shared interactions $\mathfrak{I}_{|AB} = \mathfrak{I}_{|A} \cap \mathfrak{I}_{|B}$, the set of interactions they have both observed. This results in subsets of observations for both agents: $\mathfrak{D}_{A|B} \subseteq \mathfrak{D}_A$ are A ’s observations of the interactions shared with B . These are observations based on interactions they know they’ve shared. Due to the assumption that all the observers know the other observers in interactions, each agent can find this set. We justify this assumption by noting that even if this is not the case a priori, the set of shared interactions is easy to establish through prior communication.

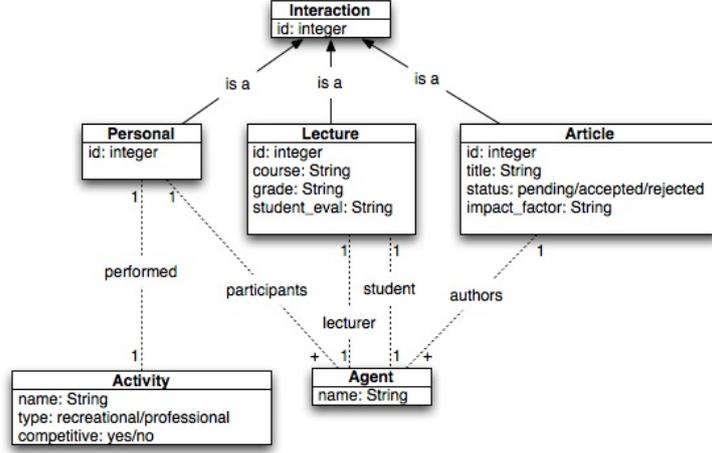


Figure 1. The ontology for an example \mathcal{L}_{Domain} , in a UML-like representation

The agents can now talk about their trust evaluations based *only* on the interactions they *both* observed to come to their evaluations. To talk about trust, they can use \mathcal{L}_{Trust} , but we have so far not specified how to talk about the interactions or observations. To do so, we introduce another, separate, language: \mathcal{L}_{Domain} . This should be a domain dependent language in which the agents can choose to relay objective properties of the interactions they have used to form their trust evaluations.

Example: Alice’s agent (\mathcal{L}_{Domain}). *An example of a domain language in which agents can talk about interactions is the one given in Figure 1. Playing racquetball can be modeled as an activity which is performed in a personal interaction. The other interactions that can be talked about are of a more professional nature.*

The agents align by gossiping about different targets: communicating their trust evaluations of a target in \mathcal{L}_{Trust} and about the interactions these evaluations are based on in \mathcal{L}_{Domain} . Gossip sent by an agent B is defined as the tuple $\langle T, \beta, \psi \rangle$, with $\beta \in Eval_{s_B}[T]$ and $\psi \in \mathcal{L}_{Domain}$. We recall that $Eval_{s_B}[T] \subseteq \mathcal{L}_{Trust}[T]$ and thus gossip between two agents consists of a part about trust and a part about the agents’ observations which support this trust predicate. This ψ can be used to pinpoint what interactions comprise I , the set of interactions that B used to compute β . Agent A uses I to find its own trust evaluation $\alpha \in Eval_{s_A}[T]$, such that $Observe_A(I) \models_A \alpha$, which gives us the basis for a targeted combined trust model: a shared set of interactions I supporting a trust evaluation for both A and B with regards to target T . It is possible that in some situations $Observe_A(I)$ will not support any trust evaluation for agent A . In this case we cannot use the related gossip in forming the targeted alignment. However, due to the requirement that the agents only take shared interactions into account this should not happen often.

Definition 3.3 (Targeted combined trust models). A targeted combined trust model has the same structure as a trust model, defined in Definition 3.2. It is the

tuple $\langle \mathcal{P}(\mathfrak{I}_{|AB}), (Evals_A[T] \cup \mathcal{L}_{Domain} \times Evals_B[T] \cup \mathcal{L}_{Domain}), \models_{AB} \rangle$
with \models_{AB} a binary relation:

$\models_{AB} \subseteq \mathcal{P}(\mathfrak{I}_{|AB}) \times ((Evals_A[T] \cup \mathcal{L}_{Domain}) \times (Evals_B[T] \cup \mathcal{L}_{Domain}))$, such
that if $I \subseteq \mathfrak{I}_{|AB}$ and $\langle \alpha, \beta \rangle \in (Evals_A[T] \cup \mathcal{L}_{Domain} \times Evals_B[T] \cup \mathcal{L}_{Domain})$,
then $I \models_{AB} \langle \alpha, \beta \rangle$ if and only if $Observe_A(I) \models_A \alpha$ and $Observe_B(I) \models_B \beta$.

Example: Alice's agent (Combined trust model with regards to Dave). *We recall the interaction BD and Alice's observation $observe_{Alice}(BD)=play_racquetball(Bob,Dave)$. Furthermore, Alice's trust model contains the relation:*

$\{play_racquetball(Bob,Dave)\} \models_{Alice} \neg trustworthiness(Dave)$. Lets assume Bob has a similar observation: $observe_{Bob}(BD)=good_racquetball_match(Bob,Dave)$ and the relation: $\{good_racquetball_match(Bob,Dave)\} \models_{Bob} trustworthiness(Dave)$

Then the combined trust model for Dave contains the relation (with the agent's names in subscript for clarity):

$BD \models_{Alice,Bob} (\neg trustworthiness_{Alice}(Dave), trustworthiness_{Bob}(Dave))$.

Additionally, the model could contain further information communicated by Bob about the interaction in \mathcal{L}_{Domain} . Thus the relation could look like:

$BD \models_{Alice,Bob} (\neg trustworthiness_{Alice}(Dave), trustworthiness_{Bob}(Dave) \wedge racquetball_match(Bob,Dave))$.

Note that neither agent knows everything, just those parts of the model which are gossiped about. From these partial models they must extrapolate the underlying model, so as to form an alignment. The rules we have such that $I \models_{AB} \langle \alpha, \beta \rangle$ in the targeted cobined trust models form the basic building blocks of this alignment.

Definition 3.4 (Targeted alignment). Given a combined trust model for agents A and B with regards to target T , we define \Rightarrow_A^T as a binary relation $\Rightarrow_A^T \subseteq Evals_B[T] \cup \mathcal{L}_{Domain} \times Evals_A[T] \cup \mathcal{L}_{Domain}$, with $\Gamma[T] \Rightarrow_A^T \Delta[T]$ such that:

$\Gamma[T] \subseteq Evals_B[T] \cup \mathcal{L}_{Domain}$ where $\Gamma[T]$ means all trust predicates
in Γ have target T

$\Delta[T] \subseteq Evals_A[T] \cup \mathcal{L}_{Domain}$

$\exists I \subseteq \mathfrak{I}_{|AB} : \forall \gamma \in \Gamma[T] : \exists \delta \in \Delta[T] : I \models_{AB} \langle \gamma, \delta \rangle$

$\mathfrak{T}_A[T] = \{ \langle \Gamma[T], \Delta[T] \mid \Gamma[T] \Rightarrow_A^T \Delta[T] \}$ is agent A 's targeted alignment with target T . We call $\Gamma[T] \Rightarrow_A^T \Delta[T]$ a rule in this targeted alignment. The relation \Rightarrow_A^T is not symmetrical, while the combined trust model is. It therefore stands to reason there is a similar targeted alignment $\mathfrak{T}_B[T]$ with its binary relation \Rightarrow_B^T .

A targeted alignment can be interpreted as the set of relations between an agent's own model and the communication partner's model with regards to some specific target. Each rule states that if there is a set of interactions I which support all trust evaluations by agent B as well as all statements in \mathcal{L}_{Domain} about the interactions, then agent A has a trust evaluation with corresponding statements in \mathcal{L}_{Domain} which is also supported by I . The semantics are given by the agent-independent semantics of \mathcal{L}_{Domain} and each agent's semantics of

\mathcal{L}_{Trust} , however because agent A doesn't know the semantics of $Evals_B$, we can replace this with an unknown. We suppose that agent B has semantics supporting his trust evaluation, or otherwise he wouldn't have communicated it. We consider agents gossiping untruthful information as outside the scope of this work.

Example: Alice's agent (Aligning about Dave). *The rule in the alignment with regards to Dave, based on the relation in above would be:*

$$racquetball_match(Dave) \wedge trustworthy_{Bob}(Dave) \Rightarrow_{Alice}^{Dave} \neg trustworthy_{Alice}(Dave)$$

3.3. Generalization and coverage

Now that we have a way of describing the relationship between two agents' trust models with regards to a specific target, we wish to expand this idea to encompass multiple targets. We consider this problem as an inductive learning problem [11]. Given a number of targeted alignments with regards to different agents, is there an alignment that describes all (or most) of them?

To use inductive learning, we need to define what our solution should look like. This is an untargeted alignment: similar to the targeted alignment in Definition 3.4, but not restricted to just one target. A natural way of forming an untargeted alignment is by simply replacing all instances of the target agent in a targeted alignment with a free variable. In general we will say an untargeted alignment is a θ -subsumption of one or more targeted alignments. We introduce the notion of coverage to specify which targeted alignments these are.

Definition 3.5 (Coverage of alignments). For an agent A , we say an alignment \mathfrak{T}_A covers a targeted alignment $\mathfrak{T}_A[T]$, if for every rule $\Gamma[T] \Rightarrow_A^T \Delta[T] \in \mathfrak{T}_A[T]$, there is a rule $\Gamma \Rightarrow_A \Delta \in \mathfrak{T}_A$, such that Γ θ -subsumes $\Gamma[T]$ and Δ θ -subsumes $\Delta[T]$ for some θ . We introduce the function \mathbf{c} which returns the set of targeted alignments covered by a given untargeted alignment.

We can now use inductive learning to find a trust alignment that covers all the targeted alignments. The way to do this is by structuring the search space. We do this with the generality relationship.

Definition 3.6 (Generality relation). We say an alignment \mathfrak{T} is more general than an alignment \mathfrak{T}' iff $\mathbf{c}(\mathfrak{T}) \supseteq \mathbf{c}(\mathfrak{T}')$. We write this: $\mathfrak{T} \succeq \mathfrak{T}'$. If $\mathbf{c}(\mathfrak{T}) \supset \mathbf{c}(\mathfrak{T}')$ we say \mathfrak{T} is strictly more general and write $\mathfrak{T} \succ \mathfrak{T}'$

The overall trust alignment between two agents can now be found by finding a minimally general generalization, which covers all targeted alignments.

Definition 3.7 (General trust alignment). The trust alignment \mathfrak{T}_A^* of an agent A with another agent is a *minimally general generalization* of all the targeted alignments: $\forall T \in Targets : \mathfrak{T}_A[T] \in \mathbf{c}(\mathfrak{T}_A^*)$. A minimally general generalization means, that if there is any other alignment \mathfrak{T}'_A that covers all targeted alignments, then: $\mathfrak{T}'_A \succeq \mathfrak{T}_A^*$.

Example: Alice's agent (Trust alignment with Bob). *If Alice and Bob only gossip about Dave, with the targeted alignment above as a result, the general alignment*

could look something like this:

$$\text{racquetball_match}(X) \wedge \text{trustworthy}_{\text{Bob}}(X) \Rightarrow \neg \text{trustworthy}_{\text{Alice}}(X)$$

This example has necessarily been very simplistic and therefore this result seems trivial. We base this alignment on only one interaction about one single agent. The generalization in this case is just the skolemization of the targeted alignment. From this we learn that if Bob bases his evaluation “trustworthy” of a target agent on an interaction where they played racquetball together, Alice should consider this agent as “ \neg trustworthy”. However, this is a good start: next time Bob recommends a possible keynote speaker based on his racquetball games, Alice knows that she should take this to mean the opposite. However, if Bob recommends someone based on their joint experience in authoring papers, this rule says nothing about this. The alignment is not yet complete and there is no rule covering this type of interaction. In real situations the alignment will be based on multiple interactions concerning multiple agents and the resulting rules will be more robust. Furthermore, $\mathcal{L}_{\text{Trust}}$ will be more realistic than agents just being “trustworthy” or “ \neg trustworthy”, making the model far richer.

4. Analysis and conclusions

We have detailed a formal framework for aligning the trust of two agents. We do this by gossiping about a variety of different agents and inducing a general model. As a proof of concept we implemented a more detailed version of the example using Prolog and Aleph [12] which found a good alignment. This showed that also in practice, 2 computational agents with different interpretations of what trust is, can pass relevant advice to each other about the trustworthiness of others, as long as they have:

- A language in which to communicate the structure of their trust model, in a shared syntax
- A domain dependent language in which to communicate which interactions their trust is supported by
- A method of inducing a more general model from the passed gossip

In addition, it is an extensible model, which can be refined when more gossip is received. The explained framework forms the foundation of this alignment: it pinpoints which communication is necessary and gives a general model in which the alignment can be formed.

In future work we plan on giving an implementation of the model. We can also extend the model in various directions. The role an agent plays can be important in the trust model and this is not taken into account in our model. Additionally we assume agents always tell the truth. This does not necessarily have to be the case. Reasoning about the validity of agents’ information is another interesting extension of the research.

Acknowledgements

This work is supported by the Generalitat de Catalunya under the grant *2009-SGR-1434*, the Agreement Technologies Project *CONSOLIDER CSD2007-0022*, *INGENIO 2010* and the LiquidPub Project *CIT5-028575-STP*. M. Schorlemmer is supported by a *Ramón y Cajal* research fellowship from Spain's Ministry of Science and Innovation, which is partially funded by the European Social Fund.

References

- [1] Ramchurn, S.D., Huynh, D., Jennings, N.R.: Trust in multi-agent systems. *The Knowledge Engineering Review* **19**(1) (2004) 1–25
- [2] Omidyar, P.: Ebay. <http://www.ebay.com>, retrieved September 26, 2008 (1995)
- [3] Sabater-Mir, J., Paolucci, M., Conte, R.: Repute: REPUTation and imAGE among limited autonomous partners. *JASSS - Journal of Artificial Societies and Social Simulation* **9**(2) (2006)
- [4] Pinyol, I., Sabater-Mir, J.: Arguing about reputation. the Irep language. In: *Proceedings of the 8th Annual International Workshop "Engineering Societies in the Agents World" (ESAW'07)*. Volume 4995. Springer LNCS (2007) 284–299
- [5] Casare, S., Sichman, J.: Towards a functional ontology of reputation. In: *AAMAS '05: Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, New York, NY, USA, ACM (2005) 505–511
- [6] Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. *Proceedings of the 33rd Hawaii International Conference on System Sciences* **6** (2000) 6007
- [7] R., F., C., C.: *Social Trust: A Cognitive Approach*. In: *Trust and Deception in Virtual Societies*. Kluwer Academic Publishers (2001) 55–90
- [8] Hübner, J.F., Lorini, E., Herzig, A., Vercouter, L.: From cognitive trust theories to computational trust. In: *Proceedings of the 12th International Workshop on Trust in Agent Societies*, Budapest, Hungary, 10/05/2009-11/05/2009. (In Press)
- [9] Barwise, J., Seligman, J.: *Information Flow: The Logic of Distributed Systems*. Cambridge University Press (1997)
- [10] Schorlemmer, M., Kalfoglou, Y., Atencia, M.: A formal foundation for ontology-alignment interaction models. *International Journal on Semantic Web and Information Systems* **3**(2) (2007) 50–68
- [11] De Raedt, L.: *Logical and Relational Learning*. Springer Verlag (2008)
- [12] Koster, A., Sabater-Mir, J., Schorlemmer, M.: An interaction-oriented model of trust alignment. Technical Report TR-2009-04, CSIC-III A (2009) <http://www2.iii.a.csic.es/~andrew/files/techreport2.pdf>.