

Normative Monitoring of Agents to Build Trust in an Environment for B2B

Eugénio Oliveira, Henrique Lopes Cardoso, Maria Joana Urbano,
and Ana Paula Rocha

LIACC, DEI / Faculdade de Engenharia, Universidade do Porto
R. Dr. Roberto Frias, 4200-465 Porto, Portugal
eco@fe.up.pt

Abstract. Agents intending to be involved in joint operations need to rely on trust measures pointing to possible future solid partnerships. Using Multi-Agent Systems (MAS) as a paradigm for an electronic institution framework enables both to simulate and facilitate the process of autonomous agents, as either enterprises or individual representatives, reaching joint agreements through automatic negotiation. In the heart of the MAS-based electronic institution framework, a Normative Environment provides monitoring capabilities and enforcement mechanisms influencing agents' behavior during joint activity. Moreover, it makes available relevant data that can be important for building up contextual-dependent agent's trust models which, consequently, also influence future possible negotiations leading to new and safer agreements. To support agents data generation and monitoring, we here present ANTE platform, a software MAS integrating Trust models with negotiation facilities and Normative environments, for the creation and monitoring of agent-based networks.

1 Introduction

In Multi-Agent Systems (MAS), the role of an Environment has been defined by Odel and Weynes et al [1] as a support for agents to exist, and the medium for agents to communicate. However, other than being a kind of a passive component providing services to the agents and facilitating, upon request, the Environment can also be seen as an active component, gathering and exploiting all kinds of information produced during MAS activity. This is mostly the case when the MAS is seen as a support for electronic institutions, a basis for automatic B2B operations [2][5]. In this paper we look at the MAS environment as a component that actively monitors all the relevant information produced by different agents engaging in activities related with both creating joint agreements and coordinating actions that enable their execution. Information, produced along these stages becomes useful for inferring relevant trustworthy capabilities of the agents in the network. In this paper, we highlight the approach taken in ANTE (Agreement Negotiation in Normative and Trust-enabled Environments) platform [4] aiming at creating an active Environment that gathers information related with agents'

activity for the sake of decision-making [6]. ANTE encompasses a normative framework guiding agents interaction through norms that apply either sanctions or incentives. Useful information regarding agents' performance will influence future interactions and possibly agents joint work [3].

The structure of the paper is as follows: After this introduction, we first describe how to use MAS environment, based on the power of a normative environment, to produce and gather relevant information about agents activity. Then, in section 3 we introduce a computational Trust Model that benefits from such information in order to assign trust measures to agents, preventing an agent from future involvement with untrustworthy partners. Section 4 briefly presents the ANTE platform that provides all the needed services plus the environment for agents to interact in a trustworthy way. We close the paper with a brief conclusion.

2 The Relevance of the Environment

Many sophisticated problems we usually address can be classified as belonging to the Triple-D class of problems [4]. They reflect a reality that simultaneously is of a Distributed, Decentralized as well as Dynamic nature. This means that, besides input data and output actions being disperse (Distributed) at different nodes of a network, also, and most important, decision-making can be, at least partially, taken at different nodes of the (Decentralized) system. Moreover, the system trying to solve the overall problem at stake, has to deal with a changing, evolving reality (Dynamic). To work jointly in solving this kind of problems, agents need to select, in run-time, their best trustworthy partners for executing the tasks through mutual agreements to coordinate joint work. In the core of software agent's definition we find the "situatedness" property and, thus, the Environment becomes an important component to be specified and designed for multi-agent systems.

As a relevant component of MAS, although not always considered per se, Environment can be seen as an active medium to facilitate the way how to reach certain consensus, agreements or mutual contracts. Instead of looking at the Environment as a regulator through the internal mechanisms of a Normative environment [2][5][3], we here see the Environment as a medium in which notifications about the outcomes of actions agreed upon, after the establishment of joint contracts, appear together with the output of the Normative Environment that assesses them and, then, combines all those pieces of information in meaningful knowledge for future use. More precisely, through the environment embodied in ANTE, the software platform we have developed, relevant data is made available for building up contextual-dependent trust models which, as a consequence, will influence future possible agents negotiation leading to new, and better, agreements.

2.1 Environment as a Facilitator

It is out of the scope of this paper to describe other environment facilities as it is the case of the negotiation protocols and mechanisms we advocate [6] in the ANTE platform, for a dynamic and adaptive convergence towards final agreements then materialized into contracts. We, along this paper, want to mainly highlight the importance of the Environment in recording the relevant events produced by the agents and the normative environment outputs as a result of agents' activities during contract enactment. Appropriately combining the trace of such events, leads to a tightly coupled connection between electronic contract monitoring and a computational trust model for estimating agents' trustworthiness. The Normative dimension of the Environment handles a normative state NS (including all elements that characterize the current state of affairs for each established agreement), a set IR of institutional rules that manipulates NS and a set N of norms, which can be seen as a special kind of rules. NS includes institutional facts that represent institutional recognition of contract related real-world events. The Normative Environment also notifies the agents about relevant contract-related events. The same mechanism conveys information to feed the computation Trust model.

A contract $C = \langle T; CA; CI \rangle$ is monitored by the normative environment and is defined as follows: A contract C is a relation of type T within which a group of agents in set CA commits to a joint activity, under the normative context using information CI, a set of definitions regarding the role of the participants, the values to be exchanged (e.g. products and money) or any parameters defining their provision. An electronic representation of a contract includes a set of norms that specify how a specific business is to be enacted. A norm is a rule whose conditions analyze the current NS and whose conclusion prescribes obligations agents ought to fulfill. As a consequence, sanctions may be imposed by prescribing obligations upon violation events, thus producing more relevant information in the environment. In the context of a MAS in which agents represent entities looking for partners to establish agreements and commit themselves to joint actions, the role of a normative environment [3] is twofold. Given an agreement, it is necessary to check if the partial contributions of individual agents make their way in enabling a successful execution of tasks leading to the agreed overall goal. In many cases, the execution of the needed tasks is itself distributed, which requires agents to enact by themselves their part of the agreement. Monitoring this phase is therefore an important process and is possible through the analysis of the interactions and facts produced, and made visible, in the environment. Furthermore, it is possible that after successfully negotiating and reaching agreements, self-interested agents are no longer willing to fulfill their commitments, which becomes also visible to the environment, since normative rules generate notifications (prescribing sanctions) in reaction to such events.

As an example, the normative environment may apply norm n enforcing a sanction (e.g. a payment) when a violation regarding delivery has been detected. This puts in evidence the second role of a normative environment, that of enforcing norms by coercing agents to stand for their commitments. At the same

time, the whole idea of an active environment is to use the relevant information generated through the normative environment and make it useful for further agents' reasoning strategy.

2.2 Information in the Environment

The way agents enact their contracts provides relevant information for trust building and, as a consequence, may decisively influence partners selection for future electronic contracts. In ANTE, an image of all the relevant real-world transactions between different agents become recognized as "institutional reality" [8]. As stated above, a normative state NS records every element that is relevant in contract enactment and is composed of institutional reality elements. Agents active in the Environment subscribe to the normative environment to be notified about eminent contractual obligations and commitments to be fulfilled. Consider the following representation of a directed obligation with time constraints: $Obl_b, c(l < f < d)$. This is an obligation of agent b (the bearer) towards agent c (the counterparty) to bring about fact f between liveline l and deadline d. An obligation such as this can give rise to a rich set of events, such as temporal violations, due fulfilment or violation. As agents go on interacting according (or not) with pre-established contracts, the normative environment reacts appropriately, and produces more (normative) information about the on-going processes. The challenge is to be able to gather and manage all that produced data, to infer meaningful information about agents' future behavior.

Moreover, other than information generated during two agents' mutual interaction, there is other available information, coming from third parties, that may be relevant to characterize a specific agent. These indirect evidences about an agent's behavior (an "image" the agent builds up in the society of all currently active agents, due to previous interactions), is called reputation. Reputation is seen as the social-based process of transmitting beliefs about a specific agent (the trustee), as a consequence of social evaluation circulating and, according to [9], are represented as reported evaluations. There are computational trust models that use reputation as another piece of evidence to be taken into account about any particular agent. For that purpose, they need to estimate the credibility of both the transmitted information and of those agents reporting that information. Providing that third party agents, not directly involved in a particular negotiation are willing to provide their own perspective on a particular agent, this new piece of evidence will also be present in the environment and can also be identified as relevant to be merged with direct evidences [10]. The perspective an agent is now able to infer, through the combination of the available information, about another agent becomes wider and more independent of local knowledge. However we still did not address the problem that could arise if reputation information and other available trust-based measures happen to be of different nature and formats. How to merge them, in a coherent way, through the computational Trust and Reputation Model is a challenge for our near future research.

3 Trust

We follow a basic and established definition for Trust (although not always consensual). Closely, but not completely following some of the ideas expressed in [7], we propose the following formal definition for trust: $\text{Trust}(i; j; \varphi)$ meaning that the Truster(i) trusts Trustee(j) to do Action(α) leading to the achievement of Goal(φ) if:

$(GOAL_i\varphi) (BEL_iPOWER_j\varphi) (BEL_i\Box(\alpha \models \varphi)) (BEL_iINTEND_j\alpha)$, where \Box is the usual ‘necessity’ modal operator, $POWER_j\varphi$ states that agent j is capable of achieving φ and $INTEND_j\alpha$ is agent j ’s intention to execute action α . Intention means choice with commitment according to Cohen and Levesque’s well accepted theory.

Agents should thus rely on Trust measures to select their best potential partners, improving their chances to achieve their own future goals. It has been established that, whenever decisions have to be made for selecting potential partners in a future activity, it becomes mandatory to rely on some kind of past evidence pointing to an estimate of the other agents’ trustworthiness. We need to remember here that Trustworthiness is an intrinsic characteristic of an entity (here an agent) while Trust measures reflect other ones’ (potentially different) assessment of that particular agent under evaluation. Exhaustive discussions are taking place about all the relevant factors that can influence a trust measure characterizing the perspective each agent builds up on other agent internal trustworthiness [7]. We may identify two different categories of such factors: subjective (propensity and disposition) and, most valuable, objective (past experience regarding previous direct interactions, current social image of the agent under scrutiny, the reputation and other indirect evidences). Another perspective over the computational trust model is if, indeed, a single measure is enough to measure an agent’s trustworthiness. In fact, although we often look for competence, it sometimes is not the only dimension that matters for successful future joint activities. Other more subtle features like how benevolent and integer an agent can be, may be seen as determinant in deciding on how much shall an agent trust in another one. Another important concern is the contextual nature of trust: Situation, task and time dependence are different factors to take into account when measuring and ranking other agents’ trust.

3.1 Computational Trust Models

There has been a number of proposals for computational trust models that rely in one or several of the factors and dimensions referred above. Direct past mutual interaction experience is the most common factor these models take into consideration as it is the case of [11]. A few computational models already include some kind of agent disposition to trust and emphasize the importance of context [12] [13]. Moreover, only recently we have called the attention to the importance of the dynamics of the relationships taking place, along the time, between the different agents, leading to a relative change of the power relationship and the evolution of mutual goodwill [13].

Proposed Trust Model. Our computational Trust model, included in the ANTE software platform, has different components responsible for managing most of the factors recognized as influencing trust-based decision-making. As a model, it encompasses both the different dimensions of trustworthiness (integrity, ability, benevolence) and the truster's disposition to trust. As a software program, implementing most of the referred important aspects, although not yet all of them, it already includes an aggregator of past direct experience, Sinalpha, a Contextual Fitness component, a Social and an Integrity Tuner. Through such model we gain the ability to compute adequate Agent i estimations of Agent j trustworthiness to achieve a specific goal φ , in different environments, including those of high dynamics, where evidence on the agent under evaluation is scarce.

The two main components first aggregate past contractual behavior evidences (Sinalpha) and then consider how agents fit into the context under consideration (Contextual Fitness). The aggregator gets information from the environment about agents' performance regarding their past obligations (either fulfillment, partial fulfillment or violation). It uses a sinus-based curve that is reshaped at both top and bottom extremities [6]. It uses a function of α that presents a sinusoidal shape (see Equation 1). By setting $\delta = +0.5$, the trustworthiness value is restricted to the range $[0;1]$. $trustworthiness = \delta * (\sin\alpha + 1)$ (1) The trustworthiness score of the agent is minimum when $\alpha = 3\pi/2$ and maximum at $\alpha = 5\pi/2$. This score is updated using $\alpha(i+1) = \alpha(i) + \lambda * \omega$, where λ reflects the outcome (either positive or negative) associated with the piece of evidence being aggregated and where parameter ω defines the size of the ascending/descending step in the trustworthiness path. For example $\omega = 6$ means that any trustee that is a newcomer, could be considered by the truster as fully trustworthy after presenting six outcomes of type F (fulfilled) in a row. Asymmetry about gaining and losing Trust, maturity after reaching a certain status and distinguishing different patterns of past behaviour are properties that are featured by the model. It has been proved that the way the aggregator merges past evidences, gives a realistic idea of, although just for simple situations, agents' expected behaviours in future similar situations.

The second component of the model, responsible for social awareness, counts for the specific appropriateness of the target agent (Trustee) regarding that particular situation in which the other agent (Truster) is willing to get help, to work jointly or to engage in a relationship with [13]. It is based on the concept of a Context that captures and represents the relevant information about the current situation, that Contextual Fitness appreciates how an agent can be, at that specific time, trustworthy or not. A Context is here defined as an N-tuple including agent identity, current time, location and a set of task-related attributes: type, complexity, deadline and expected outcome.

Contextual Fitness component, a situation-aware tuner may downgrade the trustworthiness scores computed by the aggregator (Sinalpha) whenever the agent under assessment has proved to behave poorly in that specific situation. We are using the information gain metric on the previous evidences on the agent different tendencies of failure. This approach differs from other situation-aware

computational trust approaches by its flexibility and ability to reason in terms of context even when the evidence on the agent in evaluation is scarce. Briefly, we endow the environment with the capabilities of clustering different potential partners waiting for being selected and extract stereotypes characterizing in what conditions they have either fulfilled or violated their obligations. Therefore, it may be the case that a particular agent has a good trust measure, however it has an handicap regarding a specific feature that is of most relevance for the concrete partnership another agent is interested in (for example, a faster delivery time of a product). The most important to stress here is that the environment is empowered with the means to strongly influence the decision-making process of agents, by providing indicators that, once combined, become relevant for trust evaluation.

4 ANTE Platform

Grounded on the recognition of the Environment as a relevant component to give structure and support situated MAS, we have developed ANTE, a software framework in which agents can both establish agreements (and contracts) through negotiation and, simultaneously, benefit from monitoring capabilities pointing to more transparency of joint activities. Through the development of ANTE we intended to illustrate how an active environment may help in reliable and informed agents decision-making in what joint agreements is concerned [6]. ANTE supports multi-agent collective work leading to agents' intention satisfaction, by providing negotiation as a mechanism for finding mutually acceptable agreements with trusted partners and the enactment of such agreements (contracts). Gathering appropriate data and evaluating the Normative State, enables feeding the Computational Trust model which, in turn, provides information for better selecting partners for future agreements.

In ANTE, the environment actively mediates interaction among agents in several different ways: i) The Environment is, primarily, a Facilitator by providing negotiation protocols, Ontology translation services and Contract building tools. In this perspective we can see it as an Interaction-Mediation level [1] regulating access to resources and mediating interaction between agents leading to possible agreements. ii) The Environment also provides mechanisms for the normative enactment of those agreements (established contracts among agent partners), generating Normative States (NS) and becoming responsible for monitoring those NS in an active way. ANTE includes a normative framework of constitutive and institutional rules, according to the institutional reality model, operating in a rule-based process. iii) Through evaluating the contract enactment phase, the Environment improves the chances for better future negotiations by progressively updating the available Trust measures to help in the partners' selection process. ANTE includes, at present, a computational Trust model made up of two main modules: SINALPHA aggregator and contextual fitness. iv) In the ANTE specification model, another component acts as a "social Environment" and is responsible for adaptive deterrence mechanisms (through dynamically

calculated fines imposition) reacting to different agents population behaviors. In Figure 1 we display how the active environment is subject to the simultaneous intervention of the agents activity (negotiation and contract execution) which leads to “institutional facts” monitored in the active environment. Those “institutional facts”, in turn, feed and update the computational Trust Model.

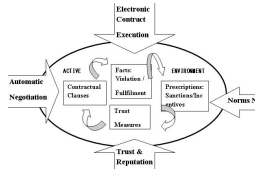


Fig. 1. The Active environment. Facts (clauses violation/fulfillment) activate norms leading to possible sanctions, changing agents trustworthiness and, thus, influencing the way agents negotiate future contracts.

Monitoring rules are responsible for gathering all the needed information that becomes relevant to feed the computational Trust model which dynamically combines that information to infer agents’ trust measures. This last component is, however, meaningless in the context of what is the aim of this paper: information merging to dynamically feed an agents’ trust model. ANTE is a modular and extensible JADE-based architecture implementation, integrated with JESS rules engine, accessible through several different GUIs for agents to announce their needs as well as for inspecting the whole negotiation process, normative states and contract results. Although ANTE has been targeting electronic contracting in B2B domains, and because it encompasses several agreement technologies like negotiation, normative environments and computational trust models, it was conceived as a more general framework having in mind a wider range of application domains including social networks monitoring.

Experiments Assessment. We have used ANTE platform both as an environment as well as a set of services suitable for achieving agents agreement in a B2B scenario. As an environment it became a useful interface between negotiation services, normative environment services and computational trust services. Many different experiments have been performed, mainly involving agents as enterprise representatives negotiating several different attributes of particular goods (fabric type, quality and time of delivery). Since the description of a full range of different partial and integrated experiments we have made, can be found in [14], we here just stress out a qualitative assessment about the summarized results coming out of the use of Trust during the partners selection negotiation phase together with normative sanctions application. Different experiments enable the agents to use Trust either to pre-select the most promising potential partners for the future joint activity (to supply a certain quantity of a certain

type of fabric, until a specific deadline) or to use trust only during the negotiation phase, to better evaluate proposals under scrutiny. Other experiments combine these two possibilities. Also, several different metrics have been used. However, we here only report on the metric reflecting the number of well succeeded contracts that have been established following the use of these referred techniques. After a number of experiments we concluded that, by using the Computational Trust Model, the outcomes were strongly influenced, besides the agents behavior type, by other factors like, for example, the degree of similarity or dissimilarity of the agent proposals under comparison. Moreover, one of the most important conclusions was that using trust for pre-selecting the agents for future negotiation together with a moderate use of Trust measures during the negotiation process itself, was the best policy for an agent to follow. Using Trust for selecting partners increased the number of succeeded contracts in this B2B scenario.

5 Conclusions

Distributed, decentralized systems for dynamic situations, appropriate for Triple D kind of problems, are well represented by means of Multi-Agent System architectures. Distributed Agents, representing different entities (like enterprises), take their own decisions (selecting partners) in a decentralized way. Moreover, they react to dynamic situations (like other agent trustworthiness change) by appropriately using trust measures to help on their own decision-making (which partner to select). Following the same line of those authors who claim that the Environment should be considered as an important, separate component in the specification and design of Multi-agent Systems, we here advocate a set of capabilities that should be made available for more secure, comprehensive and trusted agents mutual activities possibly leading to more fruitful joint agreements, as it is the case of B2B e-contracts. Besides facilitating functionalities like negotiation protocols, ontology services, normative framework application and monitoring, we here, in this paper, strongly emphasize the importance of an Environment active role also as a support for information fusion. Gathering and evaluating information related with normative states and agents behavior regarding norms abidance and contracts fulfillment is of paramount importance for feeding computational Trust models. Combining agents activity-driven information, leading to updated and contextualized agents' Trust measures become crucial to guide future agents' activity and mutual agreements. A "social-based" Environment can also have the privilege of, by reasoning about all agents activity, to detect, prevent (or incentivize), patterns in the society regarding maleficent (or beneficial) agents' behavior. Moreover, and having in mind its active role to ease and promote possible agent agreements, "social Environment" has been empowered with the capability of adaptively impose fines for the sake of better regulate the agent population activities and promoting confidence in the possibility of future better agreements. This aspect of the environment lacks of further investigation. However, in this paper, we mostly emphasize how, through a normative framework responsible for monitoring agents contractual behavior and computational

Trust models, the environment provides relevant evidences that, in the end, enables an agent to better characterize other potential partner agents and better decide about either to engage or not in a future relationship and better agreements, as it this the case for B2B contracts negotiation scenario. ANTE software platform has been developed taking all these features into consideration.

References

1. Weyns, D., Omicini, A., Odell, J.: Environment as a first class abstraction in multi-agent systems. *J. Autonomous Agents Multi-Agent Systems* 14, 5–30 (2007)
2. Dignum, F.P.M., Broersen, J., Dignum, V., Meyer, J.-J.: Meeting the deadline: Why, when and how. In: Hinchey, M.G., Rash, J.L., Truszkowski, W.F., Rouff, C.A. (eds.) FAABS 2004. LNCS (LNAI), vol. 3228, pp. 30–40. Springer, Heidelberg (2004)
3. Cardoso, H., Oliveira, E.: Social Control in a Normative Framework: An Adaptive Deterrence Approach. *Web Intelligence and Agent Systems* 9(4), 363–375 (2011)
4. Oliveira, E.: Software Agents: Can we Trust Them? In: Proceedings of IEEE 16th Int. Conf. on Intelligent Engineering Systems, INES (2012)
5. Cardoso, H., Oliveira, E.: Electronic Institutions for B2B Dynamic Normative Environments. *Artificial Intelligence & Law (Special Issue on Agents, Institutions and Legal Theory)* 16(1), 107–128 (2008)
6. Cardoso, H.L., Urbano, J., Brandão, P., Rocha, A.P., Oliveira, E.: ANTE: Agreement Negotiation in Normative and Trust-enabled Environments. In: Demazeau, Y., Müller, J.P., Rodríguez, J.M.C., Pérez, J.B. (eds.) Advances on PAAMS. AISC, vol. 155, pp. 261–264. Springer, Heidelberg (2012)
7. Falcone, R., Castelfranchi, C.: Trust and deception in virtual societies. In: Castelfranchi, C., Tan, Y.-H. (eds.), pp. 55–90. Kluwer A. P., Norwell (2001)
8. Searle, J.R.: *The Construction of Social Reality*. Free Press, New York (1995)
9. Sabater-Mir, J., Paolucci, M., Conte, R.: Repage:reputation and image among limited autonomous partners. *J. of Art. Societies and Social Sim.* 9(2), 3 (2006)
10. Urbano, J., Rocha, A.P., Oliveira, E.: A Socio-Cognitive perspective of Trust. In: Ossowski, S. (ed.) *Agreement Technologies*, ch. 23. Law, Governance and Technology Series, vol. 8, pp. 419–431. Springer (2013)
11. Sabater, J., Sierra, C.: REGRET: Reputation in gregarious societies. In: Proc. of the 5th Int. Conf. on Autonomous Agents, AGENTS 2001, pp. 194–195. ACM (2001)
12. Rehak, M., Gregor, M., Pechoucek, M.: Multidimensional Context Representations for Situational Trust. In: Proc. of the IEEE DIS 2006: Collective Intelligence and Its Applications, pp. 315–320. IEEE Computer Society (2006)
13. Urbano, J., Rocha, A.P., Oliveira, E.: A dynamic agents behavior model for computational trust. In: Antunes, L., Pinto, H.S. (eds.) EPIA 2011. LNCS, vol. 7026, pp. 536–550. Springer, Heidelberg (2011)
14. Urbano, J.: *A Situation-aware and Social Computational Trust Model*. PhD Thesis Report, Faculty of Engineering, University of Porto (2013)