

# Reputation and Social Network Analysis in Multi-Agent Systems

Jordi Sabater

IIIA - Artificial Intelligence Research Institute  
CSIC - Spanish Scientific Research Council  
Bellaterra, Catalonia, Spain

jsabater@iiia.csic.es

Carles Sierra

IIIA - Artificial Intelligence Research Institute  
CSIC - Spanish Scientific Research Council  
Bellaterra, Catalonia, Spain

sierra@iiia.csic.es

## ABSTRACT

The use of previous direct interactions is probably the best way to calculate a reputation but, unfortunately this information is not always available. This is especially true in large multi-agent systems where interaction is scarce. In this paper we present a reputation system that takes advantage, among other things, of social relations between agents to overcome this problem.

## Categories and Subject Descriptors

I.2.11 [Distributed Artificial Intelligence]: Multiagent systems; I.2.11 [Distributed Artificial Intelligence]: Intelligent agents; I.2.0 [General]: Cognitive simulation

## General Terms

Algorithms, Design, Theory

## 1. INTRODUCTION

The study and modeling of reputation has attracted the interest of scientists from different fields such as: sociology [10, 5], economics [7, 12], psychology [4, 11] and computer science [6, 3, 18]. According to the Oxford English Dictionary, *reputation* is “the common or general estimate of a person with respect to character or other qualities”. This estimate is necessarily formed and updated over time with the help of different sources of information. Up to now, the computational models of reputation have been considering two different sources: (i) the direct interactions and (ii) the information provided by other members of the society about experiences they had in the past [14, 15, 17, 18]. Those systems, however, forget a third source of information that can be very useful. As a direct consequence of the interactions it is possible (even in simple societies) to identify different types of social relations between their members. Sociologists and psychologists have been studying these social networks in human societies for a long time and

also how these social networks can be used to analyse trust and reputation [8, 5]. These studies show that it is possible to say a lot about the behaviour of individuals using the information obtained from the analysis of their social network. In this paper we extend the **Regret** [14] system to incorporate social networks in the reputation model.

In Section 2 we introduce the notion of social network analysis and its application to agent communities. In Section 3, a scenario is presented that will be used in the rest of the paper to illustrate how the system works. Sections 4—8 describe the system in detail. Finally sections 9 and 10 present the related work, conclusions and future work.

## 2. SOCIAL NETWORK ANALYSIS AND AGENT SOCIETIES

Social network analysis is the study of social relationships between individuals in a society. Social network analysis emerged as a set of methods for the analysis of social structures, methods that specifically allow an investigation of the relational aspects of these structures. The use of these methods, therefore, depends on the availability of relational rather than attribute data [16].

Relational data is represented using graphs called *sociograms*. A different sociogram is usually built for each social relation under study. Depending on the type of relation we have a directed or non-directed sociogram, with weighted edges or without.

Obviously, the more relational data the better the network analysis is. However, these data can be difficult to obtain. Sociologists usually obtain them through public-opinion polls and interviews with the individuals. This procedure is, of course, not possible in agent societies. Moreover, the analysis of human social structures is usually done by a sociologist external to the society. This external position gives the analyst a privileged watchtower to make this analysis. In our case, as we want to use social analysis as part of the reputation system to be included in each agent, each agent has to do this analysis from its own perspective.

It is beyond the scope of this paper to propose solutions about the way an agent builds such sociograms. From now on, we will assume that the agent owns a set of sociograms that show the social relations in its environment. This sociograms are not necessarily complete or accurate. We suppose they are built by each agent using the knowledge it has about the environment. Therefore, sociograms are dynamic and agent dependent.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

AAMAS'02, July 15-19, 2002, Bologna, Italy.

Copyright 2002 ACM 1-58113-480-0/02/0007 ...\$5.00.

### 3. RUNNING EXAMPLE

The scenario for this running example is a marketplace. To simplify, the buyers buy one product at a time (it is not possible to buy in bulk) and the elements that are taken into account for each transaction are price, quality and delivery date. The buyer chooses a seller. If the seller wants to deal with that buyer, then it sends an offer. The buyer has then to decide if he wants to accept the offer or not. If the buyer accepts, the transaction is performed.

It is important to note that the actual result of the transaction may not necessarily be equal to the initial offer from the seller. A swindler seller can increase the price, decrease the quality or deliver the product late. On the other hand, the buyer can decide to pay less or even not to pay at all. We consider that the process of obtaining the actual result of the transaction is atomic, that is, the buyer and the seller have to decide their strategy (to pay less, to overcharge the price, do things exactly as specified in the offer, etc.) before knowing the behaviour of the other. Although this assumption is not very realistic, it makes the example simpler and allows us to focus our attention only on reputation.

We can identify many types of relationships between agents in an e-commerce environment and it is beyond the scope of this paper to present a survey on all of them. We will refer just to a set of relations that can be found in our scenario and that will help the reader to go through the explanation of the system.

We assume that in our scenario the great majority of agents are rational and have a behaviour according to their goals and affiliation (in an environment with random or contradictory agents it makes little sense to use social relationships to predict the behaviour of agents). We note a non-directed relation of type *rel* between two agents *a* and *b* as *rel(a,b)*. The three relation types that we consider to illustrate this scenario are:

1) *Competition (comp)*. This is the type of relation found between two agents that pursue the same goals and need the same (usually scarce) resources. In this situation, agents tend to use all the available mechanisms to take some advantage over their competitors, for instance hiding information or lying. In our running example this is the kind of relation that could appear between two sellers that sell the same product or between two buyers that need the same product.

2) *Cooperation (coop)*. This relation implies significant exchange of sincere information between the agents and some kind of predisposition to help each other if possible. Notice that we talk about “sincere” information instead of “true” information because the agent who gives the information *believes* it is true. To simplify, we consider that two agents cannot have at the same time a competitive and a cooperative relation.

3) *Trade (trd)*. This type of relation reflects the existence of commercial transactions between two agents and is compatible either with cooperative or competitive relations.

As we said before, each agent owns a sociogram for each of these relation types. The three sociograms are non-directed graphs with weighted edges. Weights go from 0 to 1 and reflects the intensity of the relation.

In our scenario, the variables that appear in contracts between buyers and sellers (price, quality and delivery date) determine the reputation types. To illustrate the model, we consider four reputation types for the seller:

**to\_overcharge**: A high reputation value of this type

means that the seller tends to overcharge the prices specified in contracts.

**to\_deliver\_late**: As the name suggest, is the reputation of delivering the products later than the delivery date specified in the contract.

**quality\_swindler**: A seller with a reputation to deliver products with less quality than specified in the contract.

**swindler**: A swindler is a seller that overcharges the price and/or delivers products with a quality lower than the quality specified in the contract.

For the buyer we illustrate the model with the type:

**defaulter**: A high reputation of this type means that the buyer does not pay for the products he buys.

### 4. THE REGRET SYSTEM

The **Regret** system structure is based on what we call the three *dimensions* of reputation. If an individual is considering only its direct interaction with the other members of the society to establish reputations we say that the agent is using the *individual dimension*. If the individual also uses the information coming from other members of the society and the social relations, we are talking about the *social dimension*. Finally, we consider that the reputation of an individual is not a single and abstract concept but rather a multi-facet concept. For example, the reputation of being a good carrier summarizes the reputation of having good planes, the reputation of never losing luggage and the reputation of serving good food. The different types of reputation and how they are combined to obtain new types is the base of the third *dimension* of reputation, the *ontological dimension*. In the following sections we explain in detail each one of these dimensions. In section 8 we show how these *dimensions* are combined to obtain a single value for the reputation.

Although reputations also have a temporal aspect (the reputation value of an agent varies along time), we will omit the reference to time in the notation in order to make it more readable. We will refer to the agent that is calculating a reputation as *a* (what we call the “source agent”) and the agent that is the object of this calculation as *b* (what we call the “target agent”).

### 5. INDIVIDUAL DIMENSION

We define the *outcome* of a dialogue between two agents as:

- An initial contract to take a particular course of action and the actual result of the actions taken.
- An initial contract to fix the terms and conditions of a transaction and the actual values of the terms of the transaction.

An outcome is represented as a tuple of the form  $o = (a, b, I, X^c, X, t)$  where *a* and *b* are the agents implied in the contract, *I* a set of indexes that identify the issues of the contract,  $X^c$  and *X* two vectors with the agreed values of the contract and the actual values after its fulfillment respectively, and *t* the time when the contract was signed. We use a subindex  $i \in I$  to refer to the specific value of issue *i* in vectors  $X^c$  and *X*. For instance, in the example we have  $I = \{Price, Delivery\_Date, Quality\}$ . If we want to make reference to the *Price* value in the vector  $X^c$  we use the notation  $X^c_{Price}$ .

$ODB$  is defined as the set of all possible outcomes.  $ODB^{a,b} \subseteq ODB$  is the set of outcomes that agent  $a$  has signed with agent  $b$ . We define  $ODB_{\{i_1, \dots, i_n\}}^{a,b} \subseteq ODB^{a,b}$  as the set of outcomes that have  $\{i_1, \dots, i_n\}$  as issues in the contract. For example,  $ODB_{\{Price\}}^{a,b}$  is the set of outcomes that has agent  $a$  from previous interactions with agent  $b$  and that fix, at least, the value for the issue  $Price$ .

The *individual dimension* models the direct interaction between two agents. The reputation that takes into account this dimension is the most reliable. This is because it takes into account all the peculiarities of the target agent. We call *outcome reputation* (noted as  $R_{a \rightarrow b}(\varphi)$  where  $\varphi$  is the reputation type) the reputation calculated directly from an agent's outcomes database.

The subset of issues of an outcome taken into account to calculate a given reputation type  $\varphi$  is domain dependent. We define a *grounding relation* ( $gr$ ) as the relation that links a reputation type  $\varphi$  with a list of issues. This set of issues allows us to select the right subset of outcomes from the general outcomes' data base. Each issue is a tuple with the form  $(I_i, \{+, -\}, \alpha_i)$ . The first parameter ( $I_i$ ) is a label that identifies the issue. The second parameter ( $\{+, -\}$ ) indicates how an increment of the value of the issue affects the reputation, that is, a  $+$  means that if the value of the issue increases, the reputation also increases while a  $-$  means that if the value of the issue increases, the reputation decreases. Finally, the last parameter is the weight that that issue has in the general calculation of the reputation.

As an example, the *grounding relation* for a seller in our scenario is defined in the following table:

$\varphi$	$gr(\varphi)$
to_overcharge	$\{(Price, +, 1)\}$
to_deliver_late	$\{(Delivery\_Date, +, 1)\}$
quality_swindler	$\{(Quality, -, 1)\}$

Note that we only define the *grounding relation* for the reputation types `to_overcharge`, `to_deliver_late` and `quality_swindler`. The reputation type `swindler` is a complex reputation type calculated from more basic reputations as we will see in section 7.

To calculate an outcome reputation we use a weighted mean of the outcomes evaluation, giving more relevance to recent outcomes.<sup>1</sup>

$$R_{a \rightarrow b}(\varphi) = \sum_{o_i \in ODB_{gr(\varphi)}^{a,b}} \rho(t, t_i) \cdot Imp(o_i, gr(\varphi))$$

with  $\rho(t, t_i) = \frac{f(t_i, t)}{\sum_{o_j \in ODB_{gr(\varphi)}^{a,b}} f(t_j, t)}$  where  $t$  is the actual

time and  $f(t_i, t)$  is a time dependent function that gives higher values to values closer to  $t$ . A simple example of this type of function is  $f(t_i, t) = \frac{t_i}{t}$ .  $Imp(o_i, gr(\varphi))$  is the evaluation of the outcome  $o_i$  taking into account the actual values of a concrete set of issues, those associated with  $\varphi$  by the relation  $gr$ . The evaluation of an outcome  $o = (a, b, I, X^c, X, t)$  is defined as a function of the difference between the utility of the contract and the utility of the fulfillment of that contract:

$$Imp(o, gr(\varphi)) = g(V(X^s) - V(X^c))$$

where  $V(X^c)$  is the utility function over the vector of values  $X^c$  and  $X^s$  is a vector build using the following formula:

$$X_i^s = \begin{cases} X_i & \text{if } i \in gr(\varphi) \\ X_i^c & \text{otherwise} \end{cases}$$

In other words, we obtain this vector from the  $X^c$  vector replacing the values specified in the index  $gr(\varphi)$  by the values in the same position in vector  $X$ . Finally,  $g$  is a function that models the behaviour of the agent according to the degree of deception or reward obtained after the analysis of the outcome. An appropriate function is:

$$g(x) = \sin\left(\frac{\pi}{2}x\right)$$

Using this function, an agent penalizes a deception in the fulfillment of a contract by giving values near  $-1$  when  $V(X^s) < V(X^c)$  and values near  $1$  when  $V(X^s) > V(X^c)$ .

Besides the reputation value, it is important to know how reliable is it. There are many elements that can be taken into account to calculate how reliable an outcome reputation is but we will focus on two of them: the number of outcomes used to calculate the reputation value and the variability of their rating values. This approach is similar to that used in the Sporas system [18].

The intuition behind the number of outcomes factor (noted as  $No$ ) is that an isolated experience (or a few of them) is not enough to make a correct judgement about somebody. You need certain amount of experiences before you can say how an agent is. As the number of outcomes grows, the reliability degree increases until it reaches a maximum value, what we call the *intimate* level of interactions (*itm* from now on). From a social point of view, this stage is what we know as a close relation. More experiences will not increase the reliability of our opinion from then on. The next simple function is the one we use to model this:

$$No(ODB_{gr(\varphi)}^{a,b}) = \begin{cases} \sin\left(\frac{\pi \cdot |ODB_{gr(\varphi)}^{a,b}|}{2 \cdot itm}\right) & |ODB_{gr(\varphi)}^{a,b}| \leq itm \\ 1 & \text{otherwise} \end{cases}$$

We believe that the *itm* value is domain dependent: it depends on the interaction frequency of the individuals in that society.

The outcome reputation deviation (noted as  $Dt$ ) is the other factor that our system takes into account to determine the reliability of an outcome reputation. The greater the variability in the rating values the more volatile will the other agent be in the fulfillment of their agreements. To have a measure of this variability we take into account the impact on the expected utility of the actual execution of the contracts.

We calculate the outcome reputation deviation as:

$$Dt(ODB_{gr(\varphi)}^{a,b}) = \sum_{o_i} \rho(t, t_i) \cdot |Imp(o_i, gr(\varphi)) - R_{a \rightarrow b}(\varphi)|$$

Where  $o_i \in ODB_{gr(\varphi)}^{a,b}$ . This value goes from 0 to 1. A deviation value near 1 indicates a high variability in the rating values (that is, a low credibility of the reputation value from the outcome reputation deviation point of view) while a value close to 0 indicates a low variability (that is, a high credibility of the reputation value).

<sup>1</sup>There are many psychological studies that support recency as a determinant factor [11].

Finally, we define the reliability of an outcome reputation as a convex combination of the function  $No$  and the outcomes evaluation rating deviation  $Dt$ .

$$RL_{a \rightarrow b}(\varphi) = (1 - \mu) \cdot No(ODB_{gr(\varphi)}^{a,b}) + \mu \cdot (1 - Dt(ODB_{gr(\varphi)}^{a,b}))$$

## 6. SOCIAL DIMENSION

Although direct interaction is the most reliable source of information, unfortunately it is not always available. Not only because the agent can be a newcomer to a society but also because the society can be very large. Therefore, when the interactions with another agent are scarce it is not possible to assign it a reputation based just on direct experiences. It is in these situations when the *social dimension* of an agent may help by using information coming from other agents. In the **Regret** system we use three types of social reputation depending on the information source:

**Witness Reputation.** Based on the information about the target agent coming from other agents. We note this reputation as:  $R_{a \rightarrow b}^w(\varphi)$

**Neighbourhood Reputation.** Uses the social environment of the target agent, that is, the neighbours of the target agent and their relations with it. Noted as:  $R_{a \rightarrow b}^N(\varphi)$

**System Reputation.** It is a default reputation value based on the role played by the target agent. Noted as:  $R_{a \rightarrow b}^s(\varphi)$

Each one of these reputations requires a different degree of knowledge of the agent society and the target agent. The *System Reputation* is the easiest to calculate. We are assuming that the role an agent is playing is always “visible” information that is available to all the agents in the society. However, the role alone does not convey enough information to compute a reputation on all imaginable aspects. Also, the reliability of this type of reputation tends to be low because it doesn’t take into account the peculiarities of the individual and its environment. This is the kind of reputation that an agent can use when it is a newcomer and there is an important lack of interaction with the other agents in the society. The *Witness Reputation* and the *Neighbourhood Reputation*, on the other hand, demand from the agent a moderate to hard knowledge of the social relations in the agent community.

Sociologically speaking, this division is far to be complete. However, we consider that with these three types we maintain a good compromise between the complexity of the model and the requirements that an agent can have in an e-commerce environment. We explain below how to calculate these reputation values.

### 6.1 Witness reputation

Beliefs about the reputation of others can be shared and communicated by the members of a society. The reputation that an agent builds on another agent based on the beliefs gathered from society members (witnesses) is what we call *witness reputation*. In an ideal world, with only homogeneous and trusted agents, this information is as relevant as the direct information. However, in the kind of scenarios we are considering, it may happen that:

**The information is false.** Either because the other agents are trying to lie or because the information they own is not accurate, an agent has to be prepared to deal with false information.

**Agents hide information.** An agent cannot assume that the information is complete.

Besides that, the information that comes from other agents can be correlated (what is called the *correlated evidence problem* [13]). This happens when the opinions of different witnesses are based on the same event(s) or when there is a considerable amount of shared information that tends to unify the witnesses’ way of “thinking”. In both cases, the trust on the information shouldn’t be as high as the number of similar opinions may suggest. Because the event(s) that have generated the opinions for each agent may be hidden, the agent cannot identify directly which agents are correlated. Schillo et. al [15] propose a method based on the analysis of “lying” as a stochastic process to implicitly reconstruct witness observations in order to alleviate this problem. We take a different approach based on the social relations between agents. Analysing these relations, an agent can obtain useful information to minimize the effects of the correlated evidence problem.

We assume that the information to be exchanged between agents is a tuple where the first element is the reputation value of the target agent from the point of view of the witness, and the second element is a value that reflects how confident the witness is about the reputation value. As we said before, the witness can give wrong values or simply can decide not to give its opinion even if he has enough information to do so. We will note the tuple as  $(R_{c \rightarrow b}(\varphi), RL_{c \rightarrow b}(\varphi))$ , where  $c$  is the agent giving the information to  $a$ .

#### 6.1.1 Identifying the witnesses

The first step to calculate a witness reputation is to identify the set of witnesses (**W**). The initial set of potential witnesses might be the set of all agents that have interacted with the target agent in the past. In the example, the initial set is composed by all the agents that had had a trade relation with the target (it seems logical to think that the best witnesses about the commercial behaviour of the target agent are those agents that had a trade relation with it before). This set, however, can be very big and its members probably suffer from the correlated evidence problem.

We take the stance that grouping agents with frequent interactions among them and considering each one of these groups as a single source of reputation values minimizes the correlated evidence problem. Moreover, assuming that asking for information has a cost, it has no sense to ask the same thing to agents that we expect will give us more or less the same information. Grouping agents and asking for information to the most representative agent within each group reduces the number of queries to be done. A domain dependent sociogram is what **Regret** uses to build these groups and to decide who is their most representative agent (in our example the sociogram of the cooperative relation).

There are many heuristics that can be used to find groups and to select the best agent to ask. In the **Regret** system we use a heuristic based on the work by Hage and Harary [9]. Taking the subset of the selected sociogram over the agents that had had interactions with the target agent as the initial graph, the heuristic that **Regret** follows is:

1. To identify the *components* of the graph. A *component* is defined as a maximally connected subgraph.
2. To find the set of *cut-points* (*CP*) for each component. A *cut-point* is a node whose removal would increase the

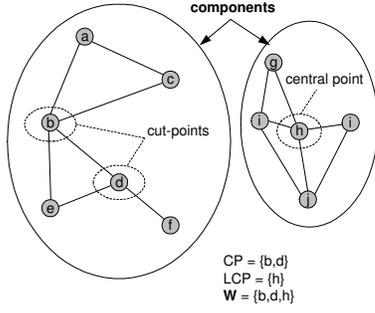


Figure 1: Witness selection within Regret.

number of components by dividing the sub-graph into two or more separate sub-graphs among which there are no connections. A cut-point can be seen from a sociological point of view as indicating some kind of *local centrality*. Cut-points are pivotal points of articulation between the agents that make up a component [16].

3. For each component that does not have cut-points, to choose as a representative for that component the node with the larger degree. If there is more than one node with the maximum degree, choose one randomly. This point is called a *central point*. The degree can be regarded also as a measure of *local centrality* [16]. We refer to this set of nodes as *LCP*.
4. The set of selected nodes is the union between the set of *cut-points* and the set of *LCP*. That is,  $\mathbf{W} = CP \cup LCP$ .

Figure 1 shows an example of the application of the heuristic. At this point, the agent has to ask for information to all the agents in the so calculated set of witnesses  $\mathbf{W}$ .

### 6.1.2 Who can I trust?

Once the information is gathered we obtain

$$\{\langle R_{w_i \rightarrow b}(\varphi), RL_{w_i \rightarrow b}(\varphi) \rangle \mid w_i \in \mathbf{W}' \subseteq \mathbf{W}\}$$

where  $\mathbf{W}'$  is the subset of agents that answered  $a$ 's query. The next step is to aggregate these values to obtain a single value for the *Witness Reputation*. However, as we said before, it is possible that this information be false so the agent has to be careful to give the right degree of importance and reliability to each piece of information. The degree of importance relies on the trust that each witness has. The system uses two different methods to calculate this trust: *social trust* and *outcome trust reputation*.

We define  $socialTrust(a, w_i, b)$  as the trust degree that agent  $a$  assigns to  $w_i$  when  $w_i$  is giving information about  $b$  and taking into account the social relations among  $a$ ,  $w_i$  and  $b$ .

**Regret** uses fuzzy rules [19] to determine how a social structure provides a reliability degree on the information coming from a given agent in that structure. The antecedent of each rule is the type and degree of a social relation and the consequent is the reliability of the information from the point of view of that social relation. In our scenario, a possible rule would be:

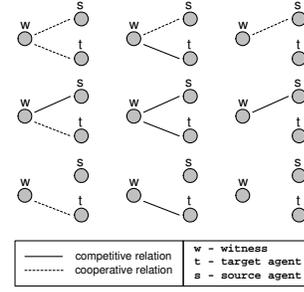


Figure 2: Relevant social structures in the example.

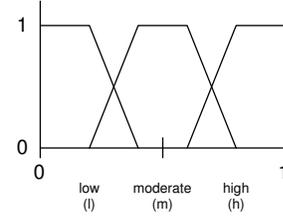


Figure 3: Intensity of a social relation.

IF  $coop(w_i, b)$  is high  
THEN  $socialTrust(a, w_i, b)$  is very\_bad

that is, if the level of cooperation between  $w_i$  and  $b$  is high then the trust from the point of view of  $a$  on the information coming from  $w_i$  related to  $b$  is very bad. The heuristic behind this rule is that a cooperative relation implies some degree of complicity between the agents that share this relation so the information coming from one about the other is probably biased.

From the set of social relations in our scenario, only the cooperative relation and the competitive relation are relevant to calculate a measure of reliability. Which relations are relevant to calculate the reliability depend on the meaning that each relation has in the specific agent society. In our scenario, for instance, a trade relation cannot cast any light on the reliability of an agent from the point of view of social analysis. In other societies, however, this could be the other way around.

Hence, together with the “no relation” possibility and with the fact that the most relevant links are between the witness and the agent and the witness and the target, there are 9 social structures to be considered as shown in Figure 2.

Figure 3 contains the fuzzy set values -defined over the intensity label on the arcs in the sociogram- for the variables  $coop(w_i, a)$ ,  $coop(w_i, b)$ ,  $comp(w_i, a)$ , and  $comp(w_i, b)$ , similarly we can define the fuzzy set values for the variable  $socialTrust(a, w_i, b)$ . Although at this moment the kind of influence of each social structure is static and based in human common sense, we plan to improve the system with a rule learning mechanism.

A second way to calculate the degree of trust of an agent is using the *outcome trust reputation* of the ‘trust’ that other agent deserves, that is,  $R_{a \rightarrow b}(\text{trust})$ . The *outcome trust reputation* is calculated like any other outcome reputation (see section 5). In our running example the *grounding relation*

for this reputation type could be:

$\varphi$	$gr(\varphi)$
trust	$\{(Price, -, 0.3), (Quality, +, 0.4), (Delivery\_Date, -, 0.3)\}$

That is, an agent that respects the price, the quality and the delivery date of a contract is a trusty agent. We could also take the informations coming from a witness into account as well as how accurate are these informations. This could become the set of outcomes used to calculate the *outcome trust reputation* for that witness.

The trust values calculated using an outcome trust reputation are more useful than those based on social relations (socialTrust) because an outcome trust reputation is based on the individual experiences and thus takes into account its particularities while the analysis of social structures rely on global expected behaviours. However, in those situations where there is not enough information to calculate a reliable outcome trust reputation, the analysis of social relations can be a good solution. Usually, social relations are easier to obtain than a set of outcomes (necessary to calculate an outcome trust reputation) especially if the agent has just arrived to a new scenario. Given that, we define  $a$ 's trust degree for an agent  $w_i$  when it is giving information about  $b$  as:

$$trust(a, w_i, b) = RL_{a \rightarrow w_i}(\varphi) \cdot R_{a \rightarrow w_i}(\varphi) + (1 - RL_{a \rightarrow w_i}(\varphi)) \cdot socialTrust(a, w_i, b)$$

That is, the agent uses the trust reputation based on direct interactions if it is reliable, if not, it uses the social trust.

Now we have all the necessary tools to calculate the *witness reputation* and its reliability considering that the information coming from the witnesses can be false. The formulae we propose to calculate these values are:

$$R_{a \rightarrow b}(\varphi) = \sum_{w_i \in \mathbf{W}'} \omega^{w_i b} \cdot R_{w_i \rightarrow b}(\varphi)$$

$$RL_{a \rightarrow b}(\varphi) = \sum_{w_i \in \mathbf{W}'} \omega^{w_i b} \cdot \min(trust(a, w_i, b), RL_{w_i \rightarrow b}(\varphi))$$

$$\text{where } \omega^{w_i b} = \frac{trust(a, w_i, b)}{\sum_{w_j \in \mathbf{W}'} trust(a, w_j, b)}$$

These formulae require some explanations. To calculate the *witness reputation* the agent uses the normalized trust of the witness to weight each opinion in the final value. For the calculation of the reliability, we want that in the final value, the contribution of each individual reliability be in the same proportion that its related reputation. Therefore, the agent uses the same weights in the reliability formula as in the reputation formula. Finally to calculate the reliability of an individual reputation, the agent uses the minimum between the trust of the witness that sent the reputation and the reliability that the witness himself gives to that reputation. We use this method to model the idea that if the witness is a trusty agent, then we can use his/her own measure of reliability for the reputation, if not, we cannot rely on his/her information and we have to calculate our own measure of trustworthiness for that reputation based on the outcomes and the social relations of that witness.

## 6.2 Neighbourhood reputation

The reputations of the individuals that are in the neighbourhood of the target agent and their relation with him are the elements used to calculate what we call the *Neighbourhood Reputation*. Neighbourhood in a MAS is not related with the physical location of the agents but with the links created through interaction. The main idea is that the behaviour of these neighbours and the kind of relation they have with the target agent can give some clues about its possible behaviour. The set of neighbours of an agent  $b$  is noted as  $\mathbf{N}_b = \{n_1, n_2, \dots, n_n\}$ .

To calculate a *Neighbourhood Reputation* we use fuzzy rules as well. These rules, that are domain dependent, relate the outcome reputation of a target's neighbour and the social relation they have, with a reputation of the target agent.

The application of these rules generates a set of *individual neighbourhood reputations* noted as  $R_{a \rightarrow b}^{n_i}(\varphi)$ . For instance, using again our running example, one rule could be:

IF  $R_{a \rightarrow n_i}(\text{swindler})$  is X AND  $\text{coop}(b, n_i) \geq \text{low}$   
 THEN  $R_{a \rightarrow b}^{n_i}(\text{swindler})$  is X  
 IF  $RL_{a \rightarrow n_i}(\text{swindler})$  is X' AND  $\text{coop}(b, n_i)$  is Y'  
 THEN  $RL_{a \rightarrow b}^{n_i}(\text{swindler})$  is T(X', Y')

In other words, we are saying that if the neighbour of the target agent is a swindler and there is a relation of cooperation between the target and this neighbour, then the target is also (assumed to be) a swindler. Finally table 1 shows a possible set of values for function  $T$ .

X'	Y'	l	m	h
vl	vl	vl	vl	vl
l	vl	vl	l	l
m	vl	l	m	m
h	l	m	h	h
vh	m	h	vh	vh

**Table 1: Function  $T$  used in reliability rules.**

The general formulae we use to calculate a *neighbourhood reputation* and its reliability are:

$$R_{a \rightarrow b}(\varphi) = \sum_{n_i \in \mathbf{N}_b} \omega^{n_i b} \cdot R_{a \rightarrow b}^{n_i}(\varphi)$$

$$RL_{a \rightarrow b}(\varphi) = \sum_{n_i \in \mathbf{N}_b} \omega^{n_i b} \cdot RL_{a \rightarrow b}^{n_i}(\varphi)$$

$$\text{where } \omega^{n_i b} = \frac{RL_{a \rightarrow b}^{n_i}(\varphi)}{\sum_{n_j \in \mathbf{N}_b} RL_{a \rightarrow b}^{n_j}(\varphi)}$$

In this case we are using the reliability of each *individual neighbourhood reputation* to weight the contribution to the final result, both for the reputation and for the reliability.

## 6.3 System reputation

The idea behind *System reputations* is to use the common knowledge about *institutional structures* and the role that the agent is playing for that *institutional structure* as a mechanism to assign default reputations to the agents. An *institutional structure* is a social structure the members of which have one or several *observable* features that unambiguously identify them as members of that social structure. The fact that there are *observable* features to identify its members is what differentiates an institutional structure from other social structures. Examples of institutional

structures in human societies are the police, a company, a club, or a family. We assume that the role that an agent is playing and the institutional structure it belongs to is something “visible” and trustworthy for each agent.

Each time an agent performs an action we consider that it is playing a single role within the institutional structure. An agent can play the role of buyer and seller but when it is selling a product only the role of seller is relevant. Although we can think up some situations where an agent can play two or more different roles at a time, we consider that there is always a predominant role so the others can be disregarded.

In **Regret** the reputations associated to each role within an institutional structure are domain dependent and part of the initial knowledge of the agent. The value for these reputations can be different depending on which institutional structure the agent belongs to. This models the idea that groups (in our case institutional structures) influence the point of view of their members [11]. Another important point is that an institutional structure does not always associate a reputation value to each contract issue.

*System reputations* are calculated using a table for each institutional structure where the rows are the possible roles, and the columns the reputation types that cannot be subdivided in more specialized reputation types (see section 7).

Table 2 shows an example of *system reputations* for agents that belong to company B from the point of view of an agent of company A. As you notice, in this case the opinion of company A towards agents in company B is not very good.

	defaulter	to_overcharge	to_deliver_late	quality_swindler
seller	0.7	-	-	-
buyer	-	0.5	0.9	0.7

Table 2: Example of *system reputations*.

Using a similar table we would define the reliability for these reputations.

*System reputations* are noted as  $R_{a \rightarrow b}(\varphi)$  and its reliability as  $RL_{a \rightarrow b}(\varphi)$ . Hence, for example, using the table defined above, we have that  $R_{a \rightarrow b}(\text{defaulter}) = 0.7$  where  $b$  is a seller that belongs to company B. Given that this is a default value for reputation used when other information sources are missing, the reliability has to be necessarily moderately low.

## 7. ONTOLOGICAL DIMENSION

Along the individual and social dimensions, reputation is always linked to a single behavioural aspect (a contract issue). With the ontological dimension we add the possibility of combining reputations on different aspects to calculate complex reputations. To represent the *ontological dimension* we use graph structures. Figure 4 shows an *ontological dimension* for a seller in the running example. In this case, the reputation of being a swindler is related with the reputation of overcharging prices and the reputation of delivering products with less quality than specified in the contracts. For the owner of this ontological structure, the delivery date is something that is not relevant to be considered a swindler.

Hence, to calculate a given reputation taking into account the *ontological dimension*, an agent has to calculate the reputation of each of the related aspects that, in turn, can be the node of another subgraph with other associated aspects. The reputation of those nodes that are related

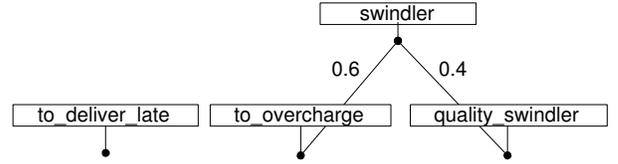


Figure 4: Ontological structure for a seller.

with an atomic aspect of the behaviour (in the example: `to_deliver_late`, `to_overcharge` and `quality_swindler`), are calculated using the individual and social dimensions. The reputation of an internal node  $\psi$  in an ontological graph is computed as follows (the computation of leaves remains as presented before):

$$R_{a \rightarrow b}(\varphi) = \sum_{\psi \in \text{children}(\varphi)} \omega_{\varphi\psi} \cdot R_{a \rightarrow b}(\psi)$$

$$RL_{a \rightarrow b}(\varphi) = \sum_{\psi \in \text{children}(\varphi)} \omega_{\varphi\psi} \cdot RL_{a \rightarrow b}(\psi)$$

For instance, using the ontological structure in figure 4 we can calculate the reputation of  $b$  as a swindler from  $a$ 's perspective using the formula:

$$R_{a \rightarrow b}(\text{swindler}) = 0.6 \cdot R_{a \rightarrow b}(\text{to\_overcharge}) + 0.4 \cdot R_{a \rightarrow b}(\text{quality\_swindler})$$

Note that the importance ( $\omega_{\varphi\psi}$ ) of each aspect is agent dependent and not necessarily static. The agent can change these values according to its mental state.

## 8. PUTTING IT ALL TOGETHER: THE REGRET SYSTEM.

The **Regret** system is an experimentation tool where the reputation of the participating agents is modeled taking into account all the aspects mentioned in this paper. In particular it defines a reputation measure (and its reliability) that takes into account the *individual dimension*, the *social dimension* and the *ontological dimension* as:

$$R_{a \rightarrow b}(\varphi) = \begin{cases} \sum_{i \in \{O, W, N, S\}} \xi_i \cdot R_{a \rightarrow b}(\varphi) & \text{if } \varphi \text{ is a leaf} \\ \sum_{\psi \in \text{children}(\varphi)} \omega_{\varphi\psi} \cdot R_{a \rightarrow b}(\psi) & \text{Otherwise} \end{cases}$$

$$RL_{a \rightarrow b}(\varphi) = \begin{cases} \sum_{i \in \{O, W, N, S\}} \xi_i \cdot RL_{a \rightarrow b}(\varphi) & \text{if } \varphi \text{ is a leaf} \\ \sum_{\psi \in \text{children}(\varphi)} \omega_{\varphi\psi} \cdot RL_{a \rightarrow b}(\psi) & \text{Otherwise} \end{cases}$$

As we have argued, the most reliable reputation is the *outcome reputation* followed by the *witness* and the *neighbourhood* reputations and finally by the *system reputation*. Therefore, we want the agent to give more relevance to the *outcome reputation* in detriment of the others. If the *outcome reputation* has a low degree of reliability (for instance because the agent does not have enough information) then the agent will try to use the *witness* and the *neighbourhood* reputations. Finally, if its knowledge of the social relationships is short, the agent will try to use the *system reputation*. Given that, the factors  $\{\xi_I, \xi_W, \xi_N, \xi_S\}$  in the general formulae are:

$$\begin{aligned}
\xi_I &= RL_{a \rightarrow b}(\varphi) \\
\xi_W &= RL_{a \rightarrow b}(\varphi) \cdot (1 - \xi_I)/2 \\
\xi_N &= RL_{a \rightarrow b}(\varphi) \cdot (1 - \xi_I)/2 \\
\xi_S &= 1 - (\xi_I + \xi_W + \xi_N)
\end{aligned}$$

## 9. RELATED WORK

The idea of using the opinion of other agents to build a reputation is not new. The work of Michael Schillo, Petra Funk and Michael Rovatsos [15] and the work of Bin Yu and Munindar P. Singh [17] are good examples of this. In both cases they use a trust-net for weighting the other agents' opinions. Our structure to calculate the witness reputation can be considered also a trust-net. In our case, however, besides the previous experiences with the witnesses we also consider the information about the agents' social relations.

The model proposed by Yu and Singh [17] merges information that comes from agents that have good reputation. Schillo et al [15] consider that the same agents that can provide you with information are also competing with you. Although agents are assumed to never lie, they can hide information or bias it to favour their goals. We go one step further and consider that the agents can also lie.

In electronic marketplaces, the reputation that a user has is the result of aggregating all the experiences of the other users that interacted with him/her in the past. Amazon Auctions [1] and eBay [2] for instance, are online auction houses where users buy and sell goods. Each time a new transaction is finished, the buyer rates the seller. These ratings are used to build the reputation of a seller. Sporas [18] is an evolved version of this kind of reputation models. Sporas introduces the notion of reliability of the reputation and is more robust to changes in the behaviour of a user than reputation systems like Amazon Auctions, based on the average of all the ratings given to the user. In all these systems each user has a global reputation shared by all the observers instead of having a reputation biased by each observer. Histos [18], also oriented to electronic commerce, is a more personalized reputation system where reputation depends on who makes the query, and how that person rated other users in the online community.

Finally, we would like to stress that unlike **Regret**, all the previous models consider reputation as a single concept instead of a multi-facet concept.

## 10. CONCLUSIONS AND FUTURE WORK

In this paper we have presented how social network analysis can be used in a reputation system that takes into account the social dimension of reputation. The system has also a hierarchical ontology structure that allows to consider several types of reputation at the same time. The combination of complementary methods that use different aspects of the interaction and social relations, allows the agent to calculate reputation values at different stages of its knowledge of the society.

The use of the social network analysis techniques as part of a reputation system opens a new field for experimentation. Our first objective is to validate the system in a realistic e-commerce environment where social relations are an important factor. To be able to exploit all the capabilities

of the **Regret** system we need environments more sophisticated than the actual e-markets like Amazon Auctions or eBay. We are working in several tools that allow the specification and implementation of these kind of e-markets.

Once you introduce the social dimension in reputation systems and the agents start to take into account social relations, it becomes more and more important to consider not only which is the reputation of the other agents, but what can an agent do to get and maintain a good reputation. Using the **Regret** system, we want to study reputations from this new perspective. Finally, it is important to study mechanisms that allow agents to build and maintain sociograms.

## 11. ACKNOWLEDGMENTS

This work has been supported by the European project SLIE, IST-1999-10948, and the Spanish MCYT project e-INSTITUTOR, MCYT 2000-1414.

## 12. REFERENCES

- [1] *Amazon Auctions*. <http://auctions.amazon.com>.
- [2] *eBay*. <http://www.eBay.com>.
- [3] A. Abdul-Rahman and S. Hales. Supporting trust in virtual communities. In *Proceedings of the Hawaii Int. Conference on System Sciences, Maui, Hawaii, 2000*.
- [4] D.B. Bromley. *Reputation, Image and Impression Management*. John Wiley & Sons, 1993.
- [5] V. Buskens. The social structure of trust. *Social Networks*, (20):265—298, 1998.
- [6] C. Castelfranchi and R. Falcone. Principles of trust for mas: Cognitive anatomy, social importance, and quantification. In *Proceedings of the 3th International Conference on Multi-Agent Systems*, pages 72—79, 1998.
- [7] M. Celentani, D. Fudenberg, D.K. Levine, and W. Psendorfer. Maintaining a reputation against a long-lived opponent. *Econometrica*, 64(3):691—704, 1996.
- [8] P. Dasgupta. Trust as a commodity. In D. Gambetta, editor, *Trust: Making and Breaking Cooperative Relations*, pages 49—72. Blackwell, 1998.
- [9] P. Hage and F. Harary. *Structural Models in Anthropology*. Cambridge University Press, 1983.
- [10] P. Hage and F. Harary. *Island Networks*. Cambridge University Press, 1996.
- [11] M. Karlins and H. Abelson. *Persuasion, how opinion and attitudes are changed*. Crosby Lockwood & Son, 1970.
- [12] R. Marimon, J.P. Nicolini, and P. Teles. Competition and reputation. In *Proceedings of the World Conference Econometric Society, Seattle, 2000*.
- [13] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.
- [14] J. Sabater and C. Sierra. Regret: A reputation model for gregarious societies. In *Proceedings of the Fourth Workshop on Deception, Fraud and Trust in Agent Societies, Montreal, Canada*, pages 61—69, 2001.
- [15] M. Schillo, P. Funk, and M. Rovatsos. Using trust for detecting deceitful agents in artificial societies. In *Applied Artificial Intelligence, Special Issue on Trust, Deception and Fraud in Agent Societies*, 2000.
- [16] J. Scott. *Social Network Analysis*. SAGE Publications, 2000.
- [17] Bin Yu and M.P. Singh. A social mechanism of reputation management in electronic communities. In *Cooperative Information Agents, CIA-2000, Boston, MA, USA*, pages 154—165, 2000.
- [18] G. Zacharia. Collaborative reputation mechanisms for online communities. Master's thesis, Massachusetts Institute of Technology, September 1999.
- [19] L.A. Zadeh. Fuzzy logic and approximate reasoning. *Synthese*, 30:407—428, 1975.