Chapter 1

Introduction

1.1 Motivation

The sociologist Niklas Luhmann said "A complete absence of trust would prevent [one] even getting up in the morning" [Luhmann, 1979]. *Trust* is necessary in our everyday life. It is part of the "glue" that holds our society together. Without trust, governments could not rule and people cannot work cooperatively together. Trust helps to reduce the complexity of decisions that have to be taken in the presence of many risks.

Similarly, *reputation* is a universal concept that has been present in human societies for a long time. From ancient Greeks to modern days, from Vietnamese to Bedouins, the concept of reputation plays a very important role in human social organization. Reputation is one of the most relevant elements that we use to build trust in others.

Until recently, both concepts were applicable only to human societies and therefore they were a study field for sociologists, philosophers and psychologists. The irruption of Internet and the emergence of virtual (not necessarily human) societies add a new dimension to these old but very important concepts.

The scientific research in the area of trust and reputation mechanisms for virtual societies is a recent discipline oriented to increase the reliability and performance of electronic communities by introducing in such communities these well known human social control mechanisms. Computer science has moved from the paradigm of an isolated machine to the paradigm of a network of systems and of distributed computing. Likewise, artificial intelligence is quickly moving from the paradigm of an isolated and non-situated intelligence to the paradigm of situated, social and collective intelligence. The new paradigm of the so called intelligent or adaptive agents and Multi-Agent Systems (MAS) together with the spectacular emergence of the information society technologies (specially reflected by the popularization of electronic commerce) are responsible for the increasing interest on trust and reputation mechanisms applied to electronic societies.

An agent is a computer system capable of flexible autonomous action in a dynamic, unpredictable and open environment endowed with the capacity to interact with other systems (artificial or natural). Agents are often deployed in environments in which they

interact, compete, and maybe cooperate, with other agents that have possibly conflicting aims. Such environments are known as multi-agent systems and are called to become a key element of the information society. In this context, trust and reputation play a similar role that in human societies.

Up to now, the design of such systems has been approached using traditional software development methods. However, the special characteristics of these systems suggest the necessity of more specific techniques adapted to its peculiarities.

1.2 Overview and main contributions

The work of this thesis contributes to the state of the art in two areas.

First, in the area of computational trust and reputation models, we present ReGreT. ReGreT is a modular trust and reputation system oriented to complex e-commerce environments where social relations play an important role. ReGreT follows a mathematical approach to the problem, with social ingredients that improve the calculation of trust and reputation values. There is a lot of work done in computational trust and reputation models, however little attention has been given to the social aspect of both concepts. With our model we want to strengthen a line of research that we think is under-explored and very promising.

The main characteristics of the ReGreT system can be summarized as follows:

- It takes into account direct experiences, information from third party agents and social structures to calculate trust, reputation and credibility values.
- It has a trust model based on direct experiences and reputation.
- It incorporates an advanced reputation model that works with transmitted and social knowledge.
- It has a credibility module to evaluate the truthfulness of information received from third party agents.
- It uses social network analysis to improve the knowledge about the surrounding society (specially when no direct experiences are available).
- It provides a degree of reliability for the trust, reputation and credibility values that helps the agent to decide if it is sensible or not to use them in the agent's decision making process.
- It can adapt to situations of partial information and improve gradually its accuracy when new information becomes available.
- It can manage at the same time different trust and reputation values associated to different behavioural aspects. Also it can combine reputation and trust values linked to simple aspects in order to calculate values associated to more complex attributes.

The study of trust and reputation has many applications in Information and Communication technology. Trust and reputation systems have been recognized as key factors for successful electronic commerce adoption. These systems are used by intelligent software agents both as a mechanism of search for trustworthy exchange partners and as an incentive in decision-making about whether or not to honour contracts. Reputation is used in electronic markets as a trust-enforcing, deterrent, and incentive mechanism to avoid cheaters and frauds. Another area of application in agent technology is teamwork and cooperation.

Besides technological issues, if we want people to massively enter the information society era, we have to improve the real and perceived security of electronic interactions. For instance, it is well known that lack of trust is one of the main reasons for consumers, as well as companies, not engaging in electronic commerce. This lack of confidence is even worse when users have to rely on autonomous agents that act on their behalf. With no doubt, low level security measures are important and necessary. However, building user confidence in e-Commerce (and electronic interactions in general) is more than secure communication via electronic networks as can be obtained with, for example public key cryptography techniques. Here is where trust and reputation mechanisms come to scene. As in human societies, electronic communities have started to use trust and reputation as a social control mechanism that complements more expeditious approaches (based on social constructs and their corresponding disciplinary actions to punish deception and fraud). Further study on trust and reputation mechanisms will directly contribute increasing both the reliability and performance of electronic communities and the confidence that humans deserve on information society technologies.

The second contribution is in the area of agent design. As we have pointed out in the motivations section, traditional software development methods do not fully cover the requirements for the design and implementation of autonomous agents. Several approaches have been proposed to overcome this problem. However these proposals usually leave a gap between specification and implementation, enforcing a particular view of architecture upon the specification or do not make explicit structures for modelling the components of an architecture or the relationships between them.

Following the steps of Parsons, Sierra and Jennings [Parsons et al., 1998] we propose the use of multi-context systems for the design and implementation of autonomous agents.

Multi-context systems provide an overarching framework that allows distinct theoretical components to be defined and interrelated. Such systems consists of a set of contexts, each of which can informally be considered a set of formulae written in a (possibly) different logic, and a set of bridge rules for transferring information between contexts. From a software engineering perspective, multi-context systems support modular decomposition and encapsulation but what is more important for our purposes, from a logical modelling perspective they provide an efficient means of specifying and executing complex logics. We extend the multi-context theory with several elements that make the specification of dynamic components, as those needed to build autonomous agents, easier.

As a nexus between both contributions, we show how it is possible to specify the

ReGreT system using the multi-context approach.

1.3 Publications

The work presented in this thesis has generated the following set of publications:

- J. Sabater, C. Sierra, S. Parsons, N. Jennings (2002) Engineering executable agents using multi-context systems, Journal of Logic and Computation. Vol 12, n 3, pp. 413-442.

- J. Sabater, C. Sierra (2002) Reputation and social network analysis in multi-agent systems, Proc. "First International Conference on Autonomous Agents and Multiagent systems (AAMAS-02)", Bologna, Italy, (July 15-19), pp. 475-482.

- J.Sabater, C. Sierra (2002) Social aspects or ReGreT, a reputation model based on social relations, Proc. "5è Congrés Català d'Intel.ligència Artificial (CCIA-02)", Castelló de la Plana, Spain, (October 24-25), pp. 336-343

- J. Sabater, C. Sierra (2002) Social ReGreT, a reputation model based on social relations, SIGecom Exchanges. ACM, Vol 3.1, pp 44-56. (http://www.acm.org/sigecom/exchanges/volume_3_(02)/3.1-Sabater.pdf)

- J. Sabater, C. Sierra (2001) ReGreT: A reputation model for gregarious societies (v.1), Proc. "Fourth Workshop on Deception Fraud and Trust in Agent Societies", Montreal, Canada, (May 28, June 9), pp. 61-70.

  (Also published at Proc. "4t Congrés Català d'Intel.ligència Artificial", Barcelona, Spain, (October 24-25), pp. 214-222)

- J. Sabater, C. Sierra (2001) ReGreT: A reputation model for gregarious societies (v.2), Proc. "Fifth International Conference on Autonomous Agents", Montreal, Canada, (May 28, June 9), pp. 194-195

- J. Sabater, C. Sierra, S. Parsons, N. R. Jennings (1999) Using multi-context systems to engineer executable agents Proc. "6th Int.Workshop on Agent Theories Architectures and Languages (ATAL-99)", Orlando, Florida, USA, (July 15-17), pp. 131-148.

  Revised version in: Intelligent Agents VI (eds N. R. Jennings and L. Lesperance) LNAI 1757 pp. 277-294.

  (Also published at UKMAS 99 (2nd workshop of the UK special interest group on multi-agent systems, Bristol, UK, (December 6-7)))

  (Reduced version published -in Catalan- at "2n Congrés Català d'Intel.ligència Artificial", Girona, Spain, (October 25-27), pp. 185-191.)

1.4 Structure of the thesis

The thesis is organized in eight chapters and two appendices that are distributed in three blocks: related work (chapter 2), the ReGreT system (chapters 3, 4, 7 and appendix A) and the use of multi-context systems to engineer autonomous agents (chapters 5, 6 and appendix B).

Chapter 2: we analyse a representative set of computational trust and reputation models as well as some frameworks/test-beds currently used to evaluate these models. We also propose a classification of the models according to a set of relevant aspects associated to trust and reputation.

Chapter 3: this chapter introduces the *SuppWorld* framework, a flexible framework specially designed to test trust and reputation models in a complex environment where social relations play an important role.

Chapter 4: in this chapter *ReGreT* is presented, a trust and reputation system that gives special relevance to the social relationships among individuals in virtual societies.

Chapter 5: is devoted to the use of multi-context systems as a means of specifying and implementing agent architectures. We propose some extensions to the multi-context base theory and, by means of two didactic examples, we show how it works in practice.

Chapter 6: this chapter puts together the two main threads of this thesis: the trust and reputation models and the multi-context approach for the design of autonomous agents. We take the *ReGreT* system described in chapter 4 and specify it by using the multi-context approach presented in chapter 5.

Chapter 7: an initial set of experimental results for the ReGreT system are presented in this chapter. Using the framework described in chapter 3 we deploy a set of scenarios that are used to test the ReGreT system capabilities.

Chapter 8: summarizes the conclusions of this work and shows future directions.

Appendix A: this appendix presents the unit theories of the *ReGreT* system specification made in chapter 6.

Appendix B: this appendix details the *SuppWorld* framework configuration files used in the experiments of chapter 7.