Objective Trust-based Agents: Trust and Trustworthiness in a Multi-agent Trading Society

Mark Witkowski and Jeremy Pitt
{m.witkowski, j.pitt}@ic.ac.uk
Intelligent and Interactive Systems Group
Department of Electrical and Electronic Engineering
Imperial College of Science, Technology and Medicine
Exhibition Road, London, SW7 2BT, United Kingdom

Abstract

In this paper we develop a notion of "objective trust" for Software Agents, that is trust of, or between, Agents based on actual experiences between those Agents. Experiential objective trust allows Agents to make decisions about how to select other Agents when a choice has to be made. We introduce a mechanism for such an "objective Trust-Based Agent" (oTB-Agent). The trust one Agent places in another is dynamic, updated on the basis of each experience. We summarise experimental results obtained from a simulated trading environment.

1. Introduction

The trust relationship that might exist between two parties (human or artificial) has, in its broadest sense, proved difficult to define ([2], [3], [4]). We synthesise the following as a working definition, suited to the purposes of this paper. "*Trust is the assessment by which one individual, A, expects that another individual, B, will perform (or not perform) a given action on which its (A's) welfare depends, but over which it has limited control*". Trust therefore implies a degree of dependency of A on B, this dependency may be reciprocal. Because of its many implications for automated electronic trading and commerce, trust between autonomous software Agents is the subject of on-going research. For instance, Castelfranchi and Falcone [1] argue that trust is subjective and different beliefs can be combined to create a *Degree of Trust* measure. Jonker and Treur [3] present a formalised framework for the description of trust based on sequences of experiences between Agents. We take a practical and experimental approach to trust through experience.

2. The trading scenario

Figure one illustrates the experimental trading scenario we have used to test and evaluate oTB-Agent based trading. Its purpose is to allow us to investigate the practical consequences of trust relationships in autonomous trading under controlled conditions, and to make observations about the role of trust and trustworthiness in e-commerce.

Trading is divided into *trading cycles*, each involves three transaction steps. First, the *bid step*, in which customer Agents receive demand for commodity units from "end-users" and issue bids to supplier Agents to meet that demand. Second, the *offer step*, in which supplier Agents select customer agents to serve and make *offers* of units in response to the bids they have received. Third, the *utilisation step*, in which customers distribute the units they have been offered to their end-users, and notify the supplier that offered them units if they utilised all the allocation they were offered. Any units not taken up by customer Agents are lost, to the immediate detriment of the supplier. Customer Agents can *overbid* for more units than they need in an attempt to safeguard their supply in times of shortage, at the risk of appearing untrustworthy.

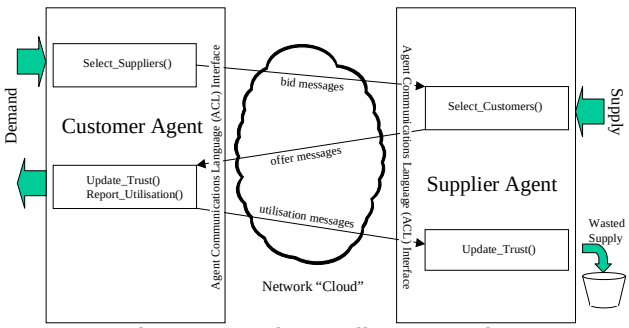

Figure one: The Trading Scenario

Every Agent maintains a *trust vector*, t, quantitatively recording the "opinion", based on direct experience, that Agent holds about the trustworthiness of the other Agents with which it can trade. The trust vector forms the primary source for selecting trading partners, and is itself updated after each transaction. In the experiments described here customer Agents preferentially select suppliers they trust to offer them the units they require. Similarly, supplier Agents attempt to distribute their supply of units preferentially to

customer Agents they trust to pass those units on to its end-users. While demand may vary between trading cycles, the total amount of units available (supply) is taken as fixed.

There is no overall control or centralised mediation in this model. Each Agent makes its current trading decisions based on its past experiences of trading with other Agents, selecting Agents it most trusts to trade with.

3. The trust functions

A customer Agent's trust vector is updated on the basis of the perceived reliability of a supplier, offers made in response to bids. A supplier Agent's trust vector is updated by comparing utilisation to offers made. The trust dependency relationship is therefore reciprocal between customer and supplier. The trust function takes two parameters, α ($0 \leq \alpha \leq 1$), the degree to which a positive experience enhances trust, and β ($0 \leq \beta \leq 1$), the degree to which a negative experience damages the relationship. Each element (${}_m t_n$) in Agent m's trust vector, t, is updated after a transaction with a trading partner Agent (n) thus:

${}_m t_n \leftarrow {}_m t_n - (\beta * {}_m t_n)$, if a request was made, but no corresponding offer received, or

${}_m t_n \leftarrow {}_m t_n + (\alpha * (1 - {}_m t_n))$, if request satisfied, or

${}_m t_n \leftarrow {}_m t_n + ((\alpha * (\text{request} - \text{offer})) * (1 - {}_m t_n))$, if offer < request, or

${}_m t_n$ is left unchanged otherwise.

These formulations are normalised such that a sequence of positive experiences asymptotically moves trust ${}_m t_n$ towards 1.0 (full trust), but a sequence of negative experiences moves it towards 0.0 (complete distrust).

The function matches our intuition that trust is most enhanced by getting exactly what we expected, partially enhanced by getting some of our request and damaged by being excluded. The formulation also conforms to our expectation that recent experiences should be given greater weight that earlier ones, a mechanism for the forgiveness of past transgressions. Similar formulations find application in areas such as reinforcement learning.

5. Experimental findings (summary)

We have performed an extensive series of experiments with this mechanism, and have observed the following points, which appear to have some parallels with human commerce and important implications for e-commerce:

- Agents tend to form tight trading partnerships, and in static supply:demand conditions trade preferentially within those cliques ("trust-begets-trust")
- When demand exceeds supply, some customer Agents fail to develop relationships with a sufficient number of suppliers to meets their needs (and vice-versa when supply exceeds demand) – "trust begets viability"
- When demand increases to exceed supply, supplier Agents discard less trusted customers first – "trust begets loyalty"
- Greedy Agents, those who wastefully overbid, can be successful by nurturing preferential trust relationships with specific suppliers, at the expense of their overall reputation within the community – "the selfish can appear trustworthy"

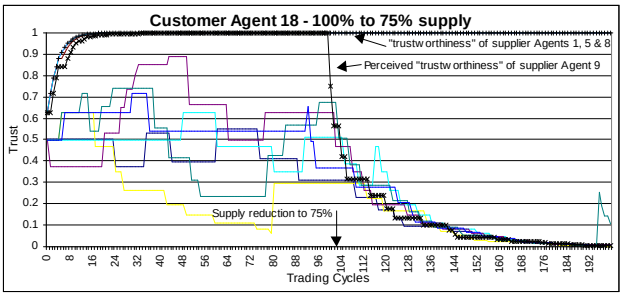

Figure two: Effect of changing demand on trust

Figure two illustrates the effect on perceived trust as demand changes relative to supply. It shows the trust vector values held by one customer Agent (number 18 of 20) for each of 10 possible supplier Agents. When supply equals demand (cycles 0-99) Agent 18 develops (and trades with) four suppliers regularly (suppliers 1, 5, 8 and 9). When supply is reduced to 75% of demand at cycle 100, supplier 9 drops customer 18, its least favoured customer (and so now appears untrustworthy to Agent 18). Close inspection reveals that the relationship between customer 18 and supplier 9 was the last to develop and was the weakest of the four. Note also the general tightening of relationships as supply is reduced.

Further discussion on the nature of trust, details of the algorithm and experimental results may be found in [4].

Acknowledgements: This work was partially sponsored by EC-funded projects ACTS-AC333 MARINER and ALFEBIITE (IST-1999-10298)

References

[1] Castelfranchi, C. and Falcone, R. (1998) "Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification", ICMAS-98, pp. 72-79.

[2] Gambetta, D. (ed.) (1988) "Trust: Making and Breaking Cooperative Relations", Basil Blackwell, Oxford.

[3] Jonker, C.M. and Treur, J. (1999) "Formal Analysis of Models for the Dynamics of Trust Based on Experiences", Proc. MAAMAW '99, Springer Verlag, Berlin, pp. 221-232.

[4] Witkowski, M., Artikis, A. and Pitt, J. (2000) "Trust and Cooperation in a Trading Society of Objective-trust Based Agents", Autonomous Agents 2000, Proc. Workshop on Trust, Fraud and Deception.