Instrumenting Multi-Agent Organisations with Artifacts to Support Reputation Processes

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. Conceiving reputation as a collective process, i.e. a kind of *shared voices* as proposed by Conte & Paolucci, is not a simple task. In this paper, we propose a first step in this direction by instrumenting multi-agent organisation with an *artifact* that publishes some 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. We propose that the evaluation of the agents is not simply based on their obedience to norms, but also considers their proactiveness 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 supported by a set of organisational artifacts as proposed in the ORA4MAS approach.

Keywords: organisation, artifacts, norm enforcement, reputation.

1 Introduction

The concept of multi-agent *organisation* is becoming widely accepted as an instrument for open systems not only to help the coordination of autonomous agents but also to control their autonomy [3, 13]. 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

* Supported by ANR Project ForTrust.

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 [11].

The success of this organisational approach depends on how the compliance to the rules is ensured inside the system. An approach to deal with this issue is to use 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 (e.g. when agents have to select partners to cooperate with), 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 by using artifacts as proposed in ORA4MAS [15].

The next section presents a general analysis of the norm enforcement in the context of organisations and the main concepts used in this paper. In the sequence (Sec. 3), these concepts are reified on the $\mathcal{M}$OISE$^+$ organisational model on which our proposal is based both at the modelling language level and at the organisation management level. Our proposal (to use artifacts for supporting reputation processes in open organisations) is then detailed in Sec. 4. We finish the paper discussing related works, specially those that consider the relation between organisation and reputation.

2 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 norms 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 because it is normally used in the PhD selection process.

In this scenario, we can identify several *norms* that limit the autonomy of Alice and that she accepted when entered in the master course. Roughly a norm is an obligation, permission, or interdiction to perform some action or achieve some goal. A norm may also have a condition that states when it is active and a

deadline to be fulfilled (write a thesis in three years).[1] Despite the disposition of the agents to follow these norms or not, the organisation should have instruments to ensure that they are followed. Before presenting these instruments, the next subsection presents two types of mechanism to implements them.

2.1 Regimentation and enforcement

In the above scenario, we can distinguish a sort of norms that can be ensured by the organisation itself. For example, the norm '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 norm. However, norms like 'write a paper in English' do not need to (or cannot) be forced by the organisation. The organisation cannot force students to write a paper in the same way it can force them to access only authorised computers. We are thus considering two main mechanisms to implement norm in a MAS:[2]

- *Regimentation* is a mechanism that simply prevents the agents to perform actions that are forbidden by a norm. More precisely, we regiment some actions in order to preserve important features of the system (e.g. the access to the computers).
- *Enforcement* is a mechanism which is applied after the detection of the violation of some norm. While regimentation is a preventive mechanism, enforcement is a reactive one. From the point of view of the agents, they may decide to obey or not the norm according to their local view of the organisation. From a system point of view, the fulfilled/unfulfilled of the norms should be detected, evaluated as a violation or not, and then judged as worth of sanction/reward or not.

These two mechanisms allow us to balance (i) the ensuring of very important properties of the system by means of regimentation and, by means of enforcement, (ii) the agents' autonomy required to keep the possibility to adapt and evolve. The norms of the MAS can be instrumented either as regimentations or enforcement mechanisms depending on which side the designer wants to give more weight. Briefly, regimentation should be used to fully constrain the actions of the agents and enforcement should be used when some violation is allowed (or even desired).

[1] We are aware that the concept of norm is broader and more complex than that used in this paper (e.g. [21] and the Deontic Logic in Computer Science workshop series [7]). For the present paper however this simple and informal definition is enough to discuss the proposal.

[2] This classification is based on the proposal described in [9, 6]. However, we present them in a more specific context: regimentation is applied only to the interdiction of organisational actions and enforcement is applied to the other cases.

2.2 Norm management

In the context of platforms for MAS organisation the regimentation mechanism is often used. Agents run on an infrastructure that ensures that all norms will be respected, as in the case of AMELI [5] where norms are ensured by means of governors, $\mathcal{S}$-$\mathcal{M}$OISE$^+$ [12] by organisational proxies, and ORA4MAS [15] 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 not indeed sent. The action of the agent is not executed since it cannot violate the norm entitled by the protocol. Organisational infrastructures normally use regimentation as an instrument to implement its norms.

Some organisational models have however norms that cannot be implemented by regimentation. 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 [13]. It is very difficult to have instruments that regiment this norm in a MAS, specially in open systems where the internal state of the agents is neither visible nor controllable by the organisation. Enforcement mechanisms are thus required to implement the norms in cases like that.

The enforcement mechanism normally considers two main steps: violation detection and sanction application. The detection of violation is certainly a hard task in MAS and several proposals have been presented (e.g. [22]). However, as stressed in [9] 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.

In this paper, we propose to instrument the organisation with an artifact that could help in the first phase of a sanction system based on reputation, as described in the introduction: evaluation of the behaviours of agents within an organisation. The proposed artifact 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 norms. The section also describes how they are managed within the ORA4MAS approach.

3 The $\mathcal{M}$oise$^+$ Organisational Model and its Artifacts

The $\mathcal{M}$OISE$^+$ model proposes an organisational modelling language that explicitly decomposes the specification of organisation into structural, functional, and deontic dimensions [13]. The structural dimension specifies 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 constraints related to this role. The functional dimension specifies 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 leaves-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.

As an illustrative and simple example of an organisation specified using $\mathcal{M}$OISE$^+$, we consider agents that aim at writing a paper and therefore have 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 inherit all properties defined for the role `author`. The cardinalities and links of this group are specified, using the $\mathcal{M}$OISE$^+$ notation, in Fig. 1(a): the group `wpgroup` can have from one to five agents playing `writer` and exactly one playing `editor`; the `editor` has authority over `writer` and every agent playing `author` (and by inheritance everyone playing `writer` or `editor`) has the possibility to communicate with every agent playing `author` (communication link from `author` to `author`). 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.

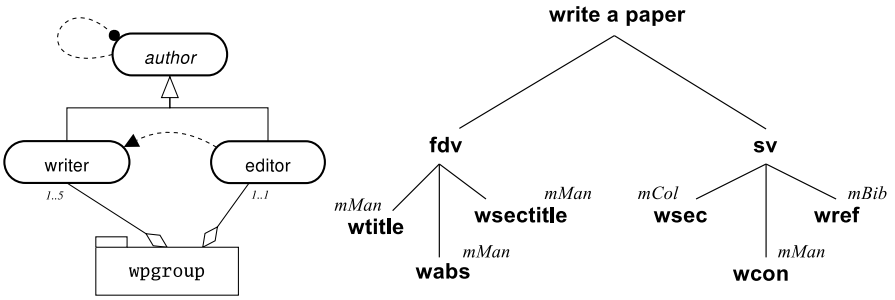
To coordinate the achievement of the goal of writing a paper, a scheme is defined in the functional specification of the organisation (Fig. 1(b)). In this scheme, a draft version of the paper has to be initially defined (identified by the goal *fdv* in Fig. 1(b)). This goal is decomposed into three sub-goals: write a title, an abstract, and the section titles. 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 which have specific cardinalities (cf. Fig. 1(c)): $mMan$ for the general management of the process (one and only one agent can commit to it), $mCol$ for the collaboration in writing the paper's content (from one to five agents can commit to it), and $mBib$ for getting the references for the paper (one and only one agent can commit to it). 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 their subgoals. The deontic relation from roles to missions is specified in Fig. 1. 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 the agents are supposed to interpret. This language is founded on components represented by predicates and functions. We present here only those components that are used in the sequel of the paper. Considering an organisational specification, $\mathcal{G}$ is the set of all group specifications, $\mathcal{R}$ is the set of all roles, $\mathcal{S}$ is the set of all scheme specifications, $\mathcal{M}$ is the set of all missions, and Φ is the set of all goals.

- $compat(g, \rho, C)$: is a predicate that is true when the role ρ ($\rho \in \mathcal{R}$) is compatible with all roles in the set C ($C \subseteq \mathcal{R}$) when played in the group g ($g \in \mathcal{G}$);
- $mission\_scheme(m, s)$ is a predicate that is true when the mission m ($m \in \mathcal{M}$) belongs to the scheme s ($s \in \mathcal{S}$);

– $goal\_mission(\varphi, m)$: is a predicate that is true when the goal φ ($\varphi \in \Phi$) belongs to the mission m ($m \in \mathcal{M}$);
– $obl(\rho, m)$: is a predicate that is true when the role ρ has an obligation relation to the mission m;
– $goal\_role(\varphi, \rho)$: is a predicate that is true when the role ρ is obliged to the goal φ, this predicate is defined as follows

$$goal\_role(\varphi, \rho) \leftrightarrow goal\_mission(\varphi, m) \ \wedge \ obl(\rho, m)$$



(a) Structural Specification (b) Functional Specification

mission	cardinality
$mMan$	1..1
$mCol$	1..5
$mBib$	1..1

(c) Missions

role	deontic relation	mission
editor	*permitted*	$mMan$
writer	*obliged*	$mCol$
writer	*obliged*	$mBib$

(d) Deontic Specification

Fig. 1. Graphical representation of the organisational specification for the writing paper example with the $\mathcal{M}$OISE$^+$ OML

3.1 ORA4MAS: Managing Organisation with Organisational Artifacts

The $\mathcal{M}$OISE$^+$ model is implemented, on one hand, by an organisational modelling language to program declarative organisation specifications, and, on the other hand, by organisational artifacts, as those proposed in ORA4MAS approach [15], that interpret the specification and manage the organisation. The conception of the artifacts follows the A&A (Agents and Artifacts) model [18]. 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.

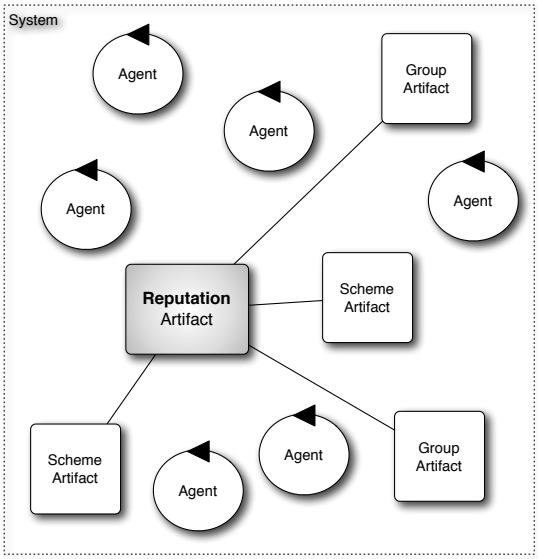


Fig. 2. Agents and Organisational Artifacts.

As depicted in Fig. 2, 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 offer all organisational services required in an organisational management platform. There are three main types of artifacts in the figure: group, scheme, and artifacts for the reputation processes. The latter will be explained in the next section. *Group artifacts* maintains the state of an instance of group type and offer 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 `adoptRole` operation. Similarly, a *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 and the scheme shows the players and possible goals.

More precisely, from the observable properties of all organisational artifacts, we can define the following sets, predicates, and functions:

- $\mathcal{A}$: 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 that the goal φ is already achieved by the agent α;
- $possible(\varphi)$: it is true that the goal φ is possible. Possible goals are those that are not achieved yet and that all pre-condition 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 that the scheme s has finished successfully.

Besides providing operations and information to the agents, ORA4MAS artifacts are also responsible to (i) ensure that all regimented norms are followed by the agents and (ii) detect the violation of norms. However they 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. As well as 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.

3.2 $\mathcal{M}$oise$^+$ norms

Based on the $\mathcal{M}$OISE$^+$ specification and a platform like ORA4MAS that provides runtime information of the current state of the organisation, we can write several norms to constrain the agent's behaviour. However it is not the focus of this paper to present how the overall organisational specification is translated to norms. Two examples are thus presented to illustrate the use of norms in a $\mathcal{M}$OISE$^+$ based organisation. In the following these norms are represented as a pair where the first argument is the condition part stating when the norm is active and the second argument is the action part stating an obligation, permission, or interdiction.

Example 1: roles are incompatible unless explicitly stated the contrary in the specification. Thus, if it is stated that two roles ρ_1 and ρ_2 are compatible inside a group g ($compat(g, \rho_1, \{\rho_2\})$), it implies that an agent that plays ρ_1 in the group g cannot perform the operation `adoptRole(`ρ_i, g`)` for any $i \neq 2$. This constraint on role adoption is formalised by the following norm:

$$\begin{array}{l}(plays(\alpha, \rho, gb) \ \wedge \ compat(g, \rho, C), \\ \forall_{\rho_i \in \mathcal{R} \setminus C} \ forbiden(\alpha, \texttt{adoptRole}(\rho_i, g)))\end{array} \quad (1)$$

The condition of the norm (the first line) is a conjunction of predicates. Its evaluation is given by the particular circumstance of the group (that defines whether $plays(\alpha, \rho, gb)$ holds or not) and the structural specification being used (that defines whether $compat(g, \rho, C)$ holds or not). The action part of the norm (the last line) states that it is forbidden for agent α to execute the action `adoptRole` on any role that does not belong to the set of compatible roles C. Based on this norm, as soon as an agent adopts a role (activating the norm), the adoption of other roles that are not explicitly stated compatible are forbidden for it.

Once this two norm is of the type 'action interdiction', they can be easily implemented by regimentation: whenever the `adoptRole` operation is requested by the agent α, if the condition of the norm holds, the execution of the corresponding operation is denied.

Example 2: once an agent α is committed to a mission m, it is obliged to fulfil the possible goals of the mission. The norm below specifies that rule.

$$(committed(\alpha, m, s) \ \wedge \ goal\_mission(\varphi, m) \ \wedge \ possible(\varphi, s), \\ obliged(\alpha, \varphi)) \tag{2}$$

While the first norm can be easily implemented in the organisational artifacts (`adopt_role`(ρ, g) is an organisation action under the control of the artifact), the implementation of this latter example 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.

4 Instrumenting Reputation Processes with Artifacts

The reputation is widely cited as an instrument to enforce norms [9, 8, 22, 16]. 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 [2, p. 101], this section details such artifact in the context of the ORA4MAS approach. It provides first class constructs which can be easily used to enrich the support of reputation processes.

4.1 Agent's reputation

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

This very artifact is linked to all organisational artifacts of the ORA4MAS and can be observed by all agents inside the organisation. Other artifacts notify 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 [2], reputation is a *shared voice* circulating in a group of agents. This 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 obliged goals an agent achieves. The goals an agent is obliged to achieve are defined by norms (as that presented in the Example 2). All obliged goals that were not yet achieved are considered as a violation.[3] 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):

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

$$o_m(\alpha,m) = \frac{\#\{\varphi \mid obliged(\alpha,\varphi) \ \wedge \ goal\_mission(\varphi,m) \ \wedge \ achieved(\alpha,\varphi)\}}{\#\{\varphi \mid obliged(\alpha,\varphi) \ \wedge \ goal\_mission(\varphi,m)\}}$$

$$o_r(\alpha,\rho) = \frac{\#\{\varphi \mid obliged(\alpha,\varphi) \ \wedge \ goal\_role(\varphi,\rho) \ \wedge \ achieved(\alpha,\varphi)\}}{\#\{\varphi \mid obliged(\alpha,\varphi) \ \wedge \ goal\_role(\varphi,\rho)\}}$$

$o(\alpha) = 1$ 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 obliged 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:

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

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

$$p_r(\alpha,\rho) = \frac{\#\{\varphi \mid achieved(\alpha,\varphi) \ \wedge \ \neg obliged(\alpha,\varphi) \ \wedge \ goal\_role(\varphi,r)\}}{\#\{\varphi \mid committed(\alpha,m,\_) \ \wedge \ goal\_mission(\varphi,m) \ \wedge \ goal\_role(\varphi,r)\}}$$

[3] We still do not consider the temporal dimension of the obligations. For instance, once an obliged 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 obliged 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.

4.2 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 obliged goals.

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

5 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 [4] where a police uniform gives some reputation to an agent wearing it because of the organisation represented by the uniform. The REGRET [19] and FIRE [14] 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 an 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 [10]. They also take an agent centred approach and 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, Da Silva et al [20] proposed an approach that considers both an agent and an organisation 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 'evaluators'. 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 [2] – 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.

6 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 [1], 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 [17], and perform an evaluation in a real scenario.

References

1. C. Castelfranchi. Formalizing the informal? dynamic social order, bottom-up social control, and spontaneous normative relations. *Journal of Applied Logic*, 1(1–2):47–92, 2003.
2. R. Conte and M. Paolucci. *Reputation in Artificial Societies: Social Beliefs for Social Order*. Kluwer, 2002.
3. V. Dignum and F. Dignum. Modeling agent societies: Co-ordination frameworks and institutions. In P. Brazdil and A. Jorge, editors, *Proceedings of the 10th Portuguese Conference on Artificial Intelligence (EPIA'01)*, LNAI 2258, pages 191–204, Berlin, 2001. Springer.
4. B. Esfandiari and S. Chandrasekharan. On how agents make friends: Mechanisms for trus acquisition. In *Proceedings of 4th workshop on deception, fraud and trus in agent societies*, 2001.

5. M. Esteva, J. A. Rodríguez-Aguilar, B. Rosell, and J. L. Arcos. AMELI: An agent-based middleware for electronic institutions. In N. R. Jennings, C. Sierra, L. Sonenberg, and M. Tambe, editors, *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'2004)*, pages 236–243, New York, 2004. ACM.
6. N. Fornara and M. Colombetti. Specifying and enforcing norms in artificial institutions. In A. Omicini, B. Dunin-Keplicz, and J. Padget, editors, *Proceedings of the 4th European Workshop on Multi-Agent Systems (EUMAS 06)*, 2006.
7. L. Goble and J.-J. C. Meyer, editors. *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, 2006.
8. A. Grizard, L. Vercouter, T. Stratulat, and G. Muller. Peer-to-peer normative system to achieve social order. In P. Noriega, J. Vázquez-Salceda, G. Boella, O. Boissier, V. Dignum, N. Fornara, and E. Matson, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, pages 274–289. Springer, 2007. Revised Selected Papers.
9. D. Grossi, H. Aldewered, and F. Dignum. *Ubi Lex, Ibi Poena*: Designing norm enforcement in e-institutions. In P. Noriega, J. Vázquez-Salceda, G. Boella, O. Boissier, V. Dignum, N. Fornara, and E. Matson, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, pages 101–114. Springer, 2007. Revised Selected Papers.
10. R. Hermoso, H. Billhardt, and S. Ossowski. Integrating trust in virtual organisations. In P. Noriega, J. Vázquez-Salceda, G. Boella, O. Boissier, V. Dignum, N. Fornara, and E. Matson, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems II*, volume 4386 of *LNAI*, pages 19–31. Springer, 2007. Revised Selected Papers.
11. J. F. Hübner, J. S. Sichman, and O. Boissier. Using the MOISE+ for a cooperative framework of MAS reorganisation. In A. L. C. Bazzan and S. Labidi, editors, *Proceedings of the 17th Brazilian Symposium on Artificial Intelligence (SBIA'04)*, volume 3171 of *LNAI*, pages 506–515, Berlin, 2004. Springer.
12. J. F. Hübner, J. S. Sichman, and O. Boissier. S-MOISE+: A middleware for developing organised multi-agent systems. In O. Boissier, V. Dignum, E. Matson, and J. S. Sichman, editors, *Coordination, Organizations, Institutions, and Norms in Multi-Agent Systems*, volume 3913 of *LNCS*, pages 64–78. Springer, 2006.
13. J. F. Hübner, J. S. Sichman, and O. Boissier. 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, 2007.
14. T. D. Huynh, N. R. Jennings, and N. R. Shadbolt. FIRE: An integrated trust and reputation model for open multi-agent systems. In *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI)*, 2004.
15. R. Kitio, O. Boissier, J. F. Hübner, and A. Ricci. Organisational artifacts and agents for open multi-agent organisations: "giving the power back to the agents". In J. Sichman, P. Noriega, J. Padget, and S. Ossowski, editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems III*, volume 4870 of *LNCS*, pages 171–186. Springer, 2008. Revised Selected Papers.
16. G. Muller and L. Vercouter. Decentralized monitoring of agent communication with a reputation model. In R. Falcone, S. Barber, J. Sabater, and M. Singh, editors, *Trusting Agents for Trusting Electronic Societies*, volume 3577 of *LNCS*, pages 144–161. Springer, 2005.

17. A. Ricci, M. Piunti, L. D. Acay, R. H. Bordini, J. F. Hübner, and M. Dastani. Integrating heterogeneous agent programming platforms within artifact-based environments. In L. Padgham, D. C. Parkes, J. Müller, and S. Parsons, editors, *7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2008), Estoril, Portugal, May 12-16, 2008*, pages 225–232. IFAAMAS, 2008.
18. A. Ricci, M. Viroli, and A. Omicini. The A&A programming model & technology for developing agent environments in MAS. In M. Dastani, A. E. Fallah-Seghrouchni, A. Ricci, and M. Winikoff, editors, *Programming Multi-Agent Systems, 5th International Workshop, ProMAS 2007, Honolulu, HI, USA, May 15, 2007, Revised and Invited Papers*, volume 4908 of *LNCS*, pages 89–106. Springer, 2008.
19. J. Sabater and C. Sierra. Reputation and social network analysis in multi-agent systems. In *First International Conference on Autonomous Agents and Multiagent systems (AAMAS-02)*, pages 475–482, 2002.
20. V. T. d. Silva, R. Hermoso, and R. Centeno. A hybrid reputation model based on the use of organization. In this volume.
21. R. Tuomela and M. Bonnevier-Tuomela. Norms and agreement. *European Journal of Law, Philosophy and Computer Science 5*, pages 41–46, 1995.
22. J. Vázquez-Salceda, H. Aldewereld, and F. Dignum. Norms in multiagent systems: some implementation guidelines. In *Proceedings of the Second European Workshop on Multi-Agent Systems (EUMAS 2004)*, 2004. http://people.cs.uu.nl/dignum/papers/eumas04.PDF.