Effects of expressiveness and heterogeneity of reputation models in the ART testbed: Some preliminary experiments using the SOARI architecture

Luis G. Nardin[1], Guillaume Muller[1], Anarosa A. F. Brandão[1], Laurent Vercouter[2], and Jaime S. Sichman[1]

[1] Laboratório de Técnicas Inteligentes - EP/USP
Av. Luciano Gualberto, 158 – trav. 3
05508-900 – São Paulo – SP – Brasil
{luis.nardin,anarosa.brandao,jaime.sichman}@poli.usp.br,
Guillaume.Muller@freesurf.fr

[2] École Nationale Supérieure des Mines de Saint-Étienne
158, cours Fauriel, 42023 Saint-Étienne Cedex 2, France
Laurent.Vercouter@emse.fr

Abstract. Trust and reputation have proved to help protect societies against harmful individuals. Inspired by these principles, many computational models have been and new ones continue to be proposed in the literature to protect multi-agent systems. In an open system, where few assumptions can be made on the internals of the agents, it is possible that different agents use different trust and reputation models. Since agents have to exchange information to make their trust and reputation models more robust, and since the models use different internal concepts and metrics, it is very important to consider the interoperability of these models. Based on experiments, this paper illustrates the usefulness of the SOARI architecture, which allows heterogeneous agents to interoperate more expressively about reputation.

1 Introduction

Agents present the capabilities of both acting autonomously and engaging in social activities. In open environments, where agents can enter or leave the environment at any time, taking part in such social activities may expose them to risks, for instance, when taking decisions based on information provided by malevolent agents. In order to avoid such risks, solutions based on trust models where implemented [5, 17, 6, 14, 13, 10]. Most of these models are based on the concept of reputation.

In order to accelerate the reputation evaluation and to improve the robustness of their reputation models, the agents generally exchange information about the reputation of third parties. However, since there is no consensus about a single unifying reputation definition, the semantics associated with reputation differs from one model to another. This semantic heterogeneity raises an interoperability problem among existing reputation models, which is addressed by SOARI [11] architecture.

In this paper, we present results of experiments where SOARI is used to enable interoperability of two reputation models: Repage [13] and L.I.A.R. [10]. These experiments evaluate the impact that the reputation models interoperability may cause

on agents evaluation accuracy. More specifically, this paper answers the following two questions: (1) is there any improvement in the reputation evaluation accuracy when enabling a more expressive communication? (2) How does the heterogeneity influence the evaluation accuracy of the dishonest agents' reputation?
The rest of the document is organised as follows. Section 2 presents briefly the platforms used to run the experiments (ART and FOREART testbeds) as well as the SOARI architecture. In Section 3, the results and analysis of the experiments are shown. Finally, our conclusions and future work are presented in Section 4.

2 Background Work

The ART testbed (Agent Reputation and Trust testbed) [7] is currently the unique platform freely available to perform benchmarks with heterogeneous reputation models. We first briefly present its scenario, because it is the basis for the experiments. However, this platform does not allow the agents to communicate about reputation using their distinct semantically reputation model concepts, thus losing expressiveness. The FOREART testbed [3], which is an extension of the ART testbed, allows a more expressive communication among the agents. In order to reach this goal, this latter platform uses FORE (Functional Ontology of Reputation) [4] as a common vocabulary. In this platform, interoperability is obtained by translating concepts from a source model (expressed in ontological terms) to concepts of FORE, and then by translating the result from FORE into concepts of a target model (also expressed in ontological terms). The SOARI architecture is then used to implement the FOREART testbed's agents thus enabling a more expressive communication about reputation among them. The resulting platform is the basis of the experiments described in the next section.

2.1 The ART testbed

In AAMAS'04 TRUST workshop, it was admitted that the diversity in the internals and metrics employed by current models of trust and reputation made it difficult to establish objective benchmarks. In order to design a testbed platform to enable comparison, the ART testbed initiative was launched.
The resulting testbed platform (programmed in Java) simulates an art appraisal game, where agents evaluate paintings for clients and gather opinions and reputations from other agents to produce accurate appraisals. More precisely, a game proceeds as a series of the following time steps[3]: (i) the platform assigns clients (i.e. paintings) to each appraiser. Appraisers receive larger shares of clients (thus larger amount of money) if they have produced more accurate appraisals in the past; according to the era each painting belongs to, an appraiser is more or less accurate in its evaluations; (ii) reputation transactions occur, where appraisers can exchange reputation information about third parties for given eras; (iii) certainty transactions occur, where appraisers can exchange

[3] Those time steps refer to ART testbed platform used on the competition of 2008, which implements slightly different time steps sequence than the one of the previous years' platform, described in [8].

how certain they are about a specific era; (iv) opinion transactions occur, where appraisers can exchange expert opinions about a specific painting; (v) finally, the appraisers are required to send weights to the platform; those weights represent the intensity with which the appraiser considers the opinion of each other appraiser; the platform then computes the final appraisal of each appraiser as a weighted mean of the opinions it has purchased; this step ensures the same computation for everybody, therefore, (a) only the trust models are evaluated (not the expertise in art) and (b) cheating is impossible. The winner of the game is the agent that has the higher final bank balance.
In this scenario, the need for reputation modelling comes from the duality of the need for cooperation to evaluate some of the paintings (because the agents are only competent in some eras) and the competition to earn the biggest part of the client pool.
More details about the ART testbed can be found in [8].

2.2 The FOREART testbed

The interaction which involves the reputation transaction is a moment where agents have to exchange information from their reputation models, meaning that interoperability among reputation models is required. In the current version of the platform, interoperability is obtained by asking the developers of each agent to map their reputation model evaluations into a single value in the domain [0:1]. This common model is too simple and the mapping of complex internal reputation models into a simplistic one results in loss of expressiveness and details. It is thus impossible to perform finer agent interactions about reputation.
The addition of semantic data to this common model may improve the agent performance during the process of reputation building, while allowing interoperability between different reputation models. Therefore, the FOREART testbed platform was implemented as an extension of ART by modifying its engine to allow the exchange of messages related to reputation transactions that involve semantic content. The messages' content is a string (instead of couples *(agent, painting era)* or numerical value) and it is expected that this string is queries and answers written in an ontology query language. The chosen query language is related to the inference engine that is used to reason about the queries. The first version of FOREART uses nRQL [9] and Racer [12]. Nonetheless, FOREART agents were implemented according to the general agent architecture proposed to support reputation interaction with semantic content [16]. The general architecture main modules are the Interaction Module (IM), the Reputation Mapping Module (RMM) and the Reputation Reasoning Module (RRM). These modules are responsible for dealing with the translation between FORE and the agent internal reputation model expressed as ontology, and the reasoning about exchanged messages.
More information on this platform can be found in [15, 16, 3, 2].

2.3 SOARI: Service Oriented Architecture for Reputation Interaction

However, because of some drawbacks of the general agent architecture [16], the SOARI architecture was proposed (Figure 1). The SOARI is a service-oriented architecture to support the semantic interoperability among agents that implement heterogeneous reputation models. The main underlying idea of SOARI is that the mapping between different ontologies (by using FORE as an interlingua) may be realised off-line, and be available on-line as a service for the agents that use the same reputation model. Hence, it extends the FOREART agent architecture in two ways: (i) it subdivides the Reputation Mapping Module (RMM) in two distinct and specialised modules: the Ontology Mapping Service (OMS) and the TRANSLATOR module (in grey in the figure), and (ii) it performs the ontology mapping and translation functions as a service outside the agent architecture.

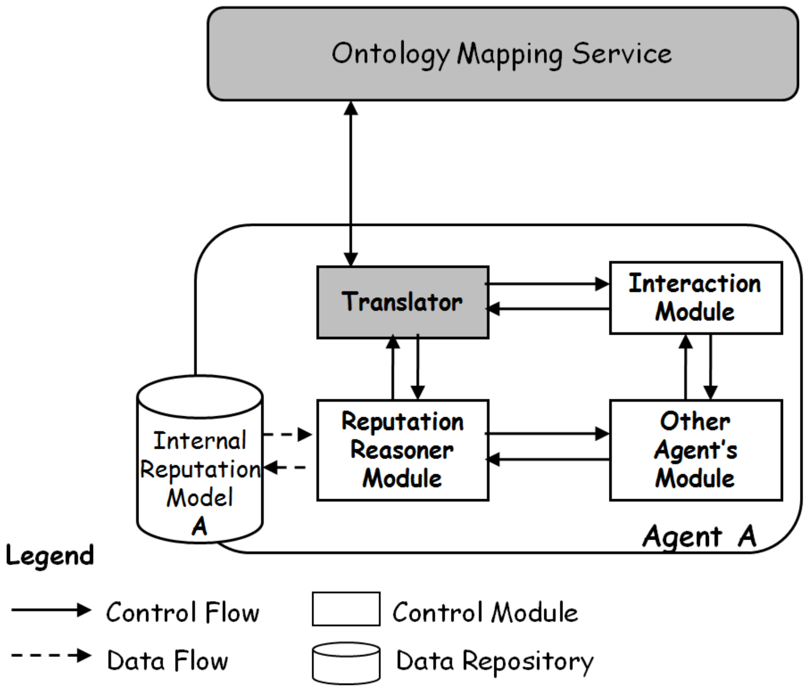

Fig. 1. Service Oriented Architecture for Reputation Interoperability

The OMS module is a service outside the agent that implements the mapping and translation ontology functions and presents two main functionalities: (i) to map concepts from the target's reputation model ontology to the concepts of the common ontology; and (ii) to answer concept translation requests from the TRANSLATOR module.
The TRANSLATOR module resides inside the agent and it translates reputation messages. It has four main activities: (i) to translate the reputation messages from the common ontology to the target agent's reputation model ontology whenever the message comes from the Interaction Module (IM); (ii) to translate the reputation messages from

the agent's reputation model ontology to the common ontology whenever the message is sent to IM; (iii) to trigger some function in the Reputation Reasoning Module (RRM) based on the interpretation of messages written using the reputation model ontology; and (iv) to create a message using the reputation model ontology whenever requested by RRM.
More information on this architecture can be found in [11].

3 Experiments

This section intends to answer two questions: (1) is there any improvement in the accuracy of the agents' reputation evaluation when enabling more expressive communication about reputation? (2) how does the heterogeneity influence the accuracy of dishonest agent's reputation evaluation?
In order to answer those questions, some experiments were performed using the ART and FOREART testbeds and the SOARI architecture. In those experiments, one agent deliberately lies about the other agents' reputation and about paintings evaluation. The analysis was performed to determine how accurate the other agents are in the evaluation of the reputation of the liar agent.
In a practical point of view, all the experiments were performed using a modified FOREART testbed. In the remaining, the term ART thus refers to situations when the reputation communication among the agents is limited to numeric (numeric communication) and FOREART when it is performed using strings (symbolic communication).
The experiments include two types of agents: *Honest* and *Dishonest*. The *Honest* agents answer to the requests only when they have expertise about the requested painting era and with information coherent to their internal state. The *Dishonest* agents answer to all the requests, even when they do not have expertise about that painting era and they never answer the requests with information coherent to their internal state.

3.1 Agent Model

The agent models in the testbed platforms are implemented by extending the abstract *Agent* class and filling up the methods that describe the agent's behaviour [8]. These methods correspond globally to the steps described on section 2.1.
In the begin of each time step, a set of paintings is assigned to the agent for appraisal. For each painting assigned, the agent performs reputation transactions. First, it requests to other agents in the testbed platform the reputation of possible appraisers of that painting. Then, it answers to reputation requests received from other agents. If it is a *Dishonest* agent, it accepts all the requests. Otherwise, it accepts the requests only if it has the expertise higher than a predefined expertise threshold ($expertisethreshold = 0.7$). To all the accepted requests, the agent answers with a reputation value, which does not reflect its internal reputation evaluation if it is a *Dishonest* agent.
After performing the reputation transaction, the agent performs certainty transactions. It first selects a group of agents and requests to them their certainty about a specific painting era. In the sequence, it answers to certainty requests received from other agents. If

it is an *Honest* agent and its expertise is higher than a predefined expertise threshold ($expertisethreshold = 0.7$), it answers with its expertise value. However, if it is a *Dishonest* agent, it answers with the maximum between 1 and its expertise value plus 0.5.

After performing the certainty transactions, the agent requests the opinion of the agents it trusts (i.e. which reputation value is higher than a trust threshold that in Repage is $Image >= 0.5$ and/or $Reputation >= 0.8$, and L.I.A.R. is $X >= 0.7$, where $X = \{DIbRp,\ IIbRp,\ ObsRcbRp,\ EvRcbRp\ or\ RpRcbRp\}$) or the agents from which it received a certainty value higher than a predefined certainty threshold ($certaintythreshold = 0.5$).

Finally, in order for the simulator to compute the opinions, the agents provide to it the weight of each opinion provided by the other agents.

3.2 Experiments Description

The main objective of these experiments was to identify the mean value of the reputation assigned by the *Honest* agents to the *Dishonest* agent. In order to enable comparison between the experiments, the initial painting era knowledge and clients distribution were identical in all the experiments. Moreover, all the agents used the same configuration parameters (Table 1) and agent model (see Section 3.1) in all the simulations.

To reach this goal, we considered the execution of 10 simulations ($p = 10$) for each experiment with 100 cycles each. Each simulation was composed of 11 agents ($n = 11$), where 10 agents were *Honest* and 1 agent was *Dishonest* ($i = [1, 10]$ and $j = 11$). The mean value of the reputation assigned to the *Dishonest* agent by each *Honest* agent (r_j) considered only the value obtained in the last simulation cycle ($l = 100$ and $m = 100$). The value of the last simulation cycle was used because we considered it the most accurate reputation evaluation.

Table 1. Testbed's configuration parameters

Parameter	Value
averageClientsPerAgent	4
numberOfPaintingEras	20
cp_opinionCost	10
cp_certaintyCost	2
f_clientFee	100
nb_certaintyMsg	20
nb_opinionMsg	5

Formally, consider a set of n agents, where $i = \{1, 2, \ldots, n-1\}$ are *Honest* agents and $j = n$ is a *Dishonest* agent. Moreover, consider that r_{ij}^{sk} is the reputation value assigned by the agent i to the agent j in cycle k on simulation s. Typically, the reputation value assigned by agent i to agent j on simulation s corresponds to the mean reputation value of a set of cycles. Thus, $r_{ij}^{s} = \dfrac{\sum_{k=l}^{m} r_{ij}^{sk}}{m-l+1}$, where l and m represents, respectively, the lower and upper cycle limits. The mean reputation value assigned by the *Honest* agents to the *Dishonest* agent on simulation s is $r_{j}^{s} = \frac{\sum_{i=1}^{n-1} r_{ij}^{s}}{n-1}$. Finally, given a set of simulations $s = 1, \ldots, p$ that compose an experiment, the mean value of the *Dishonest* agent is $r_j = \frac{\sum_{s=1}^{p} r_{j}^{s}}{p}$.

The experiments performed were classified based on two dimensions: (1) reputation models used by the agents in the experiment (Repage, L.I.A.R. or both), and (2) reputation communication method (numeric or symbolic) (Table 2). Moreover, the mixed experiments are split in two others based on the reputation model of the *Dishonest* agent. This distinction is indicated by the *D/L.I.A.R.* and *D/Repage* suffix in the experiment's name. In the other experiments, the *Dishonest* agent uses the same reputation model than the *Honest* agents.

Table 2. Summary of experiments

ID	Experiment name	Reputation Model	Reputation Communication
exp1	ART/L.I.A.R.	L.I.A.R.	Numeric
exp2	ART/Repage	Repage	Numeric
exp3.1	ART/Mixed-D/L.I.A.R.	L.I.A.R. and Repage	Numeric
exp3.2	ART/Mixed-D/Repage	L.I.A.R. and Repage	Numeric
exp4	FOReART/L.I.A.R.	L.I.A.R.	Symbolic
exp5	FOReART/Repage	Repage	Symbolic
exp6.1	FOReART/Mixed-D/L.I.A.R.	L.I.A.R. and Repage	Symbolic
exp6.2	FOReART/Mixed-D/Repage	L.I.A.R. and Repage	Symbolic

3.3 Experiments Results and Analysis

Here, we present an analysis of the results obtained from the experiments in order to answer the two questions posed at the beginning of this section. The complete raw results data can be obtained at http://www.lti.pcs.usp.br/results.pdf. The analysis methodology used to answer the questions raised on this section is based on the Student's T-Test [1].

The analysis performed on this section was based on the L.I.A.R. and Repage reputation models attributes. For reputation model attribute, we mean the different concepts of reputation defined in each reputation model. The L.I.A.R. reputation model defines five different types of reputation: *Direct Interaction-based Reputation* (DIbRp); *Indirect Interaction-based Reputation* (IIbRp); *Observation Recommendation-based Reputation* (ObsRcbRp); *Evaluation Recommendation-based Reputation* (EvRcbRp); and *Reputation Recommendation-based Reputation* (RpRcbRp). Further details about those types of reputation can be obtained in [10].
The Repage reputation model defines two reputation concepts: *Image* and *Reputation*. Further details about those can be obtained in [6].

Effect of the expressiveness of communication. In order to analyse the effects of the more expressive communication, it was verified if the mean value of the *Dishonest* agent's attributes (r_j) obtained on the numerical experiments (ART experiments) were higher than the similar ones obtained on the symbolic experiments (FOREART experiments). If so, then it means that the *Dishonest* agent was better identified in the symbolic experiments than in the numerical experiments. Thus, using Student's T-Test, a set of hypotheses was required to demonstrate it. The general form of the hypotheses is:

The mean value of the reputation model attribute from ART experiments is higher than the same attribute's mean value from the FOREART experiments. This hypothesis, from the point of view of the reputation model attribute is expressed mathematically as $Q^X_{ART} > Q^X_{FOReART}$, where X is a L.I.A.R. or Repage reputation model attribute.
In order to validate this hypothesis using the Student's T-Test, the following test is performed:
$H0 : Q^X_{ART} <= Q^X_{FOReART}$
$H1 : Q^X_{ART} > Q^X_{FOReART}$

The complete set of hypotheses to demonstrate the effects of the more expressive communication are presented on Table 3.

Table 3. Expressiveness hypotheses

Hypothesis	Reputation Model	Attribute
A	L.I.A.R.	DIbRp
B	L.I.A.R.	IIbRp
C	L.I.A.R.	RpRcbRp
D	Repage	Image
E	Repage	Reputation

When applied to the results of the following pairs of experiments: (exp1, exp4), (exp2, exp5), (exp3.1, exp6.1), (exp3.1, exp6.2), (exp3.2, exp6.1) and (exp3.2, exp6.2), considering the risk level (α) of 0.01 and the degree of freedom of 18, those hypotheses generate the results presented in Table 4 (✔ means that $H0$ was rejected, which confirms the hypothesis; ✘ means that $H0$ was not rejected, thus the hypothesis cannot be confirmed; and $-$ (dash) means that the hypothesis is not applicable for the pair of experiments).

Table 4. Expressiveness hypotheses result

Pair	**Hypotheses**				
	A	**B**	**C**	**D**	**E**
(exp1, exp4)	✘	✘	✘	-	-
(exp2, exp5)	-	-	-	✔	✔
(exp3.1, exp6.1)	✘	✘	✘	✔	✔
(exp3.1, exp6.2)	✘	✘	✘	✔	✔
(exp3.2, exp6.1)	✘	✘	✘	✘	✔
(exp3.2, exp6.2)	✘	✘	✘	✘	✔

Analysing the information in Table 4, we can verify that in most of the cases the hypotheses D and E reject the $H0$ (indicated by ✔) confirming those hypotheses, while the hypotheses A, B and C do not (indicated by ✘). From the reputation model point of view, the hypotheses D and E are associated to the Repage reputation model ($Image$ and $Reputation$ attributes), while the hypotheses A, B and C are associated to the L.I.A.R. reputation model ($DIbRp$, $IIbRp$ and $RpRcbRp$ attributes). Therefore, we can conclude that a more expressive communication about reputation has a positive effect in the accuracy of the reputation evaluation to agents that use the Repage reputation model. However, it was not possible to infer that the more expressive communication benefits or harms the agents that use the L.I.A.R. reputation model.
Based on these results, we conclude that the Repage reputation model has some intrinsic or some implementation characteristics that enables it to benefit from the more expressive communication.

Effect of the reputation model heterogeneity. The analysis of the effect of reputation model heterogeneity was performed by testing if the mean value of the *Dishonest* agent's reputation model attributes (r_j) obtained on experiments with homogeneous reputation model were higher than the similar ones obtained on mixed experiments. Thus, to demonstrate it using Student's T-Test a set of hypotheses was required. The general form of the hypotheses is:

The mean value of the reputation model attribute from experiments with homogeneous reputation model is higher than the same attribute's mean value from mixed experiments. This hypothesis, from the point of view of the reputation model attribute is expressed mathematically as $Q^X_{P/M} > Q^X_{P/Mixed}$, where M is the reputation model (L.I.A.R. or Repage), X is its attribute and P is the testbed platform (ART or FOREART).

In order to validate this hypothesis using the Student's T-Test, the following test is performed:

$H0 : Q^X_{P/M} <= Q^X_{P/Mixed}$

$H1 : Q^X_{P/M} > Q^X_{P/Mixed}$

The complete set of hypotheses to demonstrate the effects of heterogeneous reputation models are presented on Table 5.

Table 5. Heterogeneous hypotheses

Hypothesis	Reputation Model	Attribute	Platform
F	L.I.A.R.	DIbRp	ART
G	L.I.A.R.	IIbRp	ART
H	L.I.A.R.	RpRcbRp	ART
I	Repage	Image	ART
J	Repage	Reputation	ART
K	L.I.A.R.	DIbRp	FOREART
L	L.I.A.R.	IIbRp	FOREART
M	L.I.A.R.	RpRcbRp	FOREART
N	Repage	Image	FOREART
O	Repage	Reputation	FOREART

When applied to the results of the following pairs of experiments: (exp1, exp3.1), (exp1, exp3.2), (exp2, exp3.1), (exp2, exp3.2), (exp4, exp6.1), (exp4, exp6.2), (exp5, exp6.1) and (exp5, exp6.2), considering the risk level (α) of 0.01 and the degree of freedom of 18, those hypotheses generate the results presented in Tables 6 and 7 (✔ means that $H0$ was rejected, which confirms the hypothesis; ✘ means that $H0$ was not rejected, thus the hypothesis cannot be confirmed; and $-$ (dash) means that the hypothesis is not applicable for the pair of experiments).

Table 6: Hypotheses result ART

Pair	Hypotheses				
	F	G	H	I	J
(exp1, exp3.1)	✘	✘	✘	-	-
(exp1, exp3.2)	✘	✘	✘	-	-
(exp2, exp3.1)	-	-	-	✘	✘
(exp2, exp3.2)	-	-	-	✘	✘

Table 7: Hypotheses result FOREART

Pair	Hypotheses				
	K	L	M	N	O
(exp4, exp6.1)	✘	✘	✔	-	-
(exp4, exp6.2)	✘	✘	✘	-	-
(exp5, exp6.1)	-	-	-	✘	✘
(exp5, exp6.2)	-	-	-	✘	✘

Analysing the Tables 6 and 7, we can infer that in most of the cases the hypotheses did not reject $H0$ (indicated by ✘). This leads us to the conclusion that reputation model heterogeneity does not have any effect on the accuracy of the *Dishonest* agent reputation evaluation.

4 Conclusions

In this paper we presented some experiments using the SOARI architecture integrated into the FOREART testbed. Those experiments were performed to answer two questions: (1) is there any improvement in the reputation evaluation accuracy when enabling a more expressive communication? and (2) how does the heterogeneity influence the evaluation accuracy of the dishonest agents' reputation? The results obtained do not allow us to conclude that a more expressive communication about reputation or reputation model heterogeneity provides an accurate reputation evaluation of other agents. However, the results have shown the Repage reputation model benefits from the symbolic communication, which leads us to think that there are some intrinsic or implementation model's characteristics that provided it.
Since those were some preliminary experiments, the results obtained may be related to the fact that the experiments may not be the ideal ones to assess the effects of communication expressiviness and reputation model heterogeneity. Therefore, as a future work, we intend to design better experiments using and not using the ART and FOREART testbeds.
Moreover, we intend to perform experiments using more and different reputation models, thus expanding the analysis related to the effects of heterogeneity on the accuracy of reputation evaluation. Based on those results, we expect to have enough information to perform a detailed analysis to identify the relationship between the reputation models characteristics and the benefits of using the SOARI architecture.

Acknowledgements

This project is partially supported by FAPESP/Brazil. Guillaume Muller is supported by CNPq/Brazil under grant 382737/2008-3. Anarosa A. F. Brandão is supported by CNPq/Brazil under grant 310087/2006-6. Laurent Vercouter was partially supported by FAPESP/Brazil under grant 2005/02902-5. Jaime S. Sichman is partially supported by CNPq/Brazil.

References

1. S. Boslaugh and P. A. Watters. *Statistics in a Nutshell*. O'Reilly Media, Inc., 2008.
2. A. A. F. Brandão, L. Vercouter, S. J. Casare, and J. Sichman. Exchanging reputation values among heterogeneous agent reputation models: an experience on ART testbed. In *AAMAS '07: Proceedings of the* 6^{th} *International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 1–3, New York, NY, USA, 2007. ACM Press.

3. A. A. F. Brandão, L. Vercouter, S. J. Casare, and J. S. Sichman. Extending the art testbed to deal with heterogeneous agent reputation models. In C. Castelfranchi, S. Barber, J. Sabater, and M. Singh, editors, *Proceedings of the* 10^{th} *Workshop on Trust in Agent Societies*, Honolulu, Hawaii, 2007.
4. S. Casare and J. S. Sichman. Using a functional ontology of reputation to interoperate different agent reputation models. *Journal of the Brazilian Computer Society*, 11(2):79–94, 2005.
5. C. Castelfranchi and R. Falcone. Principles of trust in mas: Cognitive anatomy, social importance, and quantification. In *ICMAS '98: Proceedings of International Conference on Multi-Agent Systems*, pages 72–79, Washington, USA, 1998. IEEE Computer Society.
6. R. Conte and M. Paolucci. *Reputation in Artificial Societies. Social Beliefs for Social Order*. Kluwer, Boston, 2002.
7. K. K. Fullam, T. B. Klos, G. Muller, J. Sabater-Mir, A. Schlosser, Z. Topol, K. S. Barber, J. Rosenschein, L. Vercouter, and M. Voss. A specification of the Agent Reputation and Trust (ART) testbed: Experimentation and competition for trust in agent societies. In F. Dignum, V. Dignum, S. Koenig, S. Kraus, M. P. Singh, and M. Wooldridge, editors, *AAMAS '05: Proceedings* 4^{th} *International Conference on Autonomous Agents and Multi-Agent Systems*, pages 512–518. ACM Press, Utrecht, The Netherlands, July 2005.
8. K. K. Fullam, T. B. Klos, G. Muller, J. Sabater-Mir, Z. Topol, K. S. Barber, J. S. Rosenschein, and L. Vercouter. The Agent Reputation and Trust (ART) testbed architecture. In C. Castelfranchi, K. S. Barber, J. Sabater-Mir, and M. P. Singh, editors, *Proceedings* 8^{th} *Workshop on Trust in Agent Societies*, pages 50–62, Utrecht, The Netherlands, July 2005.
9. V. Haarslev, R. Möller, and M. Wessel. Querying the semantic web with racer + nrql. In *ADL '04: Proceedings of the KI-2004 International Workshop on Applications of Description Logics*, 2004.
10. G. Muller and L. Vercouter. L.i.a.r. achieving social control in open and decentralised multi-agent systems. Technical Report 2008-700-001, École Nationale Supérieure des Mines de Saint-Étienne, Saint-Étienne, France, 2008.
11. L. G. Nardin, A. A. F. Brandão, J. S. Sichman, and L. Vercouter. *SOARI: A Service Oriented Architecture to Support Agent Reputation Models Interoperability*, volume 5396 of *Lecture Notes in Computer Science*. Springer, Heidelberg, Germany, 2008.
12. Racer Systems GmbH and Co. KG, Hamburg, Germany. *RACERPro User's Guide Version 1.9.2*, 2007.
13. J. Sabater-Mir, M. Paolucci, and R. Conte. Repage: Reputation and image among limited autonomous partners. *Journal of Artificial Societies and Social Simulation*, 9(2), 2006.
14. J. Sabater-Mir and C. Sierra. Social regret, a reputation model based on social relations. *SIGecom Exch.*, 3(1):44–56, 2002.
15. L. Vercouter, S. J. Casare, J. S. Sichman, and A. A. F. B. ao. An experience of reputation models interoperability using interactions based on a functional ontology. In *Iberagents 2006*, Ribeirão Preto, SP, Brazil, 2006.
16. L. Vercouter, S. J. Casare, J. S. Sichman, and A. Brandão. An experience on reputation models interoperability based on a functional ontology. In *IJCAI '07: Proceedings of the* 20^{th} *International Joint Conference on Artificial Intelligence*, pages 617–622, Hyderabad, India, 2007.
17. G. Zacharia and P. Maes. Trust management through reputation mechanisms. *Journal of Applied Artificial Intelligence*, 14(9):881–907, 2000.