Reputation-based Agreements: From Individual to Collective Opinions *

Roberto Centeno
CETINIA
University Rey Juan Carlos
Madrid, Spain
roberto.centeno@urjc.es

Ramón Hermoso
CETINIA
University Rey Juan Carlos
Madrid, Spain
ramon.hermoso@urjc.es

Viviane Torres da Silva
Universidade Federal
Fluminense (UFF)
Rio de Janeiro, Brazil
viviane.silva@ic.uff.br

ABSTRACT

Reputation mechanisms have been developed during last few years as valid methods to allow agents to better select partners in organisational environments. In most of works presented in the literature, reputation is summarised as a value, typically a number, that represents an opinion sent by an agent to another about a certain third party. In this work, we put forward a novel concept of *reputation-based agreement* in order to support the reputation definition, as well as, some desirable properties about it. We define a reputation service that collects opinions from agents, so creating *agreements* over *situations*. This service will also be in charge of presenting the information by using different *informative mechanisms*. On the other hand, we analyse how to enforce agents to send their opinions to the reputation service by adding *incentive mechanisms*. Finally, two different case studies are presented to exemplify our work.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Intelligent Agents*

General Terms

Theory, Design

Keywords

Agreement, Reputation, Organisation, Trust

1. INTRODUCTION

Reputation mechanisms have been proved to be successful methods to build multi-agent systems where agents' decision-making processes to select partners are crucial for the system functioning [5][6][10]. In models such as in [6][10] the authors focus on letting the agent the duty of requesting opinions, aggregating replies and inferring conclusions from the gathered information. Although reputation gathering process from the agent's point of view is an important issue, in this work we propose a complementary approach that endows organisations with a reputation service that may help agents to make decisions when their own information is scarce.

In this paper we introduce the concept of *reputation-based agreement* as the cornerstone of the reputation service in an organisational multi-agent system. An *agreement* is usually defined as a meeting of minds between two or more parties, about their relative duties and rights regarding current or future performance. Around this concept new paradigms have emerged [1][2] aimed at increasing the reliability and performance of agents in organisations by introducing in such communities these well-known human social mechanisms. With this in mind, we propose a novel approach for the meaning of reputation. From a global point of view, a *reputation-based agreement* is a meeting point on the behaviour of an agent, participating within an organisation, with regard to its reputation. Agreements are evaluated by aggregating opinions sent by participants about the behaviour of agents. We also define some properties that describe different types of agreements. Besides, information about reached agreements will be provided to agents by using the concept of informative mechanism [3].

The second part of the paper tackles the problem of how to make agents to collaborate sending their opinions to the reputation service in a pro-active manner. We will examine the concept of incentive mechanism [3] as a way of manipulating participants, in order to get more collaboration sending their opinions about different situations they have been involved in.

The paper is organised as follows: Section 2 formalises the reputation service, supported by the idea of reputation-based agreements. In Section 3 we illustrate all concepts introduced by means of a case study. Section 4 puts forward an incentive mechanism that enforces agents to collaborate with the reputation service. Section 5 elaborates a second case study using an incentive mechanism to enforce agents to collaborate sending their opinions to the reputation service. Section 6 discusses some related work and, finally, Section 7 summarises the paper and presents the future work.

2. A SERVICE BASED ON REPUTATION-BASED AGREEMENTS

As we have previously pointed out, the current work faces with the task of formalising a reputation service working on organisational multi-agent systems. We adhere the defini-

*The present work has been partially funded by the Spanish Ministry of Education and Science under project TIN2006-14630-C03-02 (FPI grants program) and TIN2009-13839-C03-02 and by the Spanish project "Agreement Technologies" (CONSOLIDER CSD2007-0022, INGENIO 2010)

Cite as: Reputation-based Agreements: From Individual to Collective Opinions, R. Centeno, R. Hermoso and V. Torres da Silva, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. XXX-XXX.

tion of organisation given in [4]. Summarising, an organisation is defined as a tuple $\langle \mathcal{A}g, \mathcal{A}, \mathcal{X}, \phi, x_0, \varphi, \{\mathcal{ON}^{om}, \mathcal{R}^{om}\}\rangle$ where $\mathcal{A}g$ represents the set of agents participating within the organisation; $\mathcal{A}$ is the set of actions agents can perform; $\mathcal{X}$ stands for the environmental states space; ϕ is a function describing how the system evolves as a result of agents actions; x_0 represents the initial state of the system; φ is the agents' capability function describing the actions agents are able to perform in a given state of the environment; $\mathcal{ON}^{om}$ is an organisational mechanism based on organisational norms; and $\mathcal{R}^{om}$ is an organisational mechanism based on roles that defines the positions agents may enact in the organisation (see [4] for more details).

2.1 How Agents Send Their Opinions

During an agent lifetime within an organisation, it is involved in several different *situations*. A situation is defined as a tuple $\langle \mathcal{A}g, \mathcal{R}, \mathcal{A}, T\rangle$, that represents an agent $\mathcal{A}g$, playing the role $\mathcal{R}$, while performing the action $\mathcal{A}$, through a time period T. As detailed in [4], different types of situations can be defined following this definition. For instance, situations in which an agent performs an action, regardless of the role it is playing – $\langle \mathcal{A}g, \_, \mathcal{A}, \mathcal{T}\rangle$ –, or situations in which an agent is playing a role during a time period, regardless the action it performs – $\langle \mathcal{A}g, \mathcal{R}, \_, \mathcal{T}\rangle$. Agents usually evaluate those situations in order to compile reliable information that allows them to predict the result of future situations. The rationale of the current work is that if agents share their knowledge about the situations they are involved in, this information might be useful when other agents have not enough information to select partners to interact with. This problem becomes hard when new participants join an organisation and they do not have not strong opinions yet.

Situations are evaluated from an agent's individual point of view. Thus, an evaluation may reflect the experience of the agent performing the evaluation – direct way – or the opinions provided by third parties about the evaluated situation – indirect way.

At any time, an agent can send its opinion about a particular situation to the reputation service. We call this information *reputation information message*:

Definition 1. *A reputation information message $\mathcal{R}^{info}_{ag_i \in \mathcal{A}g}$ is a tuple, representing an opinion sent by the agent ag_i to the reputation service containing an evaluation about a particular situation:*

$$\mathcal{R}^{info}_{ag_i} = \langle Sit, OpR\rangle,$$

where ag_i stands for the agent which sends the opinion; Sit is the situation being evaluated; and OpR represents the agent's opinion about the behaviour of its partner in the situation (typically a number). Therefore, an agent, by using this kind of messages, is somehow making public its opinions – evaluations – about different situations: agents, roles, etc.

2.2 Creating Reputation-based Agreements

In this section we intend to face the task of giving a novel approach for the meaning of reputation tackling this concept as a partial agreement about a certain situation. When the reputation service receives reputation information messages from agents, it aggregates them creating what we have called *reputation-based agreements*. That is, the aggregation of all the opinions regarding a particular situation is '*per se*' what a set of agents – as a whole – actually think about the aforesaid situation. Thus, a reputation-based agreement represents the consensus reached in the reputation opinions space sent by a set of agents about a particular situation.

Definition 2. *A reputation-based agreement π for a particular situation, is a tuple:*

$$\pi = \langle Sit, \mathcal{A}g, OpR, t\rangle$$

where:

- Sit stands for the situation about which the agreement is reached;
- $\mathcal{A}g$ is the set of agents that contributed to the agreement;
- OpR represents the opinion rating – whatever its representation is (qualitative, quantitative, etc.) – reached as a consequence of all opinions sent about Sit;
- t stands for the time when the agreement was reached.

Therefore, an agreement means a global opinion that a set of agents have on a certain situation. This agreement, as we put forward in the next section, can be used as a generalist expectation for a situation in which agents have no (or little) previous information about.

As we have claimed, a reputation-based agreement is reached as consequence of the aggregation of all opinions sent about a particular situation. Thus, the reputation service requires a function that is able to aggregate information reputation messages sent by agents. The aim of such a function is to create agreements from reputation opinions that agents send to the service by means of reputation information messages. We formally define the function as follows:

Definition 3. *Let f_π be a function that given all the reputation information messages sent by agents and a particular situation creates a reputation-based agreement for that situation:*

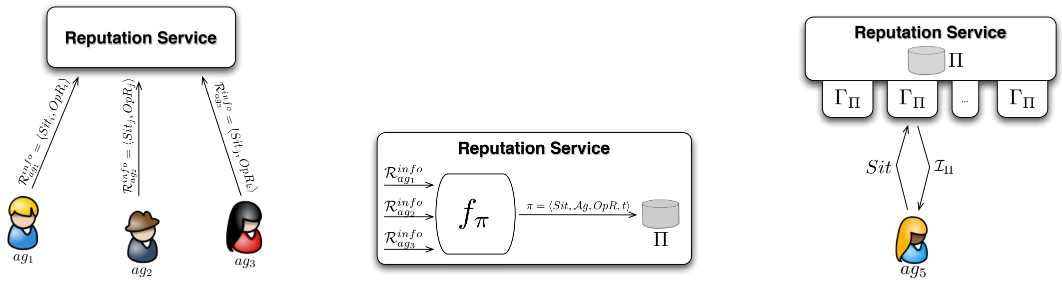
$$f_\pi : |\mathcal{R}^{info}_{ag_i \in \mathcal{A}g}| \times Sit \rightarrow \Pi$$

where:

- $|\mathcal{R}^{info}_{ag_i \in \mathcal{A}g}|$ stands for the set of reputation information messages received by the reputation service;
- Sit is the set of situations;
- Π represents the set of reputation-based agreements.

Some desirable characteristics should be taken into account when a function is designed. Therefore, we propose the following guidelines:

- an agreement about a situation should be updated when new reputation information messages are sent to the reputation service about the same situation;
- the function should take into account the temporality of the messages received, that is, two opinions should not have the same importance if they were sent in different moments;
- when an agreement about a specific situation is being created, the function should also consider opinions about a more general situation. Let us illustrate it by an example. If a reputation-based agreement is being



(a) sending opinions (b) creating agreements (c) giving information

Figure 1: Dynamics of the Reputation Service

created about the situation $\langle Harry, seller, selling-shoes, t\rangle$, the function should take into account opinions about situations such as: $\langle Harry, seller, \_, t\rangle$ or $\langle Harry, \_, \_, t\rangle$.

Following these guidelines the reputation service might use any function that is able to aggregate values. It could use a simple function to calculate the average of all opinions, or a more elaborated one that aggregates the opinions by means of complex calculation, i.e., weighting the assessment by taking into account who is giving the information.

2.3 Reputation-based Agreements: Properties

From previous definitions – definitions 2 and 3 – it is possible to define some desirable properties about reputation-based agreements. These properties should be taken into account when agreements are created and may also provide useful extra information when informing about different issues.

Property 1. *A reputation-based agreement π is **complete** iff all agents participating in an organisation, at time t, contribute to reach that agreement:*

$$\pi^* \Leftrightarrow \begin{cases} \mathcal{O} = \langle \mathcal{A}g, \mathcal{A}, \mathcal{X}, \phi, x_0, \varphi, \{\mathcal{ON}^{om}, \mathcal{R}^{om}\}\rangle \ \wedge \\ \pi = \langle Sit, \mathcal{A}g', OpR, t\rangle \ \ \wedge \\ (\mathcal{A}g = \mathcal{A}g') \end{cases}$$

That is, given a time t every participant $ag \in \mathcal{A}g$ in the organisation $\mathcal{O}$ has necessarily sent a reputation information message indicating its opinion about the situation concerning the agreement ($\mathcal{A}g = \mathcal{A}g'$). More complete agreements mean more reliable the information in the system.

Property 2. *A reputation-based agreement π is **α-consistent** iff the reputation value of π differs, at most, $1-\alpha$ from the reputation value sent by every agent that contributed to reach that agreement:*

$$\pi^\alpha \Leftrightarrow \begin{cases} \pi = \langle Sit, \mathcal{A}g, OpR, t\rangle \ \ \wedge \\ \forall ag \in \mathcal{A}g \ \ [\forall r \in Rep_{ag}^{info} [(r = \langle Sit_i, OpR_i\rangle) \wedge \\ (Sit_i = Sit) \ \ \wedge \ \ (|OpR_i - OpR| \leq 1-\alpha)]] \end{cases}$$

This property represents how agents sending their opinions about a situation agree in a certain extent. Therefore, the higher α is, the more similar the opinions are.

Property 3. *A reputation-based agreement π is **full** iff it is complete and 1-consistent:*

$$\pi^\phi \Leftrightarrow (\pi^* \ \ \wedge \ \ \pi^\alpha \ \ \wedge \ \ \alpha = 1)$$

In the case α is 1 all agents have the same opinion about a given situation. This property is very desirable when seeking reputation-based agreements, because the more agents contribute to the agreement, the stronger validity the latter gets. Thus, the likelihood of capturing what is actually happening in the organisation tends to be higher.

Property 4. *A reputation-based agreement π is **$\mathcal{R}$-consistent** iff all the agents participating in the agreement play the same role in the system:*

$$\pi_\mathcal{R} = \langle Sit, \mathcal{A}g, OpR, t\rangle \Leftrightarrow \forall ag \in \mathcal{A}g \ \ play(ag, \mathcal{R})$$

where $\mathcal{R}$ stands for the role the consistency is based on, $\mathcal{A}g$ is the set of agents that contribute to reach the agreement, and $play : \mathcal{A}g \times \mathcal{R} \rightarrow [true, false]$ is a function that returns $true$ if the agent $\mathcal{A}g$ plays the role $\mathcal{R}$.

This property is useful in cases in which a new agent, joining an organisation, wants to know what other agents – that are executing in the organisation and playing the same role – think about a given situation. For instance, someone who is thinking of *buying* something would like to know which are the opinions of those who have previously played the role *buyer*.

Property 5. *A reputation-based agreement π is **$\mathcal{R}$-complete** iff it is $\mathcal{R}$-consistent and is complete for all the agents that play the role $\mathcal{R}$ at time t:*

$$\pi^*_\mathcal{R} = \langle Sit, \mathcal{A}g', OpR, t\rangle \Leftrightarrow \\ \begin{cases} \pi_\mathcal{R} \wedge \mathcal{O} = \langle \mathcal{A}g, \mathcal{A}, \mathcal{X}, \phi, x_0, \varphi, \{\mathcal{ON}^{om}, \mathcal{R}^{om}\}\rangle \wedge \\ \forall ag \in \mathcal{A}g \ \ (play(ag, \mathcal{R}) \rightarrow ag \in \mathcal{A}g') \end{cases}$$

Property 6. *A reputation-based agreement π is **$\mathcal{R}$-full** iff it is $\mathcal{R}$-complete and is 1-consistent:*

$$\pi^\phi_\mathcal{R} \Leftrightarrow (\pi^*_\mathcal{R} \ \ \wedge \ \ \pi^\alpha \ \ \wedge \ \ \alpha = 1)$$

Although properties 1 and 3 are desirable, they are not achievable in systems that have a significant number of agents, for instance, in electronic marketplaces. However, many systems have those properties, such as closed organisational systems where the number of participants is not huge.

2.4 Providing Information about Reputation-based Agreements

Once we have defined an agreement as a distributed consensus-based expectation for a set of agents on a certain situation, we now describe how the reputation service can present the relevant information about the reached agreements to the agents participating in the organisation. Reputation-based agreements somehow capture the general thinking about a particular situation – the more α-consistent the agreement is the more reliable it is. Thus, information about the agreements reached until that moment may be very useful for agents. In particular, when agents have recently joined the organisation, they do not have any hint about situations in which they might be involved in, so if the reputation service provides information about agreements, agents may improve their utility from the very beginning.

With this in mind, we deal with the problem of how the reputation service may provide such information. To that end, we part from the notion of *informative mechanism* [3]. Those types of mechanisms are in charge of providing some kind of information to agents in order to regulate a multi-agent system. Thus, an *informative mechanism* $\Gamma : \mathcal{S}' \times \mathcal{X}' \rightarrow \mathcal{I}$ is a function that given a partial description of an internal state of an agent (S') and, taking into account the partial view that the service has of the current environmental state ($\mathcal{X}'$), provides certain information ($\mathcal{I}$). We formally define them as follows:

DEFINITION 4. *An informative mechanism providing information about reputation-based agreements is:*

$$\Gamma_{\Pi} : Sit \times \mathcal{X}' \rightarrow \mathcal{I}_{\Pi}$$

where Sit and $\mathcal{X}'$ are already defined and $\mathcal{I}_{\Pi}$ stands for the information provided by the mechanism by using the set of agreements Π reached over the situation Sit.

We have chosen a very general definition of information in order to cover all possible types of information the reputation service could offer taking into account the reputation-based agreements reached. The information provided may consist of (i) a ranking sorting the best agents for a particular situation, such as $\langle \_, \mathcal{R}, \mathcal{A}, \_ \rangle$, created from the agreements reached for that situation, (ii) a value representing the reputation value for a situation, reached as a consequence of the agreement for that situation, (iii) an information about the properties of the agreement reached for a particular situation, if it is full, complete, etc.

3. CASE STUDY: PUBS AREA

In this section, we illustrate the proposed model by means of a simple case study. The scenario we use involves five different agents: *Anna*, *John*, *Jessica*, *Albert* and *Harry* participating within an organisation. In this organisation agents can *order* and *delivery* drinks, so the action space of agents is composed of actions such as, *order-1000-drink-a*, *delivery-2000-drink-b*, where a and b represent the type of the drink agents order/deliver. That organisation is created with the aim of getting in touch pubs' owners and providers of drinks. Thus, agents join the organisation playing the roles of *pub* and *provider*, representing a pub's owner and a company provider of drinks, respectively. In our particular example, agents are playing the following roles: *Anna - pub*, *John - pub*, *Jessica - pub*, *Albert - provider* and *Harry - provider*.

In this scenario, agents representing pubs' owners are interested in collaborating by sharing information about providers. The pubs are situated in the same area and they collaborate with each other so as to foster the attraction customers to that area. That is, although they try to maximise their own benefits, one of their goals is to foster the pubs area where they are, even if that entails to exchange information about drink providers.

Therefore, after several interactions among them – performing actions of ordering and delivering different types of drinks – *Anna* decides to make public her opinion about *Albert* and *Harry* as providers. Thus, she uses the reputation information messages to send to the reputation service her opinions, as follows:

$$\mathcal{R}^{info}_{Anna} = \langle\langle Albert, provider, \_, \_\rangle, 0.2\rangle$$
$$\mathcal{R}^{info}_{Anna} = \langle\langle Harry, provider, \_, \_\rangle, 0.9\rangle$$

This information shows that *Anna* has had bad experiences while she was ordering drinks from *Albert* (0.2)[1] because *Albert* always delivers all drinks later than the agreed date. Otherwise, the second message shows that she has had good experiences with *Harry* (0.9) because *Harry*, for instance, never violates contracts and offers low prices. Similarly, *John* and *Jessica* send their opinions about *Albert* and *Harry* as providers, by using the following messages:

$$\mathcal{R}^{info}_{John} = \langle\langle Albert, provider, \_, \_\rangle, 0.2\rangle$$
$$\mathcal{R}^{info}_{John} = \langle\langle Harry, provider, \_, \_\rangle, 0.8\rangle$$
$$\mathcal{R}^{info}_{Jessica} = \langle\langle Albert, provider, \_, \_\rangle, 0.2\rangle$$

It seems that both *John* and *Jessica* agree that *Albert* is not a reliable provider. However, *Harry* is quite reliable delivering drinks, from *John*'s point of view.

When the reputation service receives this information, it is able to create reputation-based agreements by using a function that aggregates the reputation information messages. Let us suppose that it aggregates the messages by calculating the average of reputation values sent by agents over exactly the same situation[2]:

$$f_{\pi}(Sit) = \frac{\sum_{i=1}^{n} \mathcal{R}^{info}_{ag_i} = \langle Sit, OpR_i\rangle}{n}$$

Therefore, from the set of messages sent by the agents, so far, the reputation service can create two reputation-based agreements regarding to two different situations:

$$\pi_1 = \langle\langle Albert, provider, \_, \_\rangle, \{Anna, John, Jessica\}, 0.2, t\rangle$$
$$\pi_2 = \langle\langle Harry, provider, \_, \_\rangle, \{Anna, John\}, 0.85, t\rangle$$

π_1 represents that there exists an agreement within the organisation regarding to *Albert* as *provider* – regardless the action he performs – is evaluated as 0.2, and such an agreement is reached by the collaboration of *Anna*, *John* and *Jessica*, at time t. Besides, π_2 shows that there exists an agreement in which Harry is evaluated 0.85 – it is calculated as the mean of all opinions sent – as *provider* and that the agreement is reached by *Anna* and *John*, at time t.

In order to provide information about agreements the reputation service offers three different informative mechanisms:

- $\Gamma^1_{\Pi}(\langle \mathcal{A}g, \mathcal{R}, \_, \_\rangle)$ given a situation where an agent and a role are specified, it returns meta-information[3] about the reputation-based agreement reached regarding that situation;
- $\Gamma^2_{\Pi}(\langle \mathcal{A}g, \mathcal{R}, \_, \_\rangle)$ given a situation where an agent and a role are specified, it returns the reputation-based agreement reached. In particular, it returns the reputation value in the agreement of that situation;

[1] We suppose that reputation values – denoted by OpR – are in the range [0..1]

[2] It could be used whatever other function that is able to aggregate the information received from agents

[3] Meta-information means the α-consistency of the agreement, i.e., if it is full, complete, etc.

- $\Gamma^3_\Pi(\langle \_, \mathcal{R}, \_, \_\rangle)$ given a situation where a role is specified, it returns a ranking of agents playing that role, sorted by the reputation value they have as consequence of the reputation-based agreements reached until the current time t.

Let us suppose that a new pub is opened in the same area by *Alice*, so she joins the organisation playing the role *pub*. Since the pub is recently open, she needs to order drinks. Thus, she should select a provider of drinks but she does not know any provider yet. One solution could be asking to another agent about a particular provider – distributed reputation mechanism. However, this process could be very costly because she would require many queries sent to different agents to ask about different providers. Another solution is to use informative mechanisms to get information about other participants, in this case about providers. Thus, *Alice* searches for an informative mechanism that provides a ranking of "best" providers[4]. She finds Γ^3_Π that returns a ranking of agreements when it is queried based on a situation and a role. So, *Alice* performs the following query to Γ^3_Π: $\Gamma^3_\Pi(\langle \_, provider, \_, \_\rangle) \Rightarrow \{Harry, Albert\}$ and the informative mechanism returns a ranking of agents, sorted by the reputation values according to all reputation-based agreements reached at that moment, by matching the situation specified in the query with the situation of agreements. By using this information *Alice* knows that there exists an agreement within the organisation showing that *Harry* is a better provider than *Albert*. But, how good are they? To answer this question *Alice* queries the informative mechanism Γ^2_Π as follows:

$$\Gamma^2_\Pi(\langle Harry, provider, \_, \_\rangle) \Rightarrow 0.85$$
$$\Gamma^2_\Pi(\langle Albert, provider, \_, \_\rangle) \Rightarrow 0.2$$

In that moment, *Alice* is quite sure that *Harry* is much better provider than *Albert* and there exists an agreement, within the organisation, that *Harry*'s reputation delivering drinks is 0.85 and another one that *Albert* as provider is 0.2. However, *Alice* is still doubting about which provider could be the best, because she is wondering how consistent those agreements are. Thus, she queries the informative mechanism that provides meta-information about the agreement reached regarding a given situation. Therefore, she performs the following queries:

$$\Gamma^1_\Pi(\langle Harry, provider, \_, \_\rangle) \Rightarrow \pi^{0.95}$$
$$\Gamma^1_\Pi(\langle Albert, provider, \_, \_\rangle) \Rightarrow \pi^{1}$$

With this information *Alice* clears all her doubts. She knows that all opinions sent about *Albert* are coincident because the reputation-based agreement reached is 1-consistent (π^1). Besides, she knows that the opinions sent by the agents that have interacted with *Harry* are almost the same since their variability is low (0.95-consistent). With this in mind, *Alice*, finally, selects *Harry* as her provider.

In this domain, the reputation service has been an useful mechanism allowing *Alice* to select a good provider, when she did not have any previous information about providers. The mechanism worked perfectly due to agents participating in the organisation collaborated by sending their opinions to the reputation service. They were motivated to send opinions because of the own nature of the domain – pub's owners wants to create a pubs area to attract customers. Thus, pubs get benefits individually from making public their opinions about providers. But what happens when agents are not motivated by the domain to send opinions to the reputation service? Next section deals with that problem.

[4]We suppose that informative mechanisms are made available to all participants by the organisation

4. HOW TO MOTIVATE AGENTS TO SEND THEIR OPINIONS

As we have mentioned before, this section deals with the problem of motivating agents to send their reputation information messages to the system. Thus, we propose to endow the reputation system with an *incentive mechanism* [3] so as to face this task. Following the work in [3], an incentive mechanism is formalised as a function that, given a possibly partial description of an environmental state of a multiagent system, produces changes in the transition probability distribution of the system: $\Upsilon_{inc} : \mathcal{X}' \rightarrow [\mathcal{X} \times \mathcal{A}^{|\mathcal{A}g|} \times \mathcal{X} \rightarrow [0..1]]$, where $\mathcal{X}'$ stands for the partial description of an environmental state; and $\mathcal{X} \times \mathcal{A}^{|\mathcal{A}g|} \times \mathcal{X} \rightarrow [0..1]$ is the transition probability distribution of the system, that describes how the environment ($\mathcal{X}$) evolves as a result of agents' actions ($\mathcal{A}^{|\mathcal{A}g|}$) with certain probability in $[0..1]$. Hence, an incentive mechanism, producing changes in the transition probability distribution of the system, may produce changes in the consequences of agents' actions by introducing rewards and/or penalties. That is, when a mechanism is able to modify the consequences of an action, such a modification might become in a reward or a penalty for the agent. For instance, a mechanism that installs radars over a road, is an incentive mechanism, since the probability of a car – an agent – to get fined (and, thus, the probability to change to a state with less money) is higher if the car passes at prohibited velocity than without the radar. Thus, the mechanism changes the consequences of the action *passing a road at high velocity*.

Any incentive mechanism must tackle the following requirements: *i)* to choose the agents that will be affected by the mechanism; *ii)* to select the actions in which the incentive will be applied; *iii)* to find out, at least one attribute and a possible modification of this, that affects the preferences of each selected agent[5]; and finally, *iv)* to apply the modification of the parameters selected in the step *iii)* as a consequence of the selected actions in the step *ii)*, giving in such a way a reward or a penalty to these agents. Formally, an incentive mechanism is composed of $\Upsilon_{inc} = \langle \mathcal{A}g_{inc}, \bigtriangledown, \omega_{inc}\rangle$, where $\mathcal{A}g_{inc}$ is the set of agents that will be applied for the incentive mechanism, $\bigtriangledown$ stands for the set of actions in which the incentive will be applied and ω_{inc} represents the set of attributes and their selected modification to tune up the preferences of agents $\mathcal{A}g_{inc}$. Each attribute is formalised as a tuple $\langle attribute, value\rangle$.

In our particular case, we are interested in motivating all agents participating in the organisation, because the more information the system has, the more complete the reputation-based agreements will be, and consequently, the more useful the information provided will be as well. Thus, all agents within the organisation will be affected by the incentive mechanism: $\mathcal{A}g_{inc} = \mathcal{A}g$ (requirement *i)*). As we have pointed out, all agents should be motivated to send their opinions to the system. Thus, the action of *sending reputa-*

[5]It means that this attribute affects to the utility function of the agent. Because we assume that such preferences are expressed by means of an utility function.

tion information messages to the system has to be affected by the incentive mechanism (requirement *ii)*). Therefore, $\bigtriangledown = \{send(\mathcal{R}_{ag_i}^{info})\}$. In order to find out the attributes that affect agents' preferences – requirement *iii)* – there exist two different alternatives: *a)* discover the attributes by observing the performance of agents, by modifying – the mechanism – attributes randomly, what could be very costly; and *b)* introduce a new attribute in the system, becoming an attribute that influences the agents' preferences. We tend to favour the second option, introducing a new attribute to the organisation: *points*. That is, each agent is assigned with an amount of points when it joins the organisation and the incentive mechanism will modify their amont. Formally: $\omega_{inc} = \langle points_{ag_i}, value\rangle$

Furthermore, agents within an organisation must select partners to interact with, so whatever the domain of the organisation is, such agentes are interested in selecting the best partners. Thus, their utility is influenced by the selection of such partners. So, since the reputation system might provide them useful information to that end, if we associate the new attribute with the action of querying that information, the new attribute somehow becomes in an attribute that influences the agents' preferences. Therefore, the action of querying an informative mechanism has to be affected by the incentive mechanism as well. Formally: $\bigtriangledown = \{send(\mathcal{R}_{ag_i}^{info}), query(\Gamma_\Pi)\}$.

In order to complete requirement *iv)* the mechanism must decide how to modify the attribute introduced – the points each agent has – as a consequence of the actions selected to receive an incentive – to send opinions and to query information. If the incentive mechanism does not exist, agents will be interested in performing the action of querying an informative mechanism – because they might get useful information –, but they will not be interested in performing the action of sending their opinions – because they might lose utility if they make public such opinions. Hence, the mechanism has to get the opposite effect, that is, it should make more attractive the sending of opinions and less attractive the querying for information. In this way, the new attribute becomes an attribute whose modification affects the agents' preferences.

In the case of the action $send(\mathcal{R}_{ag_i}^{info})$, the mechanism should make more attractive the state in which an agent will be, when it performs that action, since they are not interested in performing it. Thus, when an agent $ag_i \in \mathcal{A}g_{inc}$ performs the action $send(\mathcal{R}_{ag_i}^{info})$ the consequence of such an action will be a modification of the value of the attribute $\omega_{inc} = \langle points_{ag_i}, value\rangle$, such that:

$$value = value + (\alpha_1 x + \alpha_2)$$

where x is the number of new reputation-based agreements that will be created with the new opinion sent and α_1, α_2 are parameters to weight the incentive ($\alpha_1 > 0$, $\alpha_2 > 0$).

We are inspired by the market and the law of demand and offer, that is, the price of a service/product is fixed based on the demand and offer this service/product has. Thus, the points an agent gets when it sends an opinion depends on how further the new opinion is. This is measured by calculating the number of reputation-based agreements it creates – parameter x in the equation. It fluctuates between 0 and 1, when an agent sends an opinion about a situation that an agreement was not reached so far, it creates as maximum one new agreement. In such a way, when more new opinions are sent, more points agents get, so consequently, agents will be motivated to send new opinions.

On the other hand, when an agent $ag_i \in \mathcal{A}g_{inc}$ performs the action $query(\Gamma_\Pi)$, the attribute $\omega_{inc} = \langle points_{ag_i}, value\rangle$ is modified as a consequence in the following way:

$$value = value - (\alpha_3 y + \alpha_4)$$

where y stands for the demand the informative mechanism being queried by the agent has. It is calculated by the number of times such a mechanism is queried; and α_3, α_4 are parameters to weight the incentive, such that $\alpha_3 > 0$ and $\alpha_4 > 0$.

Following the simile of the law of offer and demand, the more an informative mechanism is queried, the more points agents lose querying it. Therefore, it supposes that the more an informative mechanism is queried, the more useful is the information it provides. Thus, its price will be risen and consequently, agents will need more points to query. Since the only way to get points is sending opinions, agents are definitely motivated to send information. Hence, a market of points is created giving incentives to agents to share their opinions. It is important to notice that the attribute $points_{ag_i}$ cannot be negative. That is, if the consequence of performing the action $query(\Gamma_\Pi)$ was a negative value, this action will not be executed – the information will not be provided.

In order to solve a deadlock produced when agents join the organisation without any points, the incentive mechanism assigns to agents an amount of "trial" points. It should assign, at least, a minimum quantity of points to make them able to query the informative mechanism, since the available information might be useful for them. Therefore, the attributes will be initialised as follows: $\omega_{inc} = \langle points_{ag_i}, value\rangle$ where $value = \alpha_3 n + \alpha_4$ such that, α_3 and α_4 are the same as the modification of the attribute when agents query an informative mechanism; and n is the number of "free" queries the incentive mechanism assigns when agents join the organisation.

Finally, it is important to remark that those incentives should be published so as to be effective. We suppose that the organisation also publishes them together with the informative mechanisms. In Section 5 we illustrate how the reputation service, coupled with this incentive mechanism, works in domains in which agents are not motivated to share their opinions.

5. CASE STUDY: TASKS SERVERS

In this section we put all together by illustrating the dynamics of the reputation service working with the incentive mechanism introduced in the same system. This scenario involves the same five agents: *Anna*, *John*, *Jessica*, *Albert* and *Harry* participating within another organisation. In this organisation, agents can execute different tasks when another agent requests it. Thus, agents can join the organisation playing two different roles: *servers* and *customers*. The formers are able to execute the tasks that customers request them. This domain is characterised by the following aspects: *i)* when a tasks server is overload it is not able to execute more tasks; *ii)* each server has different capabilities to execute a task, i.e., the quality of the executed task may be different; and *iii)* when a task is assigned to a server and the quality of the executed task is not good enough, the task could be required to be executed again by another server. With this characteristics agents are not motivated to send their opinions, because if all agents discover the best server, the latter will unrelentingly be overload. On the other hand, if agents do not have any hint about the best servers for each task, they will select the partners – the servers – randomly, and it could imply a loss of utility, because they could need to repeat the request to a different server. Obviously, the reputation service, without the incentive mechanism, will not work in this domain, because agents will not send their opinions and, as a result, reputation-based agreements will not be created. Therefore, the organisation will be endowed with an incentive mechanism set up as follows:

$$\Upsilon_{inc} : \langle \mathcal{A}g_{inc} = \{Anna, John, Jessica, Albert, Harry\}, \\ \bigtriangledown = \{send_{ag_i}(\mathcal{R}^{info}_{ag_i}), query_{ag_i}(\Gamma_\Pi)\}, \\ \omega_{inc} = \{\langle points_{ag_i}, 130\rangle\}\rangle$$

where ag_i stands for an agent in the set $\mathcal{A}g_{inc}$; and the number 130 represents the initial points, calculated by the equation $value = \alpha_3 n + \alpha_4$ with $n = 2$ and α_3, α_4 as we will detail next. In order to modify the value of the attributes, as consequence of the execution of actions in $\bigtriangledown$, the mechanism is configured with the following parameters: $\alpha_1 = 200$ $\alpha_2 = 100$ $\alpha_3 = 10$ $\alpha_4 = 50$.

In addition, agents join the organisation playing the following roles: *Anna* - *customer*; *John* - *customer*; *Jessica* - *customer*; *Albert* - *server* and *Harry* - *server*. Within the organisation there are many agents playing the server role as well, but for the sake of simplicity we do not detail them.

After this point, agents start to interact by selecting their partners randomly, because although they have enough points to query informative mechanisms, there are not agreements available yet, since agents have not send their opinions. Since bootstrapping of the incentive mechanism is out of the scope of the paper we decided to assign no points to agents forming the initial state of the scenario. Thus, initial agents will have a value of 0 in the corresponding attribute of points. Otherwise, newcomers will have a value of 130, calculated as we have explained before.

When several interactions have been carried out, *customer* agents are aware of some useful information about *servers*. Once agents run out of points they need to send their opinions, so as to get some points in order to keep on querying. Thus, the following reputation information messages are sent to the service:

$$\mathcal{R}^{info}_{Anna} = \langle\langle Albert, server, \_, \_\rangle, 0.9\rangle \\ \mathcal{R}^{info}_{John} = \langle\langle Albert, server, \_, \_\rangle, 0.9\rangle \\ \mathcal{R}^{info}_{Jessica} = \langle\langle Albert, server, \_, \_\rangle, 0.9\rangle \\ \mathcal{R}^{info}_{John} = \langle\langle Harry, server, \_, \_\rangle, 0.9\rangle \\ \mathcal{R}^{info}_{Anna} = \langle\langle Harry, server, \_, \_\rangle, 0.2\rangle$$

As a consequence of executing those actions, the incentive mechanism modifies the values of the attribute $points_{ag_i}$ by using the equation explained in Section 4, as follows: $\langle points_{Anna}, 400\rangle$, $\langle points_{John}, 400\rangle$ and $\langle points_{Jessica}, 100\rangle$, since Anna and John contributed to create a new agreement when they sent their opinions. However, Jessica did not contribute to any new agreement.

When the reputation service receives those messages, the following reputation-based agreements are created, by using the same function – the average function – that in the case study explained in Section 3:

$$\pi_1 = \langle\langle Albert, server, \_, \_\rangle, \{Anna, John, Jessica\}, 0.9, t_1\rangle \\ \pi_2 = \langle\langle Harry, server, \_, \_\rangle, \{Anna, John\}, 0.55, t_1\rangle$$

These agreements show that all agents playing the role *customer* think the same about *Albert* as server and that Harry is evaluated with 0.55 as *server*, from the point of view of *Anna* and *John*. We suppose that the reputation service has the same informative mechanisms as in the example of Section 3: $\Gamma^1_\Pi(\langle \mathcal{A}g, \mathcal{R}, \_, \_\rangle)$, $\Gamma^2_\Pi(\langle \mathcal{A}g, \mathcal{R}, \_, \_\rangle)$ and $\Gamma^3_\Pi(\langle \_, \mathcal{R}, \_, \_\rangle)$.

At this point, our agents could query any informative mechanism because they have enough points to do it. However, we will focus on a new agent – *Alice* – that joins the organisation at this moment. Since she is a newcomer, she will be assigned 130 points – two "free" queries to an informative mechanism. Then, she performs the following queries:

$$\Gamma^3_\Pi(\langle \_, server, \_, \_\rangle) \Rightarrow \{Albert, Harry\} \\ \Gamma^2_\Pi(\langle Albert, server, \_, \_\rangle) \Rightarrow 0.9$$

After these queries *Alice* knows that the best server, according to the reputation-based agreements reached, so far, is *Albert*, with 0.9 evaluation. Then, the value of the attribute $points_{Alice}$ will be modified to 10, following the equation explained in Section 3. Since she cannot query again due to the lack of points, she will select *Albert* as server. When she performs the interaction with Albert, she wants to get more information about how the agreement about *Albert* is, because she is wondering if such an agreement is not consistent enough, she could have problems if she assigns a different task to *Albert*. Thus, she needs to send her opinion in order to get more points. After that, she gets 310 points (10 that she already has plus 300 that she gets sending an opinion about a situation that does not form part of an agreement yet). Now, she can perform a new query: $\Gamma^1_\Pi(\langle Albert, server, \_, \_\rangle) \Rightarrow \pi^*_{customer}$. With this new information, she knows that *Albert* is the best server and all customers evaluate him exactly with the same reputation ($\pi^*_{customer}$ is *customer-complete*). This information is worthy for *Alice* because she will not need to change of server.

6. RELATED WORK

In this paper we do not propose another reputation mechanism but a reputation service that generates agreements based on collected opinions about the reputations of agents in a given situation. Such service can be used by the centralised part of an hybrid reputation model [13] or by an agent participating in a distributed mechanism that is interested in aggregate the opinions it has received about its behaviour or in aggregate the opinions it has about the behaviour of another agent, for instance.

One of the main advantages of having a centralised reputation service is the feasibility for an individual to know a more consistent reputation about another agent based on numeral experiences. In the case of distributed mechanisms (such as [11][6][10]), the agent itself would need to participate in several interactions with the given agent and also to ask other agents for their experiences with others. In the case of a centralised mechanism, the agent can easily get information about the reputation showing the behaviour of other agents within the system. In [12], Sabater et al. present a centralised reputation mechanism that is incorporated as a service in Electronic Institutions (EIs). From a global perspective, this work has many similarities with ours, since uses also a reputation service in an organisational environment (EIs). However, the authors do not focus on how to exploit the collected information as agreements that can be presented to agents in different ways. In addition, none of the analysed, distributed mechanisms [11][6][10][12] detail how agents aggregate the opinions they receive.

Regarding the incentive mechanisms existing in literature to drive agents to collaborate by sending their truthfully opinions a well known work is the one by Jurca and Faltings [8]. They use a mechanism of buy/sell information using credits. The main difference is that they use it in a distributed environment, where agents send opinions among them, but not to a central repository as in our case. In [7] the authors present an approach to create rankings able not only to provide the most trustful agents but also a probabilistic evidence of such reputation values. Those rankings are also computed by a centralised system by aggregating the reputations reported by the agents. This approach and the one presented in our paper could be complementary, since that paper focuses on defining the ranking algorithms and ours focuses on describing the mechanism that allows to receive the reputation information and to provide the already evaluated agreements (for instance by using rankings). Another work that could be also complementary to the approach presented here, is the one presented in [9]. They describe the algorithm *NodeRanking* that creates rankings of reputation ratings. Therefore, our reputation service could use this algorithm so as to provide information about the reputation-based agreements reached within the organisation.

7. CONCLUSIONS AND FUTURE WORK

Summarising, this work puts forward a novel approach of reputation-based agreement concept by supporting on a reputation service that creates reputation-based agreements as aggregations of opinions sent by participants within an organisation. Besides, we also define some desirable properties that can be derived and should be taken into account when providing the information they contain. Furthermore, we also propose to use the agreements by utilising the concept of informative mechanisms [3], so providing agents with useful information. On the other hand, we propose an incentive mechanism [3] to deal with the problem of lack of collaboration from agents to send their opinions to the service. Finally, different examples have been analysed so that they illustrate how the reputation service works in two different domains: the former represents a collaborative domain where agents are interested in sharing their opinions, and the later shows a competitive domain in which the reputation service must be coupled with an incentive mechanism to motivate agents to send their opinions.

In future work we plan to experimentally test our approach by implementing a case study presented here, as well as, running several experiments comparing our approach with similar ones. We also intend to investigate new properties about reputation-based agreements to provide agents participating in an organisation with more useful information. Finally, we plan to extend the concept of reputation-based agreement by creating agreements aggregating "similar" situations, so we must go into the concept of similar situations in depth.

8. REFERENCES

[1] H. Billhardt, R. Centeno, A. Fernández, R. Hermoso, R. Ortiz, S. Ossowski, J. Pérez, and M. Vasirani. Organisational structures in next-generation distributed systems: Towards a technology of agreement. *Multiagent and Grid Systems: An International Journal*, 2009.

[2] C. Carrascosa and M. Rebollo. Modelling agreement spaces. In *IBERAMIA 2008 Workshop on Agreement Technologies (WAT 2008)*, pages 79–88. M. Esteva, A. Giret, A. Fernández, V. Julián, 2008.

[3] R. Centeno, H. Billhardt, R. Hermoso, and S. Ossowski. Organising mas: A formal model based on organisational mechanisms. In *24th Annual ACM Symposium on Applied Computing (SAC2009), Hawaii, USA, March 8-12*, pages 740–746, 2009.

[4] R. Centeno, V. T. da Silva, and R. Hermoso. A reputation model for organisational supply chain formation. In *Proc. of the 6th COIN@AAMAS'09 Budapest,Hungary*, pages 33–48, 2009.

[5] C. Dellarocas. *Handbook on Economics and Information Systems*, chapter Reputation Mechanisms. Elsevier, 2005.

[6] T. Huynh, N. Jennings, and N. Shadbolt. Fire: An integrated trust and reputation model for open multi-agent systems. In *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI)*, pages 18–22, 2004.

[7] A. Ignjatovic, N. Foo, and C. Lee. An analytic approach to reputation ranking of participants in online transactions. In *WI-IAT '08: Proceedings of the 2008 International Conference on Web Intelligence and Intelligent Agent Technology*, pages 587–590. IEEE Computer Society, 2008.

[8] R. Jurca and B. Faltings. An incentive compatible reputation mechanism. In *E-Commerce, 2003. CEC 2003. IEEE International Conference on*, pages 285–292, 2003.

[9] J. Pujol, R. Sangüesa, and J. Delgado. Extracting reputation in multi agent systems by means of social network topology. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 467–474, New York, NY, USA, 2002. ACM.

[10] J. Sabater and C. Sierra. Reputation and social network analysis in multi-agent systems. In *Proceedings of First International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 475–482, 2002.

[11] J. Sabater and C. Sierra. Review on computational trust and reputation models. *Artificial Intelligence Review*, 24(1):33–60, 2005.

[12] J. Sabater-Mir, I. Pinyol, D. Villatoro, and G. Cuní. Towards hybrid experiments on reputation mechanisms: Bdi agents and humans in electronic institutions. volume 2, pages 299–308, 2007.

[13] V. Silva, R. Hermoso, and R. Centeno. A hybrid reputation model based on the use of organization. In J. Hubner, E. Matson, O. Boissier, and V. Dignum, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems IV*, volume 5428 of *LNAI*. Springer-Verlag, 2009.