REGRET: Reputation in gregarious societies

Jordi Sabater and Carles Sierra
IIIA - Artificial Intelligence Research Institute
CSIC - Spanish Scientific Research Council
Campus UAB, 08193 Bellaterra, Catalonia, Spain
{jsabater,sierra}@iiia.csic.es

1. INTRODUCTION

In a broad sense, *reputation* is the "opinion or view of one (agent) about something". Each agent may have a different opinion about a given entity and, therefore, reputation is linked to subjectivity. This opinion is formed and updated along time through interactions. As a result of these interactions, agents record *impressions* that reflect how they value the experience. *Impressions* are the bricks that properly combined permit to value the reputation of agents.

In societies where agents belong to groups that condition their behaviour not only the direct interaction with the agent whose reputation we want to determine is important (the *individual dimension* of reputation), the interaction with other members of his group becomes very influential. Also, as in human societies, previous experiences of the members of our group with that agent or his mates may be important [1]. This is what we call the *social dimension* of reputation.

Finally, we consider that the reputation of an agent is not a single and abstract concept but rather it is a multi-facet concept. How these facets are combined to build reputations for complex concepts is what we call the *ontological dimension* of reputation.

The approach followed in this paper is similar to that used in the work of Michael Schillo et al. [3] and Bin Yu et al. [4]. Our approach radically differs from these in the sense that we take the stance that the agents' social structure is an essential factor in weighing the other agents' opinions. Another major contribution of our work is the way we model reputation as a multi-facet concept.

2. OUTCOMES AND IMPRESSIONS

We define the *outcome* of a dialogue between two agents as an initial contract to take a particular course of action or to establish the terms and conditions of a transaction, together with the actual result of the actions taken or the actual values of the terms of the transaction. We note the set of all possible outcomes as $\mathbf{O}$.

AGENTS'01, May 28-June 1, 2001, Montréal, Quebec, Canada.

We note groups of agents with upper-case letters, ($\mathcal{A}$, $\mathcal{B}$,...), and agents with indexed lower-case letters, (a_2, b_3, ...). An agent noted b_i is assumed to belong to group $\mathcal{B}$. We note by $\mathbf{A}$ the set of all agent identifiers.

We define an *impression* as the subjective evaluation made by an agent on a certain aspect of an outcome. An *impression* ι is represented by a tuple of the form:

$$\iota = (a, b, o, \varphi, t, W)$$

where $a, b \in \mathbf{A}$ are the agents who dialogue (being a who is judging), $o \in \mathbf{O}$ is the outcome, φ the variable of the outcome that is judged, t is the time when the impression is recorded, and $W \in [-1, 1]$ represents the subjective opinion of agent a with respect to φ for that particular o.

We note by I the set of all possible impressions and agent a's *impressions database* by $IDB^a \subseteq I$. We define $IDB^a_p \subseteq IDB^a$ as the set of impressions in IDB^a that satisfy the pattern p, where the general form for a pattern is:

$$\{(a, b, o, \varphi, t, W)|condition\}$$

with *condition* as a logical formula in FOL over the components of the impression. The '_' symbol is used to represent a 'don't care' value.

3. THE REGRET SYSTEM

3.1 Individual reputation

We use the term *individual reputation* to talk about the reputation value computed directly from an agent's impressions database. An individual reputation at time t from agent a's point of view and satisfying pattern p is noted as $R^t(IDB^a_p)$. To calculate this individual reputation we use a weighted mean of the impressions rating factors, giving more relevance to recent impressions:[1]

$$R^t(IDB^a_p) = \sum_{\iota_i \in IDB^a_p} \rho(t, t_i) \cdot W_i$$

where $\rho(t, t_i) = \frac{f(t_i,t)}{\sum_{\iota_j \in IDB^a_p} f(t_j,t)}$, and $f(t_i, t)$ is a time dependent function that gives higher values to values closer to t. We use the notation $R_{a \to b}(\varphi)$ to represent $R^t(IDB^a_p)$ where $p = \{(a, b, \_, \varphi, \_, \_)|true\}$ and t is the current time.

Besides the concrete reputation value, it is important to know how reliable is that value. To assess that, our model considers two elements: the number of impressions used to

[1]There are many psychological studies that support recency as a determinant factor [1].

calculate the reputation value, $Ni(IDB_p^a)$, and the variability of their rating values $Dt(IDB_p^a)$(rating deviation of the impressions). This approach is similar to that used in the Sporas system [5]. The main difference being that due to the representation of the impressions database we can define more sophisticated functions to model these two factors. We define the reliability of the individual reputation associated to an impressions database $RL(IDB_p^a)$ as a convex combination of the two elements (see [2] for the detailed definition of these two functions).

$$RL(IDB_p^a) = (1-\mu) \cdot Ni(IDB_p^a) + \mu \cdot Dt(IDB_p^a)$$

Similarly to $R_{a\to b}(\varphi)$, we use the notation $RL_{a\to b}(\varphi)$ to represent a reputation reliability.

3.2 Social reputation

In many societies, an individual inherits the reputation of the group it belongs to by default. This is so because belonging to a group gives initial expectations about the behaviour of an agent. Also, an individual tends to use the experiences of the other members of his own group as a help to shape his opinion on some matter [1]. This is what we call the *social dimension* of reputation. This dimension gives an agent the possibility to model an essential characteristic of complex societies: the group structure.

We distinguish three social reputation measures:

- The interaction with the other members of the group to which the other agent belongs to:

  $$R_{a\to\mathcal{B}}(\varphi) = \sum_{b_i\in\mathcal{B}} \omega^{ab_i} \cdot R_{a\to b_i}(\varphi)$$

  where $\sum_{b_i\in\mathcal{B}} \omega^{ab_i} = 1$.

  As in the *individual reputation* case we need to express how reliable this reputation is:

  $$RL_{a\to\mathcal{B}}(\varphi) = \sum_{b_i\in\mathcal{B}} \omega^{ab_i} \cdot RL_{a\to b_i}(\varphi)$$

- What the members of the group think about the agent being evaluated $R_{\mathcal{A}\to b}(\varphi)$.

- What the members of the group think about the oher group $R_{\mathcal{A}\to\mathcal{B}(\varphi)}$.

To calculate the last two measures and their associated reliabilities we use the same approach as for the first case.

Finally, we define the reputation measure that combines the *individual reputation* measure with the three *social reputation* measures as:

$$\begin{aligned} SR_{a\to b}(\varphi) \quad = \quad & \xi_{ab} \cdot R_{a\to b}(\varphi) + \xi_{a\mathcal{B}} \cdot R_{a\to\mathcal{B}}(\varphi) + \\ & \xi_{\mathcal{A}b} \cdot R_{\mathcal{A}\to b}(\varphi) + \xi_{\mathcal{A}\mathcal{B}} \cdot R_{\mathcal{A}\to\mathcal{B}}(\varphi) \end{aligned}$$

where $\xi_{ab} + \xi_{a\mathcal{B}} + \xi_{\mathcal{A}b} + \xi_{\mathcal{A}\mathcal{B}} = 1$. The reliability $SRL_{a\to b}$ can be calculated similarly.

3.3 Ontological dimension

In the individual and social reputation measures defined so far, reputation is always linked to a single aspect, a variable φ in the impressions. Within REGRET we permit the possibility to combine reputations on different concepts. This is done by defining an ontology represented as a cyclic graph structure. To calculate the reputation of a node in the graph, REGRET aggregates the reputation of each children that, in turn, is recurrently obtained from that of its children. The reputation of an atomic concept (a leave) is calculated using the individual and social measures presented before. The reputation of node i in an ontological graph is then computed as follows:

$$OR_{a\to b}(i) = \begin{cases} \sum_{j\in children(i)} w_{ij} \cdot OR_{a\to b}(j) & \text{if } children(i) \neq \varnothing \\ SR_{a\to b}(i) & \text{otherwise} \end{cases}$$

4. EXPERIMENTS

In order to test the model we have performed two experiments. The first one is a replica of the experiment described in [5] where a user who joins a marketplace behaves reliably until s/he reaches a high reputation value and then starts commiting fraud. With this experiment we show that REGRET can be used in simple scenarios and shows a behaviour as good as other systems like Sporas [5].

The interaction between travelers and travel agencies and the relation between travel agencies and tour operators define the scenario for the second experiment. This scenario has been designed to demonstrate how the model works in a complex environment and the influence of the social reputation measures. The results of this experiment corroborates that the use of *social reputation* leads to a significative improvement in the behaviour of an agent.

A detailed description of these experiments and their results can be found in [2].

5. ACKNOWLEDGMENTS

This work has been supported by the European project SLIE, IST-1999-10948.

6. REFERENCES

[1] M. Karlins and H. I. Abelson. *Persuasion, how opinion and attitudes are changed*. Crosby Lockwood & Son, 1970.

[2] J. Sabater and C. Sierra. Regret: A reputation model for gregarious societies. Technical report, IIIA, 2000. http://www.iiia.csic.es/Publications/Reports/2000/2000-06.pdf.gz.

[3] M. Schillo, P. Funk, and M .Rovatsos. Using trust for detecting deceitful agents in artificial societites. In *App. Art. Int., Special Issue on Trust, Deception and Fraud in Agent Societies*, 2000.

[4] Bin Yu and M. P. Singh. A social mechanism of reputation management in electronic communities. In *CIA-2000, Boston, MA, USA*, pages 154—165, 2000.

[5] Giorgios Zacharia. Collaborative reputation mechanisms for online communities. Master's thesis, MIT, September 1999.