40 SOFTWARE AGENTS IN VIRTUAL ORGANIZATIONS: GOOD FAITH AND TRUST

Francisco Andrade[1], Paulo Novais[2], José Machado[2] and José Neves[2]

[1]*Escola de Direito, Universidade do Minho, Braga, PORTUGAL*
[2]*DI-CCTC, Universidade do Minho, Braga, PORTUGAL*
fandrade@direito.uminho.pt, {pjon, jmac, jneves}@di.uminho.pt

Virtual organizations tend to play an ever more part in electronic commerce, as well as software agents, here understood as the building blocks of the methodology for problem solving that is being subscribed. Indeed, one of the issues that have to be addressed is the capability of such entities to rationally and autonomously "think" and decide. The behavior of these agents may go more and more unpredictable; they will choose their own strategies and define their own planning where are faced to a problem, being possible that they may act with good faith or with bad faith. This leads us to the absolute need of considering the major issue of trust in software agent's environments.

1. INTRODUCTION

It must be anticipated the possibility of software agents to play a determinant role in corporate bodies, in virtual enterprises ("temporary alliances of organizations that come together to share skills or core competencies and resources in order to better respond to business opportunities" (Camarinha-Matos and Afsarmanesh, 2004), in Dynamic Virtual Organisations ("...a VO that is established in a short time to respond to a competitive market opportunity, and has a short life cycle" (Camarinha-Matos and Afsarmanesh, 2004). This, intervention intends interactions based on contracts and relations of trust (Teubner, 2001). But agents operate without the direct intervention of human beings and "have some degree of control over their actions and inner states" (Weitzenboeck, 2002). Indeed, it can be can assumed that agents behave upon mental states, that is to say their behavior is a product of reasoning processes over incomplete or unknown information (Andrade et al, 2007). In this sense, agents do make options and their behavior can not be fully predicted.

Thus being, considering open distributed systems and autonomous agents that "act and interact in flexible ways in order to achieve their design objectives in uncertain and dynamic environments", is it possible to trust agents in electronic relations? Trust is mainly a belief in the honesty or reliability of someone ("a belief an agent has that the other party will do what it says it will…given an opportunity to defect to get higher payoffs"). It is clearly a requisite of the utmost importance to consider when deciding on "how, when and who to interact with" because it can not be assumed in advance whether or not agents will behave according to rules of

honesty and correctness. This issue leads us to consider the possibility of agents acting with good or bad faith. But also it forces us to view that ways of ensuring a high degree of reliability of electronic relations participants are required, namely the "need for protocols that ensure that the actors will find no better option than telling the truth and interacting honestly with each other". But trust can be perceived in different ways, from an individual perspective ("an agent has some beliefs about the honesty or reciprocities nature of its interaction partners"), or from a social or systemic perspective ("the actors in the system are forced to be trustworthy by the rules of encounter (i.e. protocols and mechanisms that regulate the system").

At the individual level, trust arises from learning (agents do learn from experience), from reputation (a view "derived from an aggregation of members of the community about one of them") or from socio-cognitive models (mainly the belief that someone is competent or willing to do something).

At the system level, trust can be ensured by constraints imposed by the system, either by using protocols that prevent agents from lying or colluding, or by making the system itself spreading agent's reputation as being truthful or liar, or even by using a system "proof" or "guarantee" of reliability "through the references of a trusted third party".

All these will be important elements to consider within open systems where agents with quite different characteristics may "enter the system and interact with one another", offering different services with different levels of efficiency. And this is of the utmost importance in considering the participation of agents in Virtual Organisations.

2. VIRTUAL ORGANISATIONS

The notion of "consortium" is quite well known in commercial law for a long time, but VE certainly enhance its use in the commercial arena. But the main characteristics of "consortium" remain in virtual enterprises. Here, the main characteristics of commercial societies simply do not exist: there is really not an entity different from the participating companies; there is not an autonomous patrimony; there are no common profits (as an own patrimony) to be distributed among partners (Abreu, 2004). But the "consortium" is valid for itself, as a way to enhance the possibilities of commerce development for its participants.

A virtual enterprise must be regarded as a legal unity based in an organization of informatics means, an autonomous instrument for the production of immaterial goods (or services) only exchangeable through Internet, in a market without any physical, local or time constraints (Abreu, 2003). But it may also be understood as an assemble of enterprises legally and economically autonomous, connected through telematic means temporarily cooperating in the fulfillment of a project or economic activity (Abreu, 2003). Authors tend to assume VE as a "temporary alliance between globally distributed independent companies working together to improve competitiveness by sharing resources, skills, risks and costs" (Camarinha-Matos et al, 2007, Crispim and Sousa, 2005), that is to say a "consortium" – two or more different entities (natural or corporate) "get obliged to undertake certain activities or assuring certain contributions in order to make it possible to achieve certain material or legal acts" (Abreu, 2004). Yet, it may as well be considered the possibility of

Virtual Organisations Breeding Environments, understood as "an association or pool of organizations and their related supporting institutions that have both the potential and the interest to cooperate with each other, through the establishment of a "base" long-term cooperation agreement" (Camarinha-Matos et al, 2005). A With respect to the computational paradigm it were considered extended logic programs with two kinds of negation, classical negation, ¬, and default negation, not. Intuitively, not p is true whenever there is no reason to believe p (close world assumption), whereas ¬p requires a proof of the negated literal. An extended logic program (program, for short) is a finite collection of rules and integrity constraints, standing for all their ground instances, and is given in the form:

p ← p1 ∧ … ∧ pn ∧ not q1 ∧ … ∧ not qm; and
   ? p1 ∧ … ∧ pn ∧ not q1 ∧ … ∧ not qm, (n,m ≥ 0)

where ? is a domain atom denoting falsity, the pi, qj, and p are classical ground literals, i.e. either positive atoms or atoms preceded by the classical negation sign ¬ (Binmore and Fun, 1992). Every program is associated with a set of abducibles. Abducibles can be seen as hypotheses that provide possible solutions or explanations of given queries, being given here in the form of exceptions to the extensions of the predicates that make the program.

In terms of a VO one of their building blocks or predicates may be given by predicate *vo*, that stands for a particular entity or organisation, which in abstract terms may be given in the form (Figure 1 and Figure 2).

vo(…).
¬vo(…) ← /The closed word assumption is being softened/
   not vo(…) ∧
   not exception$_{vo}$(…).
? (vo(…,Y) ∧ Y ≥ 0 ∧ Y ≤ 1). /This invariant states that *vo* takes accuracy values on the interval 0…1/
? ((exception$_{vo}$(…, X,Y) ∨ exception$_{vo}$(…,X,Z)) ∧ ¬(exception$_{vo}$(…,X,Y) ∧ exception$_{vo}$(X,Z))).

/This invariant states that the exceptions to the predicate *vo* follow an ***exclusive or***, and that the last two attributes of *vo* (i.e. (X,Y) and/or (X,Z)) state that there exists a functional dependency between (X,Y) and/or (X,Z) and the remaining attributes of *vo*/.

Figure 1 - The extension of predicate *vo* that stands for a particular company or organisation.

Anyway, "the concept of virtual organization (VO)", as a special kind of consortium for electronic commerce, appears as "particularly well-suited to cope with very dynamic and turbulent market conditions" and "this is largely due to the possibility of rapidly forming a consortium triggered by a business opportunity and specially tailored to the requirements of that opportunity. Implicit in this is a notion of agility, allowing rapid adaptation to a changing environment", with its content being object of a process of optimization, as it is described in (Neves et al, 2007) (Figure 3).
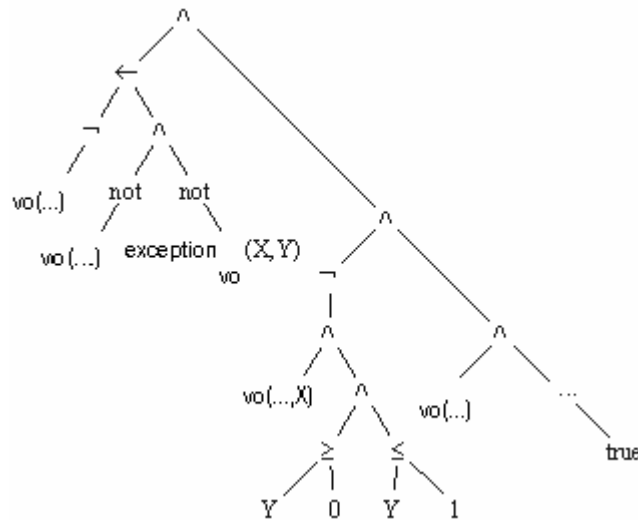
Figure 2 – The evolutionary logic program for predicate *vo.*

There might be legal issues on the use of virtual companies, since they imply cooperation agreements and might restrain concurrence between partners and or between these and third parties (prevented from accessing the agreement) which might have implications in the field of concurrence law (Abreu, 2003). Yet, these agreements may be totally legal, provided the non existence of restrictions to concurrence and the non elimination of concurrence in a substantial part of the market. These virtual enterprises may be quite interesting for the electronic commerce and software agents may certainly play an important role in it. Besides the question of the legal consideration of the electronic agents (are these mere tools used by the participants - natural or legal persons - in the "consortium", it is anyway unavoidable to consider the issue of the behaviour of agents by itself). And it is quite important that agents "know The Law" and social standards of behavior and abide to its rules. But is it possible to have agents abiding with legal and social norms? (Brazier et al, 2002).

3. SOFTWARE AGENTS AND GOOD FAITH

Software agents are computational entities with a rich knowledge component, having sophisticated properties such as planning ability, reactivity, learning capabilities, cooperation, communication and the possibility of argumentation. It is also possible to build logical and computational models having in consideration The Law norms (i.e., legislation, doctrine and jurisprudence). Agent societies may mirror a great variety of human societies, such as commercial societies with emphasis to behavioral patterns, or even more complex ones, with pre-defined roles of engagement, obligations, contractual and specific communication rules. An agent must be able to manage its knowledge, beliefs, desires, intentions, goals and values.
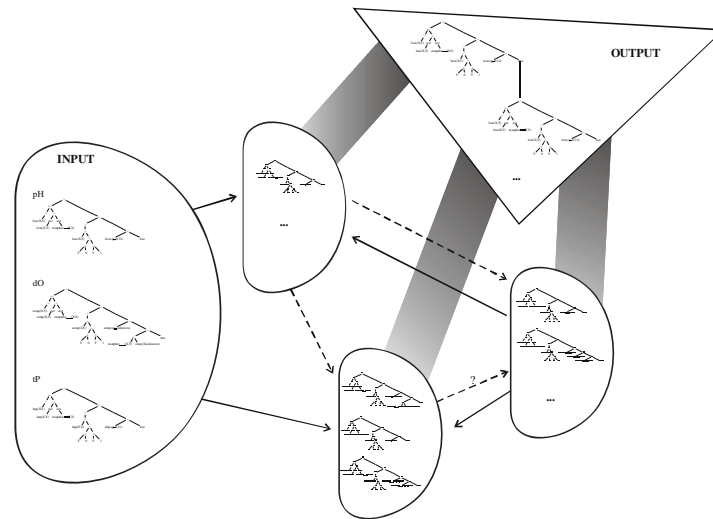
Figure 3 - A blended (i.e. INPUT) of the extensions of the predicates vo, ... whose evolution leads to an optimal Virtual Organization setting (i.e. OUTPUT).

Good faith is related to the ideas of fidelity, loyalty, honesty and trust in business (Lima and Varela, 1987). Good faith may be understood both in a psychological subjective sense and in an ethical objective sense (Lima and Varela, 1987). In the objective sense it consists "in considering correct behavior and not actor's mental attitudes" (Rotolo et al, 2005). It refers to both social norms and legal rules (Rotolo et al, 2005). In the subjective sense it has to do with knowledge and belief. "It regards the actor's sincere belief that she/he is not violating other people's rights" (Rotolo et al, 2005)

Good faith arises from general objective criteria related to loyalty and cooperation between parties. Good faith is an archetype of social behavior; loyalty in social relations, honest acting (Antunes, 1973), fidelity, reliability, faithfulness and fair dealing "and it comprises the protection of reasonable reliance" (Weitzenboeck, 2002).

Acting in bad faith in business may either "lead to the invalidation of some of the contract clauses or of the whole contract" (Rotolo et al, 2005) or even originate liabilities (Antunes, 1973).

The issue is not to wonder whether or not software agents may act in good faith or in bad faith; the question at stake is to consider that software agents acting in business relations will presumably negotiate and perform their acts according to certain standards of behavior. Yet, according to (Russel and Norvig, 2003) "an agent's behavior can be based on both its experience and the built-in-knowledge used in constructing the agent for the particular environment in which it operates", but autonomous systems will produce a behaviour "determined by its own experiences". Furthermore, agents may be "able to act strategically by calculating their best response given their opponents possible moves" (Binmore and Fun, 1992).

Good faith criteria relate to "objective standards of conduct" (Rotolo et al, 2005) that will help determine "whether the agent has observed reasonable commercial

standards of fair dealing in the negotiation and performance of the contract" (Weitzenboeck, 2002).

"The form given to the correctness rules allows to impose both positive and negative requirements to be fulfilled" (Rotolo et al, 2005).

Important issue is the one related to the attribution of the acts: should the acts of the electronic agent be attributed to the user who activated it, considering the electronic agent just as an instrument or tool at the disposal (and control??) of the user? Or should the volition of the agent be autonomously considered, since the user may not have been directly involved or consulted and may not even be aware that the agent acted at all (Weitzenboeck, 2002)? Either we assume or not the possibility, in a near future, of any sort of legal personality for software agents, the truth is that we probably should not rely on a legal fiction of attribution of the acts of agents to humans and, at least, it must be considered the autonomous will of the agent for purposes of good faith, bad faith, error on declaration and divergences between will and declaration. And the fact that an agent acts on good or bad faith will surely be of the utmost importance for all those (software or humans) who have to deal with it. Trust will thus become an unavoidable question for agents contracting.

3. TRUST AND SMART CONTRACTS

Trust is intimately related to beliefs. "Trust is a belief an agent has that the other party will do what it says it will (being honest and reliable)" (Ramchurn, et al, 2004). As (Rotolo et al, 2005) put it "agent x thinks that agent y not only is able to do certain actions but that y is willing to do what x needs" (formalization of a "Normative version" of Good Faith"). We can distinguish different levels of trust, the individual level ("an agent has some beliefs about the honesty or reciprocities nature of its interaction partners") and the system level ("the actors in the system are forced to be trustworthy by the rules of encounter (i.e. protocols and mechanisms) that regulate the system" (Ramchurn, et al, 2004).

At the system level it will assume the most relevance special protocols made in "such a way that they prevent agents from manipulating each other (e.g. through lies or collusion) so as to satisfy their selfish interests" (Ramchurn, et al, 2004).

At this system level it will be interesting to focus our attention in the figure of "smart contracts" which are seen as a "set of promises, specified in digital form, including protocols in which the parties perform on these promises" (Szabo, 1996). These contracts are really program codes imposing by itself an enforcement of the contract (the "terms of the contract are enforced by the logic of the program's execution", turning the breach of the contract, at least, quite expensive) (Miller and Stiegler, 2003). Indeed, one of the most relevant and difficult issues about inter-systemic electronic contracting is the one related to enforcement. Enforcement may be seen at two different levels, "publicly, through the court system, and privately, largely through reputation" (Friedman, 2000). The perspective of smart contracts tries to escape these difficulties - "Instead of enforcement, the contract creates an inescapable arrangement" (Miller and Stiegler, 2003).

Smart contracts can thus enhance the trust in electronic contracting. An interesting idea of this model is to use contracts as games. These games have rules - either fixed ones or sets of rules the players will choose (Miller and Stiegler, 2003) - which have to be followed in order to play the game. This is the upcoming of a new

and much more reliable form of "adherence contract", in which the contract is seen as an electronic game, managed or arbitrated by a board manager (a trusted third party)(Miller and Stiegler, 2003), which does not itself play the game but only allows the parties to make legal moves. And of course the board manager may be either a human or a software agent. Here, the contract may contain contractual clauses embedded in the software "in such a way to make breach of contract expensive (if desired, sometimes prohibitively so) for the breacher" (Szabo, 1996). In this model, trust is enhanced "by virtue of the different constraints imposed by the system" (Ramchurn, et al, 2004). As far as electronic contracting is concerned we may well consider that "public enforcement will work less well and private enforcement better than for contracts in real space at present" (Friedman, 2000). In cyberspace, reputation will turn to be a key issue and trusted third parties or Networks of Trust (widely trusted intermediary institutions or entities) must play an important role (Miller and Stiegler, 2003). Trust at the system level is based upon different possibilities, that is to say that it might be dependent on special interaction protocols (as is the case with the referred smart contracts), reputation mechanisms and security mechanisms. Reputation mechanisms are thus unavoidable to be considered as instruments required to foster the trustworthiness of electronic interactions (as it was referred already in our previous work (Andrade et al, 2005).

As far as security mechanisms are intended, we must refer the importance of authentication by trusted third parties, that is to say that information about the actors in the system specially delivered by trusted third parties may lead participants (human or software) to act upon what they think is trustable information. That is a special domain for trust and security in electronic relationships – here, as it happens with electronic signatures and time-stamp, the intervention of a trusted third party will be determinant for establishing participant's trust. Of course, this by itself will not be enough to "ensure that agents act and interact honestly and reliably towards each other. They will only represent a barrier against agents that are not allowed in the system".

In the end, trust will be highly dependent on the existence of social networks and on the traceability of past interactions among the agents in the community. This will be a fundamental issue for the existence of virtual organizations and of assuring a minimum reliability for the intervention of agents in it.

4. CONCLUSIONS

A traditional figure of commercial law, "consortium" will be highly enhanced in electronic commerce through appearance and acting of Virtual Enterprises, Virtual Organizations and Virtual Breeding Environments. These will highly depend (or its activities will be strengthened) by the use of software agents. Autonomy is a main advantage of software agents since they will act without any human intervention. Yet, the autonomy of software agents brings along several issues concerning their behavior and the legal consequences of it. One important issue has to do with good faith. Autonomous agents may act with good faith or bad faith, may comply or not with certain standards of behavior. In this sense it is important to understand if agents do comply with social or legal norms. Upon the activities and behavior of software agents engaging in business, it will be built a certain "image" of the agent and trust will be a mandatory requirement for commercial dealings. Trust may be

considered both at the individual or system level. At this level, it may become quite interesting to consider the issue of smart contracts as a way of enhancing trust and of achieving enforcement in electronic contracting.

Acknowledgments

5. REFERENCES

1. Abreu, J.C., Empresas Virtuais, i"Estudos em Homenagem ao Professor Doutor Inocêncio Galvão Telles", vol. IV – "Novos Estudos de Direito Privado", Almedina, 2003 (in Portuguese).
2. Abreu JC., Curso de Direito Comercial, Almedina, 2004 (in Portuguese).
3. Andrade F., Neves J., Novais P., Machado J., Abelha A., Legal Security and Credibility in Agent Based Virtual Enterprises, in Collaborative Networks and Their Breeding Environments, Camarinha-Matos L. Afsarmanesh H., Ortiz A., (Eds), Springer-Verlag, ISBN 0-387-28259-9, pp 501-512, 2005.
4. Andrade F., Novais P., Machado J., Neves J., Intelligent Contracting: Software agents, Corporate Bodies and Virtual Organizations, in Establishing the Foundation of Collaborative Networks, Camarinha-Matos L. Afsarmanesh H., Novais P., Analide C., (Eds), Springer-Verlag, Series: IFIP International Federation for Information Processing, ISBN: 978-0-387-73797-3, pp 217-224, 2007.
5. Antunes Varela J.M., Das Obrigações em Geral, Almedina, 1973, (in Portuguese).
6. Binmore, K., Fun and Games: A Text on Game Theory. D. C. Heath and Company, 1992.
7. Brazier, F., Kubbe, O., Oskamp, A., Wijngaards, N., Are Law-Abiding Agents Realistic? Proceedings of the workshop on the Law of Electronic Agents (LEA02), 2002.
8. Camarinha-Matos L, Afsarmanesh H, Ollus M. ECOLEAD: A holistic approach to creation and management of dynamic and virtual organizations, Collaborative Networks and Their Breeding Environments, Camarinha-Matos L. Afsarmanesh H., Ortiz A., (Eds), Springer-Verlag, ISBN 0-387-28259-9, pp 501-512, 2005.
9. Camarinha-Matos, L., Afsarmanesh H., The Emerging Discipline of Collaborative Networks. Virtual Enterprises and Collaborative Networks 2004: 3-16
10. Camarinha-Matos, L., Oliveira. A., Ratti R., Demsar D., Baldo, F., Jarimo, T., A Computer-Assisted Vo Creation Framework. Virtual Enterprises and Collaborative Networks 2007: 165-178
11. Crispim J, Sousa JP. A multi-Criteria support system for the formation of collaborative networks of enterprises, Collaborative Networks and Their Breeding Environments, Camarinha-Matos L. Afsarmanesh H., Ortiz A., (Eds), Springer-Verlag, ISBN 0-387-28259-9, pp 501-512, 2005.
12. Friedman, D., Contracts in cyberspace, American Law and Economics Association meeting, May 6, 2000.
13. Lima, Fernando Andrade Pires de / Varela, João de Matos Antunes : "Código Civil Anotado", vol. I, Coimbra Editora Limitada, 1987 (in Portuguese).
14. Miller, M., Stiegler, M., The digital path: smart contracts and the third world, Markets, Information and Communication. Austrian Perspectives on the Internet Economy Routledge 2003.
15. Neves J., Machado J., Analide C., Abelha A., Brito B. The Halt Condition in Genetic Programming. in Progress in Artificial Intelligence, Neves J., Santos M. and Machado J. (eds), Lecture Notes in Artificial Intelligence 4874 Springer, ISBN 978-3-540-77000-8, pp 160-169, 2007.
16. Ramchurn, SD., Huynh, D., Jennings, N., Trust in multiagent systems, The Knowledge Engineering Review 19 (1) 1-25, 2004.
17. Rotolo, A., G. Sartor, and C. Smith., Formalization of a 'Normative Version' of Good Faith. In A. Oskamp and C. Cevenini (eds.), Proc. LEA 2005. Nijmegen: Wolf Legal Publishers, 2005.
18. Russel S., Norvig P., Artificial Intelligence: A modern approach, Prentice-Hall, 2nd Ed., 2003, IBSN: 0-13-103805-2.
19. Szabo, N., Smart contracts: building blocks for digital markets, 1996. (http://szabo.best.vwh.net/smart.contracts.2.html)
20 Teubner G., Das Recht hybrider Netzwerke, ZHR, 2001.
21. Weitzenboeck, E., Good Faith and Fair Dealing in the Context of Contract Formation by Electronic Agents", Proceedings of the AISB 2002 Symposium on Intelligent Agents in Virtual Markets, 2-5 April 2002, Imperial College London.