Instrumenting Multi-Agent Organisations with Reputation Artifacts

Jomi Fred Hübner*, **Laurent Vercouter, and Olivier Boissier**
ENS Mines Saint Etienne / G2I / SMA
158 Cours Fauriel
42023 Saint-Etienne Cedex, France
{hubner,boissier,vercouter}@emse.fr

Abstract

Reputation is often cited as an instrument to enforce norm compliance: agents that do not follow the norms have their reputation decreased. In the case where reputation is conceived as a collective process, a kind of *shared voices* as proposed by Conte & Paolucci, to change the reputation is not a simple task. In this paper we propose to instrument multi-agent organisation with *reputation artifacts* that publishes objective evaluations of the performance of the agents with respect to their behaviour within the organisation. The members of the organisation can then read these evaluations and build up their reputation of others. The artifact serves thus as an instrument that aid in the building of the reputation of the agents. The evaluation of the agents is not simply based on their obedience to norms, but also considers their pro-activeness and their contribution to the success of collective tasks that are being executed in the organisation. This proposal is detailed and exemplified in the context of the $\mathcal{M}$OISE$^+$ organisational model and the ORA4MAS platform.

Keywords: organisation, norm enforcement, sanctions, reputation, artifacts.

Introduction

The concept of multi-agent *organisation* is becoming widely accepted as an instrument to control the autonomy of the agents in a system that has a general purpose (Dignum and Dignum 2001; Hübner, Sichman, and Boissier 2007). For example, when someone adopts the role of master student in a laboratory, she remains autonomous to perform its research but should follow some rules of the laboratory organisation. These rules vary from 'the access to computers requires an username' to 'a master thesis should be written in two years'. The agent is free to adopt the role, but once adopted the organisation expects her autonomy to be limited.

An important feature of this approach when applied to multi-agent systems (MAS) is the flexibility: the agents are neither completely autonomous to do whatever they want nor completely constrained to pre-defined behaviours. The organisation serves as a kind of 'tuning' of the autonomy level. To find out a good degree of allowed autonomy is indeed a challenge, specially in the case where the agents have to organise the system themselves (Hübner, Sichman, and Boissier 2004).

The success of this organisational approach depends on how the compliance to the rules are ensured inside the system. In this paper we discuss the use of the agent's *reputation* as an instrument to enforce the compliance to organisational rules. The general proposal is that the agent's behaviour is constantly evaluated by the organisation with respect to the roles it plays and the result of this evaluation is published to other members (phase i). This information helps then the agents to construct the reputation of others inside the organisation (phase ii). Hence the reputation influences decision processes, agents take care of their reputation and behave accordingly (phase iii). While phases ii and iii are concerned with how the agents will use the published information, the first phase can be conceived outside the agents. The main contribution of this work is to describe how this first phase of the process can be instrumented in a multi-agent organisation.

The next section presents a general view of the norm enforcement in the context of organisations and the main concepts used in this paper. In the sequence, the third section reifies these concepts on the $\mathcal{M}$OISE$^+$ organisational model on which our proposal is based. The fourth section then details the proposal of using reputation to enforce norms, particularly by means of artifacts. We finish the paper discussing related works, specially those that consider the relation between organisation and reputation.

Regimentations and Norms in Multi-Agent Organisations

To illustrate the concepts used in the sequence of this section, we will use the following scenario:

> Alice has recently started her master's course in a French research laboratory in computer science. As a master student she has thus several rules to follow: write a technical report from state of the art in the thesis' subject; write a paper in English, code programs to experiment ideas, be friendly with colleagues, use only computers allocated to the master course, do not break equipments, etc. Alice also plans to continue her studies in a PhD course in the same laboratory. She is thus concerned about her reputation during the masters be-

*Supported by ANR Project ForTrust.

cause it is normally used in the PhD selection process.

In this scenario, we can identify several rules that limit the autonomy of Alice and that she accepted when entered in the master course. Despite her disposition to follow these rules or not, the organisation should have instruments to ensure that they are followed. Before presenting these instruments, the next subsection classifies two types of rules.

Regimentations and Norms

Among the rules in the above scenario, there is a sort of rules that can be ensured by the organisation. For example, the rule 'use only computers allocated to the master course' can be ensured by user's profiles and passwords. In this example, the login procedure to access computers is the *instrument* that implements that rule. However, rules like 'write a paper in English' do not need to (or cannot) be forced by the organisation. From the set of all rules related to a role, we distinguish thus the following two sub sets.[1]

Regimentations : are rules that the agents are forced by the system to follow, they cannot be violated. More precisely, we consider a regimentation as an action interdiction. The predicate $regimentation(\alpha, a)$ is used to mean that the organisational action a is ensured to not be executed by the agent α.[2]

Norms : are rules that the agents decide to obey, norms can therefore be violated and cannot be ensured by the organisation. More precisely, we consider norms as a goal obligation. The predicate $obligated(\alpha, \varphi)$ means that an agent α has to fulfil the goal φ.[3]

One reason for an organisation to implement a particular rule as a regimentation instead of a norm is the (serious) damage caused by the no compliance to the rule (e.g. illegal access to computers). However, for some rules, even those very important, the organisation may not find ways to implement it as a regimentation. In this case, that rule can be implemented as a norm and thus be potentially violated by the agents (e.g. break equipments). Of course, there are other reasons that motivate the design of a rule as a norm rather than as a regimentation, for instance, the cost of implementation. To state in a contract that Alice is obligated to finish her thesis in two years is cheaper than the implementation of an access control system. A more detailed discussion of these two sets of rules can be found in (Grossi, Aldewered, and Dignum 2007) and (Fornara and Colombetti 2006).

[1]This classification of rules is based on the proposal described in (Grossi, Aldewered, and Dignum 2007). However, we present them in a more specific context: regimentation is related only to actions and norms to goals.

[2]We do not define regimentation as 'ensure to perform some action' because it is not possible in open systems to force agents to perform some action. A very special set of action however (those that are under the control of the system, as 'send a message'), can be ensured to not be performed by the agents.

[3]We are aware that the concept of norm is broader and more complex than ours (e.g. (Tuomela and Bonnevier-Tuomela 1995) and the Deontic Logic in Computer Sciente workshop series (Goble and Meyer 2006)). For the present paper however this simple and informal definition is enough to discuss the proposal.

Regimentations and Norms Management

In the context of platforms for MAS organisation the concept of regimentation is often used. Agents run on an infrastructure that ensures that all regimentations will be respected, as in the case of AMELI (Esteva et al. 2004) where regimentations are ensured by means of governors, $\mathcal{S}$-$\mathcal{M}$OISE$^+$ (Hübner, Sichman, and Boissier 2006) by organisational proxies, and ORA4MAS (Kitio et al. 2008) by organisational artifacts. For example, when an agent sends a message in the context of a protocol execution, if the message does not follow the rules stated by the protocol, the message is blocked. The agent cannot violate these rules. Organisational infrastructures are thus common instruments that implement regimentation.

Some organisational models have however rules that can not be implemented as regimentations. In the $\mathcal{M}$OISE$^+$ model, for example, two roles may be related by an authority link: the agent playing the role ρ_1 has to obey orders from the agent playing the role ρ_2 (Hübner, Sichman, and Boissier 2007). It is very difficult to implement instruments to force this rule in an MAS, specially in open systems where the internal state of the agents is neither visible nor controllable by the organisation. Instruments that implement those rules as norms are required in this case.

For this second kind of instruments, normally two steps are considered: violation detection and sanction application. The detection of violation is certainly a hard task in MAS and several proposals have been presented (e.g. (Vázquez-Salceda, Aldewereld, and Dignum 2004)). However, as stressed in (Grossi, Aldewered, and Dignum 2007) detection without sanction is worthless. The problem we identified and that motivated our work is that, as far as we know, no organisational platform consider the sanction issue.

The instrument for a sanction system proposed in this paper is detailed on a particular organisational model: $\mathcal{M}$OISE$^+$. The next section thus briefly describes this model based on an example and identifies some of its regimentation and norms.

Regimentations and Norms in $\mathcal{M}$OISE$^+$

$\mathcal{M}$OISE$^+$ is an organisational modelling language that explicitly decomposes the organisation into structural, functional, and deontic dimensions (Hübner, Sichman, and Boissier 2007). The structural dimension defines the *roles*, *groups*, and *links* of the organisation. The definition of roles states that when an agent decides to play some role in a group, it is accepting some behavioural rules related to this role. The functional dimension describes how the *global collective goals* should be achieved, i.e., how these goals are decomposed (in global *plans*), grouped in coherent sets (by *missions*) to be distributed to the agents. The decomposition of global goals results in a goal-tree, called *scheme*, where the leafs-goals can by achieved individually by the agents. The deontic dimension is added in order to bind the structural dimension with the functional one by the specification of the roles' *permissions* and *obligations* for missions. Instead of being related to the agents' behaviour space (what they can do), the deontic dimension is related to the agents'

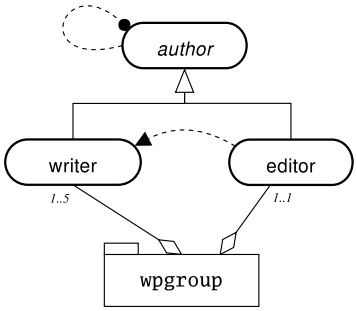

Figure 1: Structural Specification

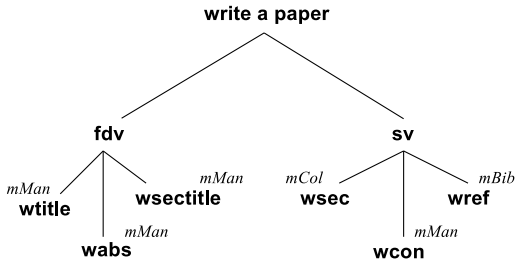

Figure 2: Functional Specification

role	deontic relation	mission	cardinality
editor	$permitted$	$mMan$	1..1
writer	$obligated$	$mCol$	1..5
writer	$obligated$	$mBib$	1..1

Figure 3: Deontic Specification

autonomy (what they should do).

As an illustrative and simple example of an organisation specified using $\mathcal{M}$OISE$^+$, we consider a set of agents that wants to write a paper and therefore has an organisational specification to help them to collaborate. The structure of this organisation has only one group (`wpgroup`) with two roles (`editor` and `writer`) that are sub-roles of the role `author`. The cardinalities and links of this group are specified, using the $\mathcal{M}$OISE$^+$ notation, in Fig. 1: the group can have from one to five writers and exactly one editor; the editor has authority on writers and every author (and by inheritance every writer and editor) has a communication link to all other authors. In this example, the `editor` and the `author` roles are not compatible, to be compatible a compatibility relation must be explicitly added in the specification.

To coordinate the achievement of the goal of writing a paper, a scheme is defined in the functional specification of the organisation (Fig. 2). In this scheme, an agent initially defines a draft version of the paper (identified by the goal *fdv* in the scheme of Fig. 2) that has the following subgoals: write a title, an abstract, and the section names. Other agents then "fill" the paper's sections to get a submission version of the paper (identified by the goal *sv*). The goals of this scheme are distributed in three missions: $mMan$ (general management of the process), $mCol$ (collaborate in the paper writing the content), and $mBib$ (get the references for the paper). A mission defines all goals an agent commits to when participating in the execution of a scheme, for example, commit to the mission $mMan$ is indeed a commitment to achieve four goals of the scheme. Goals without an assigned mission are satisfied by the achievement of its subgoals. The deontic relation from roles to missions is specified in Fig. 3. For example, any agent playing the role `editor` is permitted to commit to the mission $mMan$.

The specification of an organisation is written in a suitable language that gives the following sets, predicates and functions (only items used in this paper are listed):

- $\mathcal{G}$: the set of all groups;
- $\mathcal{R}$: the set of all roles;
- $\mathcal{S}$: the set of all schemes;
- $\mathcal{M}$: the set of all missions;
- Φ: the set of all goals;
- $compat(g, \rho, C)$: it is true that the role $\rho \in \mathcal{R}$ is compatible with all roles in $C \subseteq \mathcal{R}$ when played in the group $g \in \mathcal{G}$;
- $ms(m, s)$ it is true that the mission $m \in \mathcal{M}$ belongs to the scheme $s \in \mathcal{S}$;
- $gm(\varphi, m)$: it is true that the goal $\varphi \in \Phi$ belongs to the mission $m \in \mathcal{M}$;
- $obl(\rho, m)$: it is true that an agent playing the role $\rho \in \mathcal{R}$ is obligated by the deontic specification to commit to the mission $m \in \mathcal{M}$;
- $gr(\varphi, \rho)$: it is true that the role ρ is obligated to the goal φ, this predicate is defined as follows

$$gr(\varphi, \rho) \stackrel{\text{def}}{=} gm(\varphi, m) \ \wedge \ obl(\rho, m)$$

- $maxp : \mathcal{R} \times \mathcal{G} \rightarrow \mathbb{Z}$: a function that maps pairs of roles and groups to the maximum number of players of that role in the group.

An MAS that uses a $\mathcal{M}$OISE$^+$ specification runs on an organisational platform (as $\mathcal{S}$-$\mathcal{M}$OISE$^+$ or ORA4MAS) that provides the following runtime information:

- $\mathcal{A}$: the set of all agents inside the organisation;
- $plays(\alpha, \rho, g)$: it is true that the agent α plays the role ρ in the group g (g is an instance of a group in $\mathcal{G}$);
- $committed(\alpha, m, s)$ it is true that the agent α is committed to the mission m in the scheme s (s is an instance of a scheme in $\mathcal{S}$);
- $achieved(\varphi, \alpha)$: it is true if the goal φ is already achieved by the agent α;
- $possible(\varphi)$: it is true if the goal φ is possible. Possible goals are those that are not achieved yet and that all precondition goals are satisfied. For example, the goal 'to write the conclusion of the paper' can be achieved only after the goal of writing sections was achieved;

- $succeeded(s)$ it is true if the scheme $s \in \mathcal{S}$ finished successfully;
- $players : \mathcal{R} \times \mathcal{G} \rightarrow \mathbb{Z}$: a function that maps pairs of roles and groups to the number of current players of that role in the group.

Based on the $\mathcal{M}$OISE$^+$ specification and the platform, we can write several rules that can be applied to the agent's behaviour. However it is not the focus of this paper to present how the overall specification is translated to regimentations and norms. Four examples are thus presented to illustrate the translation.

Example 1 (regimentation). An organisational specification can declare that two roles ρ_1 and ρ_2 are compatible inside a group g ($compat(g, \rho_1, \{\rho_2\})$) and thus an agent that plays ρ_1 in g cannot perform the action $adopt\_role(\rho_i, g)$ (where $i \neq 2$). This rule is formalised in the following implication:

$$plays(\alpha, \rho, g) \ \wedge \ compat(g, \rho, C)$$
$$\Rightarrow \forall_{\rho_i \in \mathcal{R} \setminus C} \ regimentation(\alpha, adopt\_role(\rho_i, g))$$

The condition of the rule is a conjunction of predicates which evaluation is given by the specification ($compat(g, \rho, C)$) and the platform ($plays(\alpha, \rho, g)$); the conclusion is a regimentation on an organisational action ($adopt\_role(\rho_i, g)$)

Example 2 (regimentation). The number of players of a role in a group may be limited by a declaration such as $card(\rho, g, max)$, we can have thus the following regimentation based on the cardinality declaration:

$$players(\rho, g) \geq maxp(\rho, g)$$
$$\Rightarrow \forall_{\alpha \in \mathcal{A}} \forall_{\rho \in \mathcal{R}} \ regimentation(\alpha, adopt\_role(\rho, g))$$

Among the $\mathcal{M}$OISE$^+$ specification elements, the authority link, the commitment to obligated missions, and the achievement of goals should be expressed as norms.

Example 3 (norm). From the deontic relations $obl(\rho, m)$ included in the organisation specification and the roles played by the agents, we directly infer a norm:

$$plays(\alpha, \rho, g) \ \wedge \ obl(\rho, m) \ \wedge \ ms(m, s)$$
$$\Rightarrow obligated(\alpha, DONE(commit\_mission(m, s)))$$

This is a particular case of a norm where the goal is to perform an organisational action (commit to a mission in a scheme). Note that the predicate obl is static (a potential obligation), based on the specification, and $obligated$ is dynamic (a concrete obligation), based on the current state of the organisation (α is playing a role).

Example 4 (norm). Once an agent α is committed to a mission m, she is obligated to fulfil the possible goals of the mission.

$$committed(\alpha, m, s) \ \wedge \ gm(\varphi, m) \ \wedge \ possible(\varphi)$$
$$\Rightarrow obligated(\alpha, \varphi)$$

While the regimentations described in the two first examples can be easily implemented in the organisational infrastructure ($adopt\_role(\rho, g)$ is an organisation action under the control of the infrastructure), the implementation of the two latter examples is not so easy: how can we detect that some agent is not pursuing a goal without accessing its internal state; how can we enforce agents to follow their organisational obligations. The next section deal with these problems.

Reputation Artifact

The reputation is widely cited as an instrument to enforce norms (Grossi, Aldewered, and Dignum 2007; Grizard et al. 2007; Vázquez-Salceda, Aldewereld, and Dignum 2004; Muller and Vercouter 2005). However few proposals are detailed in the context of an organisational infrastructure that aims to enforce its norms. Inspired by the concept of *reputation artifact* proposed in (Conte and Paolucci 2002, p. 101), this section details such artifact in the context of the ORA4MAS organisational infrastructure (Kitio et al. 2008). This infrastructure was chosen because it is based on the concept of artifact and thus our reputation artifact are easily integrated.

ORA4MAS

ORA4MAS is based on the A&A (Agents and Artifacts) model (Ricci, Viroli, and Omicini 2007). In this model, the environment is not a merely passive source of agent perceptions and target of agent actions, but a first-class abstraction that can be suitably designed to encapsulate some fundamental functionalities and services, supporting MAS dimensions such as coordination and organisation. In particular A&A introduces a notion of *artifact* as first-class abstraction representing function-oriented dynamic entities and tools that agents can create and use to perform their individual and social activities. Thus, while agents are goal-oriented pro-active entities, artifacts are function-oriented passive entities, designed by MAS designers to encapsulate some kind of functionality, by representing (or wrapping existing) resources or instruments mediating agent activities.

Each artifact is mainly composed of two interfaces: usage and link interfaces. The *usage interface* include (1) a set of operations that agents can trigger to get artifact services and behaviours, and (2) a set of *observable properties* that the agents can inspect (observe) without necessarily executing operations on it. The execution of an operation upon an artifact can result both in changing the artifact's inner (i.e., non-observable) state, and in the generation of a stream of *observable events* that can be perceived by agents that are using or simply observing the artifact. The *link interface* provides operations to another artifact enabling composed functionalities. Agents exploit artifacts functionality (that is, they *use* artifacts) by acting on artifact usage interface which functions as a control panel, and can be aware of artifact observable state by observing observable properties.

As depicted in Fig. 4, agents are situated in an environment with artifacts that they can use for different services. In the particular case of ORA4MAS, we are emphasising the organisational artifacts that offers all organisational services required in an organisational management platform. There are three main types of artifacts in the figure: group, scheme, and reputation artifacts. The latter will be explained in the

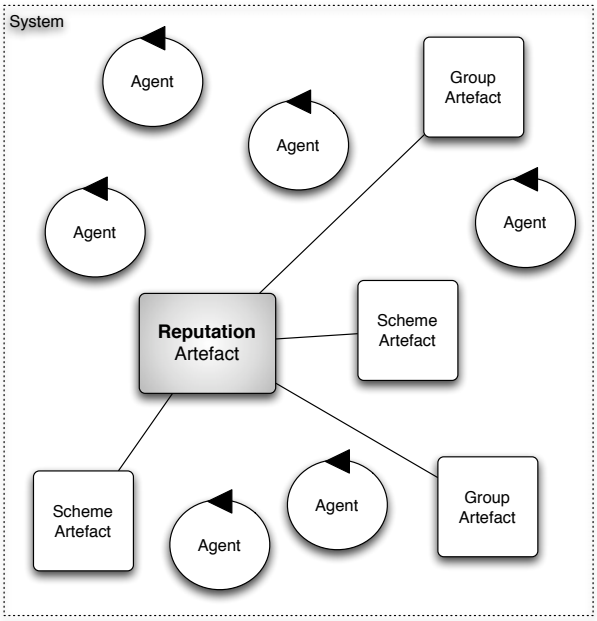

Figure 4: Agent and Artifact system

next subsection. Group artifacts maintains the state of an instance of group type and offers operations related to this group. For example, when an agent wants to adopt a role in a group, she should go to the corresponding artifact and trigger the 'role adoption' operation. Similarly, scheme artifact offers operations related to the execution of an instance of a scheme, e.g. commitment to missions. As observable properties, the group artifact shows the current players of the group (a table based on the predicate $plays(\alpha, \rho, g)$) and the scheme shows the players (a table based on predicate $committed(\alpha, m, s)$) and possible goals (based on predicate $possible(\varphi)$).

ORA4MAS artifacts are responsible to (i) ensure that all regimentations are followed by the agents and (ii) detect the violation of norms, but do not implement the violation policies that conclude to sanctions (these policies are delegated to organisational agents). All violated norms can be both displayed as observable properties of the artifact and sent to the reputation artifact. Besides a clear separation of concerns between agents and artifacts, the A&A approach simplifies the decentralisation of the infrastructure once one artifact is loosely coupled to others.

Agent's reputation

We are proposing to add a new special artifact in the system that serves as an indirect sanction instrument for norms enforcement. While direct sanction are applied when the violation is detected, indirect sanctions have long term results, as is the case of reputation.

The reputation artifact is linked to all organisational artifacts and can be observed by all agents inside the organisation. Other artifacts inform it about the current state of the organisation and then this information is used to compute an *evaluation* for each agent inside the organisation. This evaluation is published as an observable property of the artifact. It is important to notice that the evaluation is not the reputation of the agent, as remarked in (Conte and Paolucci 2002), reputation is a *shared voice* circulating in a group of agents. The reputation artifact is indeed an instrument to influence the reputation of the agent.

Several criteria may be used to evaluate an agent inside an organisation. Herein we choose to evaluate an agent in the context of the roles and missions she is engaged. Three criteria are used: obedience, pro-activeness, and result.

The *obedience* of an agent is computed by the number of obligated goals an agent achieves. The goals an agent is obligated to achieve are defined by norms (as those presented in the Examples 3 and 4). All obligated goals that were not yet achieved are considered as a violation.[4] The general mission obedience function ($o : \mathcal{A} \rightarrow [0, 1]$) and the obedience in the context of a particular mission ($o_m : \mathcal{A} \times \mathcal{M} \rightarrow [0, 1]$) and role ($o_r : \mathcal{A} \times \mathcal{R} \rightarrow [0, 1]$) are calculated as follows (in the equations $\#$ is a function that returns the size of a set):

$$o(\alpha) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ obligated(\alpha, \varphi)\}}{\#\{\varphi \mid obligated(\alpha, \varphi)\}}$$

$$o_m(\alpha, m) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ obligated(\alpha, \varphi) \ \wedge \ gm(\varphi, m)\}}{\#\{\varphi \mid obligated(\alpha, \varphi) \ \wedge \ gm(\varphi, m)\}}$$

$$o_r(\alpha, \rho) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ obligated(\alpha, \varphi) \ \wedge \ gr(\varphi, r)\}}{\#\{\varphi \mid obligated(\alpha, \varphi) \ \wedge \ gr(\varphi, r)\}}$$

$o(\alpha) = 1$ means that the agent α achieved all its obligation and $o(\alpha) = 0$ means she achieved none. $o_m(\alpha, m) = 1$ means that the agent achieved all goals when committed to the mission m, and $o_r(\alpha, \rho) = 1$ means that the agent achieved all goals when playing the role ρ.

The *pro-activeness* of an agent is computed by the number of goals an agent achieves such that she is not obligated to fulfil that goal in a scheme. The general pro-activeness function ($p : \mathcal{A} \rightarrow [0, 1]$) and the pro-activeness in the context of a particular mission ($p_m : \mathcal{A} \times \mathcal{M} \rightarrow [0, 1]$) and role ($p_r : \mathcal{A} \times \mathcal{R} \rightarrow [0, 1]$) are calculated as follows:

$$p(\alpha) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ \neg obligated(\alpha, \varphi)\}}{\#\Phi \ \#\mathcal{S}}$$

$$p_m(\alpha, m) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ \neg obligated(\alpha, \varphi) \ \wedge \ gm(\varphi, m)\}}{\#\{\varphi \mid committed(\alpha, m, \_) \ \wedge \ gm(\varphi, m)\}}$$

$$p_r(\alpha, \rho) = \frac{\#\{\varphi \mid achieved(\alpha, \varphi) \ \wedge \ \neg obligated(\alpha, \varphi) \ \wedge \ gr(\varphi, r)\}}{\#\{\varphi \mid committed(\alpha, m, \_) \ \wedge \ gm(\varphi, m) \ \wedge \ gr(\varphi, r)\}}$$

[4] We still do not consider the temporal dimension of the obligations. For instance, once an obligated goal is possible for an agent, it is violating the corresponding norm until the achievement of the goal because there is not timeout assigned to the obligation.

$p(\alpha) = 1$ means that the agent achieved all goals she is not obligated to (a highly pro-active behaviour) and $p(\alpha) = 0$ means the contrary.

The *results* of an agent is computed by the number of successful execution of scheme where she participates. It does not depend on the achievement of the goals in the scheme. It means the agent somehow share the success of the scheme execution and likely has helped for the success. The general results function ($r : \mathcal{A} \rightarrow [0, 1]$) and the results in the context of a particular mission ($r_m : \mathcal{A} \times \mathcal{M} \rightarrow [0, 1]$) and role ($r_r : \mathcal{A} \times \mathcal{R} \rightarrow [0, 1]$) are calculated as follows:

$$r(\alpha) = \frac{\#\{s \mid committed(\alpha, \_, s) \ \wedge \ succeeded(s)\}}{\#\{s \mid committed(\alpha, \_, s)\}}$$

$$r_m(\alpha, m) = \frac{\#\{s \mid committed(\alpha, m, s) \ \wedge \ succeeded(s)\}}{\#\{s \mid committed(\alpha, m, s)\}}$$

$$r_r(\alpha, \rho) = \frac{\#\{s \mid committed(\alpha, m, s) \ \wedge \ succeeded(s) \ \wedge \ obl(\rho, m)\}}{\#\{s \mid committed(\alpha, m, s) \ \wedge \ obl(\rho, m)\}}$$

$r(\alpha) = 1$ means that all schemes the agent participated have finished successfully and $r(\alpha) = 0$ means the contrary.

Unlike the previous two criteria, the results value of an agent cannot be increased by the agent itself. This evaluation depends on the performance of all agents committed to the same scheme, creating thus a dependence among them. The selection of good partners is therefore important and the reputation artifact could be used for that purpose.

The aforementioned criteria are combined into a single overall evaluation of an agent ($e : \mathcal{A} \rightarrow [0, 1]$) by the following weighted mean:

$$e(\alpha) = \frac{\gamma\, o(\alpha) + \delta\, p(\alpha) + \epsilon\, r(\alpha)}{\gamma \ + \ \delta \ + \ \epsilon}$$

$$e_m(\alpha, m) = \frac{\gamma\, o(\alpha, m) + \delta\, p(\alpha, m) + \epsilon\, r(\alpha, m)}{\gamma \ + \ \delta \ + \ \epsilon}$$

$$e_r(\alpha, \rho) = \frac{\gamma\, o(\alpha, \rho) + \delta\, p(\alpha, \rho) + \epsilon\, r(\alpha, \rho)}{\gamma \ + \ \delta \ + \ \epsilon}$$

The factors γ, δ, and ϵ are used to define the importance of the obedience, pro-activeness, and results values respectively.

All these objective values provided by the reputation artifact can then be used by agents to compute the reputation of others. It is possible that in one organisation where violation is the rule, if you are a strong violator of norms, your reputation is perhaps greater that in an organisation where violation is not at all the rule.

Example

This subsection illustrates the evaluations performed by the reputation artifact based on a small history of the organisation created to write papers and presented in the second section. Three instances of the scheme were executed as shown in Table 1, the first and third executions have finished with a paper written, but the second has failed. In the first scheme Bob has chosen Alice as a partner and in the second scheme the partner was Marc. Even though all goals were achieved in the second scheme, the overall scheme failed. One possible reason is the competence of Marc to achieve his goals. In the third scheme Bob decided to work with both Alice and Marc. The scheme finished successfully. Note however that Marc did not achieve the goal of compiling the references. This task was done by Alice, even though wref was not her goal.

In the Table 2 the evaluation of the three agents are shown. Only the obligation criteria is presented in all contexts (missions and roles), for the others the general evaluation is included in the table. The values used for γ, δ, and ϵ are respectively 1, 5, and 2. With these parameters, pro-activeness is the more important criteria resulting in Alice as having the best evaluation since she was the only one that performed not obligated goals.

Related Works

Some works that consider both the organisation and the reputation are concerned to the problem of how an agent can use the position of another agent in a organisation as an evaluation criteria. This approach is well illustrated in the example cited by (Esfandiari and Chandrasekharan 2001) where a police uniform gives some reputation to an agent wearing it because of the organisation represented by the uniform. The REGRET (Sabater and Sierra 2002) and FIRE (Huynh, Jennings, and Shadbolt 2004) reputation models also take this direction and use the organisation as yet another source of information (as direct interaction and witness) to form the reputation of a target agent. The organisation gives a kind of 'label' (as a uniform or a role) to the agents. Summing up, they have an agent centred approach and thus collective issues like norm enforcement and sanctions are not considered.

On one hand, our proposal is complementary to the approach used in the works cited above given an organisation centred view of the problem. Although we do not consider how the agents build the reputation of others, we provide an objective and detailed source of information to the agents' reputation model. The information published in the reputation artifact has two important features: (i) it is not a simple label assigned to agents ('Bob plays editor') but an evaluation of the performance of the agent in an organisational context (role or mission); and (ii) it does not depend on a subjective evaluation, but is rather precisely computed. On the other hand, we differ from the agent centred approach placing the reputation artifact *inside* the organisation. It is supposed to be used by agents of the organisation to chose partners and to improve the overall organisational performance, working as norm enforcement instrument.

Another important work in the domain is presented in (Hermoso, Billhardt, and Ossowski 2007). They also take an agent centred approach, but propose to consider the place of an agent in the organisation in different contexts. The three levels of evaluation described in our evaluation mechanism (general, role, mission) are inspired by their work.

In a recent work, (Silva, Hermoso, and Centeno 2008) proposed an approach that considers both an agent and an or-

Scheme	Agent	Role	Mission	Achieved Goals	Unachieved goals
s_1	Bob	editor	mMan	wtitle, wabs, wsectitle, wcon	
	Alice	writer	mCol	wsec	
	Alice	writer	mBib	wref	
s_2	Bob	editor	mMan	wtitle, wabs, wsectitle, wcon	
	Marc	writer	mCol	wsec	
	Marc	writer	mBib	wref	
s_3	Bob	editor	mMan	wtitle, wabs, wsectitle, wcon	
	Alice	writer	mCol	wsec, wref	
	Marc	writer	mCol	wsec	
	Marc	writer	mBib		wref

Table 1: Example of an history of the organisation

Agent	o_{mMan}	o_{mCol}	o_{mBib}	o_{editor}	o_{writer}	o	p	r	e
Bob	12/12	–	–	12/12	–	12/12	0/18	2/3	0.29
Alice	–	2/2	1/1	–	3/3	3/3	1/18	2/2	0.41
Marc	–	2/2	1/2	–	3/4	3/4	0/18	2/3	0.26

Table 2: Example of observable properties of the reputation artifact

ganisation centred approach. Agents evaluate others regarding the compliance of their behaviour vis-à-vis the norms. The evaluation and the reasons for such evaluation are then sent to the organisation. One advantage of their proposal is that the agents' evaluations are distributed, since they are performed by agents. This feature requires however that the system is also concerned of the reputation of the agent as 'evaluation sources'. As in our approach the evaluation is performed by the infrastructure, we can assume the correctness and objectiveness of the information. Another difference is that our evaluation is not based only on norm conformity, the pro-activeness of the agents is also taken into account.

Our approach also shares one property with traditional reputation systems as eBay: the centralisation and publication of the information. Although the evaluations of our proposal are published in one artifact, they are computed by several distributed artifacts (scheme and group artifacts). Another difference is that the evaluation is not performed by users but based on precise metrics with a clear meaning.

Although several authors comment that reputation can be increased or decreased as a kind of sanction, they do not tackle the problem of how to increase/decrease reputation. It is a problem specially when considering the definition of reputation as proposed by (Conte and Paolucci 2002) – reputation is something outside the agents, but known by them. In this case, to change the reputation is neither to simply change a value in a database nor to answer this value when requested (serving as a witness). The public character of the value is important, and it is achieved by our proposal of reputation artifact.

Conclusion and Perspectives

This paper presented work in progress that includes reputation as an instrument to enforce norms inside organisations. Its contribution is twofold: (i) a detailed agent evaluation process that considers the agents obedience, pro-activeness and results in three levels (general, role, mission); and (ii) the use of artifacts as instruments for an indirect sanction system. The inclusion of pro-activeness leads us to a system that is not based only on obedience, as pointed out for example by (Castelfranchi 2003), sometimes the agents should break the rules. The inclusion of results forces the agents to choose good partners in the execution of collective tasks. To choose good partners, the reputation artifact can be used, improving thus the importance and effect of this artifact. Although we have presented the concept of reputation artifact in the case of ORA4MAS and $\mathcal{M}$OISE$^+$, its application on other infrastructures is straightforward.

As future work, we intend to study "the agents' side" (phases ii and iii cited in the introduction): how the information provided by the reputation artifact can be concretely used by the reasoning mechanisms of the agents and how the reputation of the agents are formed. We also plan to implement our proposal in an agent programming language where artifacts are well integrated, as those proposed in (Ricci et al. 2008), and perform an evaluation in a real scenario.

References

Castelfranchi, C. 2003. Formalizing the informal? dynamic social order, bottom-up social control, and spontaneous normative relations. *Journal of Applied Logic* 1(1–2):47–92.

Conte, R., and Paolucci, M. 2002. *Reputation in Artificial Societies: Social Beliefs for Social Order*. Kluwer.

Dignum, V., and Dignum, F. 2001. Modelling agent societies: Co-ordination frameworks and institutions. In Brazdil, P., and Jorge, A., eds., *Proceedings of the 10th Portuguese Conference on Artificial Intelligence (EPIA'01)*, LNAI 2258, 191–204. Berlin: Springer.

Esfandiari, B., and Chandrasekharan, S. 2001. On how agents make friends: Mechanisms for trust acquisition. In *Proceedings of 4th workshop on deception, fraud and trust in agent societies*.

Esteva, M.; Rodríguez-Aguilar, J. A.; Rosell, B.; and Arcos, J. L. 2004. AMELI: An agent-based middleware for electronic institutions. In Jennings, N. R.; Sierra, C.; Sonenberg, L.; and Tambe, M., eds., *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'2004)*, 236–243. New York: ACM.

Fornara, N., and Colombetti, M. 2006. Specifying and enforcing norms in artificial institutions. In Omicini, A.; Dunin-Keplicz, B.; and Padget, J., eds., *Proceedings of the 4th European Workshop on Multi-Agent Systems (EUMAS 06)*.

Goble, L., and Meyer, J.-J. C., eds. 2006. *Proceedings of the 8th International Workshop on Deontic Logic in Computer Science, DEON 2006, Utrecht, The Netherlands, July 12-14, 2006*, volume 4048 of *Lecture Notes in Computer Science*. Springer.

Grizard, A.; Vercouter, L.; Stratulat, T.; and Muller, G. 2007. Peer-to-peer normative system to achieve social order. In Noriega, P.; Vásquez-Salceda, J.; Boella, G.; Boissier, O.; Dignum, V.; Fornara, N.; and Matson, E., eds., *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, 274–289. Springer. Revised Selected Papers.

Grossi, D.; Aldewered, H.; and Dignum, F. 2007. *Ubi Lex, Ibi Poena*: Designing norm enforcement in e-institutions. In Noriega, P.; Vásquez-Salceda, J.; Boella, G.; Boissier, O.; Dignum, V.; Fornara, N.; and Matson, E., eds., *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, 101–114. Springer. Revised Selected Papers.

Hermoso, R.; Billhardt, H.; and Ossowski, S. 2007. Integrating trust in virtual organisations. In Noriega, P.; Vásquez-Salceda, J.; Boella, G.; Boissier, O.; Dignum, V.; Fornara, N.; and Matson, E., eds., *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, 19–31. Springer. Revised Selected Papers.

Hübner, J. F.; Sichman, J. S.; and Boissier, O. 2004. Using the MOISE+ for a cooperative framework of MAS reorganisation. In Bazzan, A. L. C., and Labidi, S., eds., *Proceedings of the 17th Brazilian Symposium on Artificial Intelligence (SBIA'04)*, volume 3171 of *LNAI*, 506–515. Berlin: Springer.

Hübner, J. F.; Sichman, J. S.; and Boissier, O. 2006. S-MOISE+: A middleware for developing organised multi-agent systems. In Boissier, O.; Dignum, V.; Matson, E.; and Sichman, J. S., eds., *Proceedings of the International Workshop on Organizations in Multi-Agent Systems, from Organizations to Organization Oriented Programming in MAS (OOOP'2005)*, volume 3913 of *LNCS*, 64–78. Springer.

Hübner, J. F.; Sichman, J. S.; and Boissier, O. 2007. Developing organised multi-agent systems using the MOISE+ model: Programming issues at the system and agent levels. *International Journal of Agent-Oriented Software Engineering* 1(3/4):370–395.

Huynh, T. D.; Jennings, N. R.; and Shadbolt, N. R. 2004. FIRE: An integrated trust and reputation model for open multi-agent systems. In *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI)*.

Kitio, R.; Boissier, O.; Hübner, J. F.; and Ricci, A. 2008. Organisational artifacts and agents for open multi-agent organisations: "giving the power back to the agents". In Sichman, J.; Noriega, P.; Padget, J.; and Ossowski, S., eds., *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, volume 4870 of *LNCS*, 171–186. Springer. Revised Selected Papers.

Muller, G., and Vercouter, L. 2005. Decentralized monitoring of agent communication with a reputation model. *Trusting Agents for trusting Electronic Societies, Lecture Notes in Computer Science* (3577):pp. 144–161.

Ricci, A.; Piunti, M.; Acay, L. D.; Bordini, R. H.; Hübner, J. F.; and Dastani, M. 2008. Integrating heterogeneous agent programming platforms within artifact-based environments. In *7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008), Estoril, Portugal, May 12-16, 2008*. ACM Press.

Ricci, A.; Viroli, M.; and Omicini, A. 2007. The A&A programming model & technology for developing agent environments in MAS. In Dastani, M.; El Fallah Seghrouchni, A.; Ricci, A.; and Winikoff, M., eds., *Post-proceedings of the 5th International Workshop "Programming Multi-Agent Systems" (PROMAS 2007)*, volume 4908 of *LNAI*, 91–109. Springer.

Sabater, J., and Sierra, C. 2002. Reputation and social network analysis in multi-agent systems. In *First International Conference on Autonomous Agents and Multiagent systems (AAMAS-02)*, 475–482.

Silva, V. T. d.; Hermoso, R.; and Centeno, R. 2008. A hybrid reputation model based on the use of organization. In Hübner, J. F., and Boissier, O., eds., *Proceedings of the Fifth Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems (COIN@AAMAS 2008)*.

Tuomela, R., and Bonnevier-Tuomela, M. 1995. Norms and agreement. *European Journal of Law, Philosophy and Computer Science 5* 41–46.

Vázquez-Salceda, J.; Aldewereld, H.; and Dignum, F. 2004. Norms in multiagent systems: some implementation guidelines. In *Proceedings of the Second European Workshop on Multi-Agent Systems (EUMAS 2004)*.