Governing Multi-Agent Systems

Viviane Torres da Silva[1], Fernanda Duran[2], José Guedes[2] & Carlos J. P. de Lucena[2]

[1]Departamento Sistemas Informáticos y Computación, UCM
C/ J. G. Santesmases, s/n, Madrid, 28040 – Spain
viviane@fdi.ucm.es

[2]Computer Science Department, PUC-Rio
R. M de S. Vicente, 225, Rio de Janeiro/RJ, 22453-900 – Brazil
{fduran,jguedes,lucena}@inf.puc-rio.br

Abstract

In order to cope with the heterogeneity, autonomy and diversity of interests among the different agents in open multi-agent systems, several governance mechanisms have been defined. Governance mechanism enforce the behavior of agents by establishing a set of norms that describe actions that agents are prohibited, permitted or obligated to do. In this paper we present a governance mechanism that enforces not only dialogical actions but also non-dialogical ones. Although several governance mechanisms have been proposed, none of them satisfactorily deals with non-dialogical actions. Our proposed mechanism is based on testimonies provided by agents about the behavior of other agents. The governance mechanism provides decisions pointing out if norms have been violated or if false testimonies have been supplied. The decisions are based not only on testimonies and depositions provided by the agents but also on the agents' reputations supplied by a reputation system that is part of the mechanism.

Keywords: Open multi-agent system, governance, reputation, norm

1. INTRODUCTION

Open multi-agent systems are societies in which autonomous, heterogeneous and independently designed entities can work towards similar or different ends [13]. In order to cope with the heterogeneity, autonomy and diversity of interests among the different members, governance (or law enforcement) systems have been defined. Governance systems enforce the behavior of agents by establishing a set of norms that describe actions that agents are prohibited, permitted or obligated to do [4][19]. Such systems assume that norms can be violated by agents and that their internal state is neither observable nor controllable.

In this paper we propose a governance mechanism based on testimonies provided by witnesses about facts or events that they know are related to norm violations. Agents are inserted in an environment where they can perceive the changes occurred in it. Since agents are able to observe these changes, they can provide testimonies about actions or messages that are in violation of a norm. The main advantages between our approach and the ones proposed in the literature such as [5][14][15][10][20] are: (i) Our proposed mechanism does not influence the agents' privacy since it does not interfere in the interaction between agents. Some of the analyzed mechanism [14][15][10] intercepts the messages sent from an agent to another if such messages are violating a norm. In our point of view, the monitoring and the interception of the messages are violating the agents' privacy; (ii) By using our approach, it is possible to govern not only dialogical-actions but

also non-dialogical ones. Non-dialogical actions are not related to the interactions between agents but to tasks executed by agents that characterize, for instance, the access to resources, their commitment to play roles or their movement in environments and organizations. The majority of other approaches only concern about the compliance of messages with the system norms. In [5] the authors propose an access control mechanism to handle the access to resources. However, such governance is restricted and only applied to resources that are inserted in *tuple* centre environments; and (iii) Other approaches such as [20] claim that the governance system enforces only the observable behavior of agents in terms of public messages and visible actions. In our approach, private messages and also private actions can be enforced. Private messages that violate norms can be testified by agents that are involved in the interactions. Such agents can testify about messages they should have received or about messages they should have not received. Private actions that are executed in the scope of a group and are violating norms can be testified by any member of the group that knows such norms and has seen the actions being executed or has perceived facts or events that reflect the execution of such actions. The same can be said about actions that should have been executed but were not. Related facts of events cannot be observed and, therefore, agents can testify stating that the actions (probably) were not executed. In addition, private actions that are executed in the scope of one single agent and that are violating norms can be testified by any agent that knows the norms and that perceives facts or events that are related to the execution of such actions. The same can be said about actions that an agent should have execute but have not. Other agents that know the norms that regulate such actions can testify if they cannot observe the related facts or events.

Since our proposed mechanism is based on testimonies provided by autonomous, heterogeneous and independently designed agents that can lie, it is necessary to verify the trustfulness of such testimonies before using them to blame the agents being accursed of violating norms. Therefore, our governance mechanism bases the judgment of the testimonies on the reputation of the involved agents. The governance mechanism is composed of three subsystems: (i) the judgment subsystem is responsible for receiving the testimonies and for providing a decision (or verdict) pointing out to the reputation and sanction subsystems if an agent has really violated a norm; (ii) the reputation subsystem evaluates the reputation of agents according to the decisions provided by the judgment subsystem about violated norms and false testimonies.

This system also provides such updated reputations to the judgment system or to any application agent whenever it is requested; (iii) the sanction subsystem applies the sanctions specified in norms to witness agents or to defendant agents, according to the judgment decision. In this paper we focus on the judgment and the reputation systems.

The reputation model implemented by the reputation system combines the characteristics of centralized and decentralized approaches such as [8][11][17][21]. In our approach, as well as in FIRE[11] and Regret[17], agents are able to evaluate the past behavior of other agents and store the reputation of each agent with whom they have interacted with, what characterizes a decentralized approach. In addition, our approach also provides organizations with the ability to evaluate and store the reputations of agents. We assume that large-scale multi-agent systems are composed of (a hierarchy of) organizations where agents are playing roles. Each system organization should implement the proposed governance system and, therefore, its three subsystems, characterizing a (semi-)centralized approach.

Since our reputation model puts together centralized and decentralized approaches, several problems found when separately analyzing such approaches could be solved: (i) agents do not need to meet frequently in order to have consistence reputations of other agents, what occurs in some decentralized approaches. They can consult the organizations that store reputations evaluated based on several interactions with the agents; (ii) to find out someone that can provide the reputation of an agent is not time expensive since the organization can provide such information. In large-scale systems, to look for someone that can provide the reputation of an agent may be time expensive when executing in a decentralized approaches that do not offer any other way to know the reputation of an agent; (iii) the reputations provided by organizations are not overestimated since they are reliable systems that do not make distinction between the agents. In FIRE, agents receive certified reputations from those they have interacted with and can provide such reputations to agents with whom they have never interacted. The agents can overestimate their reputation since they can provide only the highest certified reputations; and (iv) the reputations provided by organizations are not biased on others' opinion. Centralized approaches simply put together agents' point of view. In our approach, the reputations are evaluated according to the characteristics of the violated norms or false testimonies.

The paper presents in Section 2 an overall view of the testimony-based governance mechanism. Section 3 details the judgment process used by the mechanism while Section 4 describes the reputation subsystem. In Section 5, a case study where we apply our approach is illustrated. Section 6 presents some related work and, finally, section 7 concludes our work.

2. THE GOVERNANCE MECHANISM

The governance mechanism presented here is based on testimonies provided by agents attesting facts or events that may be norm violations. Since every agent knows sets of norms, it can report to the governance mechanism their violation. Agents can, for instance, witness about the breaking of interaction protocols or disallowed resource accesses.

2.1. GOVERNANCE MECHANISM ASSUMPTIONS

The testimony-based governance mechanism is funded in the following assumptions.

Assumption I: *Every agent should know every norm applied to itself.*

Such as in the real world where everyone should know a code of behavior, we assume that every agent should know all norms that can be applied to their messages or actions independently of the system environment in which it is executing. When an agent enters in the environment to play a role, the environment /system must be able to provide to the agent all the norms applied to that specific role. This is important because the mechanism assumes that an agent acting in violation of a norm chooses to do so being aware of that. If the agent is not able to understand the norms, it should not be part of the system.

Assumption II: *Every agent should know every norm that influences its behavior and should be able to observe violations of such norms.*

Agents should know the norms that regulate the behavior of other agents when the violations of such norms influence their own execution. When an agent violates a norm other agents are (usually) affected by such violation. Therefore, when entering in an environment, the environment should also inform to the agents about the norms applied to other agents that may influence their behavior in order to testify to the governance system about their violations. The possible violation of such norms motivates the agents to be aware of them.

Assumption III: *Every agent can give testimonies about norm violations.*

Since an agent knows norms that are applied to other agents, the agent is able to state that one of these norms is being violated. Every time an agent perceives the violation of a norm, it must be able to give a testimony to the governance mechanism. The mechanism provided a component that can be used by agents to help them analyzing their beliefs in order to find out well-known facts or events that may be norms violations.

Assumption IV: *Some violations might be ignored / not observed.*

The proposed mechanism does not impose that an agent must give its testimony whenever it notices a norm violation. Agents should be well motivated in order to provide their testimonies. Besides, the mechanism does not guarantee that all violations will be observed by at least one agent. It may be the case that a violation occurs and no agent testifies about it. The actions that should or should not be executed and messages that should or should not be sent are specified in the norms known by the agents that are able to testify about their violations. In order to minimize violations of norms not observable, the application must carefully define and associate the norms with the agents. For each norm there must be at least two agents related to it: one of them is the agent that should behave according to the norm and the other one is the witness that can provide testimonies about the violations. It guarantees that for every norm there is an agent that is able to testify about it but it does not guarantee that the agent will do so.

Assumption V: *Agents can give false testimony.*

In an open system, agents are independently implemented, i.e. the development is done without a centralized control and the governance mechanism cannot assume that an agent was properly designed. Therefore, there is no way to guarantee that all testimonies are related to actual violations. So, the governance mechanism should be able to check and assert the truthfulness of the testimonies.

Assumption VI: *The mechanism can have a law-enforcement agent force.*

The mechanism can introduce agents which have the sole purpose of giving testimonies. The testimonies of these agents can always be considered to be truthful and the judgment subsystem can directly state that a norm was violated and a penalty should be assigned. Note that those agents must only testify if they are sure about the culpability of the application agents. They must be aware that an agent may violate a norm due some major force or to another agent fault, for instance.

2.2 THE GOVERNANCE MECHANISM ARCHITECTURE

The governance mechanism architecture defines three subsystems, as illustrated in Figure 1. The judgment subsystem is responsible for receiving the testimonies and for providing a decision (or verdict) pointing out to the reputation and sanction subsystems if an agent has really violated a norm. The system may use

different strategies to judge the violation of the different norms specified by the application. Such strategies might use the agents' reputation afforded by the reputation system to help providing the decision. It is well established that trust and reputation are important in open systems and can be used as a form which agents can reason about the reliability of other agents [16]. In [16] trust is defined as subjective probability with which agents assess that other agents will perform a particular action. We adapt this definition to our approach stating that reputation is defined as a subjective probability with which agents assess that other agent will provide trustful testimonies. The reputation subsystem evaluates the reputation of agents according to the decisions provided by the judgment subsystem about violated norms and false testimonies. This system also provides such updated reputations to the judgment system or to any application agent whenever it is requested. Finally, the third subsystem, the sanction subsystem, applies the sanctions specified in norms to the witness agents or to the defendant agents, according to the judgment decision.
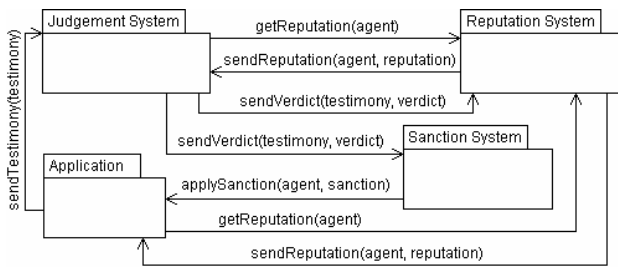

Figure 1: The architecture of the governance mechanism.

The governance mechanism was implemented by using the ASF (Agent Society Framework) framework [18]. Such framework provides support for the implementation of agents, organizations and roles. Each one of the three governance subsystems was implemented as a separated organization that interacts with a fourth organization where the application agents are situated.

3. THE JUDGMENT SUBSYSTEM

The judgment subsystem has three main responsibilities: to receive testimonies, to judge them and to provide the decision about the violation. Three different agent types were defined to deal with these responsibilities: inspector, judge and broker agents. The inspector agents are responsible for receiving the testimonies and sending them to judge agents. The judge agents examine the testimonies and provide decisions that are sent to broker agents. Broker agents are responsible for interacting with the reputation and

sanction subsystems to make the decisions effective. While judging the testimonies, judge agents may interact with brokers to get information about the reputation of agents.

3.1. THE JUDGMENT PROCESS

The judgment process is composed of seven steps where five are application independent ones. Although judgment strategies cannot be completely independent of the application norms, it is possible to define some common steps to be followed by any judgment strategy. In this section we present the seven steps that compose the judgment process.

Step I: *To verify who the witness is*

According to assumption VI, the testimony provided by some specific agents must be considered always truth. Therefore, the first step of the judgment process verifies who the witness is. If it is the case of an always truthful witness, the judgment process is finished and the verdict stating that the agent must be penalized is provided.

Step II: *To check if the norm applies to the defendant agent*

According to assumption V, agents can lie and end up accusing other agents of violating norms that are not applied to them. In order to find out if a testimony is true, the first step is to check if the norm applies to the defendant agent, i.e., if the norm is one of the norms that must be fulfilled by the agent. If the norm does not apply, the judgment process is finished and the verdict states that the defendant agent is absolved.

Step III: *To ask the defendant agent if it is guilty*

If the norm applies to the agent, the next step is to ask it if it has violated the norm it is accused of. As it happens in the real world, if the agent confesses, the judgment process is finished and the verdict states that the defendant agent is condemned. Otherwise, the judgment process continues. In cases where the defendant confesses the violation, the applied punishment is smaller than the one that would be applied if it does not confess. It intends to stimulate the agents to confess the violation.

Step IV: *To judge the testimony according to the norm (application dependent step)*

If the agent did not confess, it is necessary to carefully exam if the agent really violated the norm. In order to determine if the testimony is truth and, therefore, if the defendant agent is guilty, it may be necessary to use different strategies for different violated norm. For instance, to enforce a norm that regulates the interaction between agents or another one that regulates the access to a resource is completely different. On one hand, if the norm regulates the payment of an item and

the defendant is being accused of having not paid the witness, one possible strategy is to ask the defendant if it has the receipt signed by the witness asserting that it has received the payment. On the other hand, if the norm regulates that an agent should have not updated a resource, the judgment system could use the simple strategy that checks the resource log, in case it is provided. It is clear that such strategies are application dependent ones since they depend on the norm that is being enforced.

Step V: *To ask other agents about their depositions (application dependent step)*

If the application strategy could not decide if the defendant agent is guilty or not, the judgment system can still try another approach. Since there may be other agents that can also testify about the violation of the norm or facts related to it, the judgment system can explicitly ask them about their opinion about the violation. This step is an application dependent step because depending on the kind of question the judgment system makes to the agents, it may be necessary to interpret the answer according to the application norm being checked. For instance, two different kinds of questions can be asked to those agents: (i) Have you seen an agent violating norm n_j? (ii) What do you know about fact f_k? On one hand, the answer to the first question is a *boolean* that strictly indicates if the agent is guilt or not in the witness point of view. On the other hand, the answer to the second question must be interpreted in order to the judgment system understand the witness point of view. And such interpretation is application dependent.

Step VI: *To come up with a consensus considering the depositions*

After interpreting the depositions, the judgment system must put them together to come up with a verdict. In order to do so, our approach uses the agent reputations to help evaluating the depositions. The consensus between the depositions is provided by using subjective logic [12], as detailed in Section 3.2.2. Such an approach evaluates the depositions considering the reputations of the agents to come up with the probability of the defendant agent being guilt of violating the norm.

Step VII: *To provide the decision*

The judgment system can provide tree different decisions. It can state that (i) the defendant agent is probable guilt, (ii) the defendant is not probable guilt (the witness has lied), or (iii) the culpability of the defendant is undefined. In this case, the judge could not decide if the agent is guilt or not. Figure 2 represents the judgment process.

After producing the decision, it is necessary to send it to the reputation subsystem so that it can modify the

reputation of the accused agent, in case the judgment system has decided that the defendant agent is guilty, or the reputation of the witness, in case the judgment system has decided that it has lied. It is also important to inform the decision to the sanction subsystem to (i) punish the agent for violating a norm and to award the witness for providing the testimony or (ii) to punish the witness for providing an untruthful testimony.
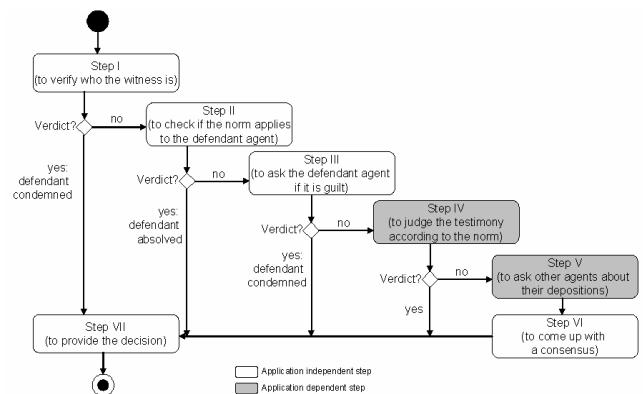

Figure 2: The state machine of the judgment process.

3.2. EVALUATING THE TESTIMONIES AND DEPOSITIONS

When there are not enough evidences to be used by the judge agent to come up with a decision, it can still make use of agents' depositions to finally provide a verdict, as described in Step V and VI. However, as stated before in assumption V, agents can give false testimonies and also false depositions. Therefore, there is a need for an approach that evaluates such testimonies and depositions considering the reliability of the agents, i.e., considering their reputations. We propose the use of subjective logic to provide a verdict stating the probability of an agent being guilt or not of violating a norm. Such an approach is used in the application independent Step VI to ponder the testimonies/depositions according to the agents' reputations and to make a consensus between them.

In [7] the authors sketched a model for e-marketplaces based on subjective logic for setting contracts back on course whenever their fulfillment deviate from what were established. Evidences from various sources are weighed in order to inform the actions that are probably violating the contracts. Subjective logic is used to support reasoning over those evidences, which involve levels of trust over parties, combining recommendations and forming consensus.

3.2.1. INTRODUCING SUBJECTIVE LOGIC.

Subjective Logic was proposed by Audun Jøsang based on the Dempster-Shafer theory of evidence [12]. This approach addresses the problem of forming a measurable belief about the truth or falsity on an atomic

proposition, in the presence of uncertainty. It translates our imperfect knowledge about reality into degrees of belief or disbelief as well as uncertainty which fills the void in the absence of both belief and disbelief [10]. This approach is described as a logic which operates on subjective beliefs and uses the term opinion to denote the representation of a subjective belief. The elements that compose the frame of discernment which is a set of all possible situations are described as follows: (i) The agent's opinion is represented by a triple $w(x) = <b(x), d(x), u(x)>$; (ii) $b(x)$ measures belief, represented as a subjective probability of proposition x of being true; (iii) $d(x)$ measures disbelief, represented as a subjective probability of proposition x of being false; (iv) $u(x)$ measures uncertainty, represented as a subjective probability that a proposition x of being either true or false; (v) $b(x)$, $d(x)$, $u(x) \in [0..1]$ and $b(x) + d(x) + u(x) = 1$; and (vi) $w^A(x)$ represents the opinion that an agent A has about the proposition x be true or false.

Subjective Logic operates on opinions about binary propositions, i.e. opinions about propositions that are assumed to be either true or false. The operators described above are to be applied over such opinions.

Recommendation (Discounting): The *discounting* operator $\tilde{\ }$ combines agent A's opinion about agent B's advice with agent B's opinion about a proposition x expressed as an advice from agent B to agent A. That means if agent B gives an advice x to agent A, and agent A has an opinion about agent B, the operator $\tilde{\ }$ can be used to form *agent A's opinion about agent B's advice x*:

(i) $w^A(B) = <b^A(B), d^A(B), u^A(B)>$ represents agent A's opinion about agent B;

(ii) $w^B(x) = <b^B(x), d^B(x), u^B(x)>$ represents agent B's opinion about x;

(iii) $w^{A:B}(x) = w^A(B) \ \tilde{\ } \ w^B(x)$ represents agent A's opinion about agent B's opinion about the preposition x. $w^{A:B}(x) = <b^{A:B}(x), d^{A:B}(x), u^{A:B}(x)>$ and is evaluated as follows:

• $b^{A:B}(x) = b^A(B) \, b^B(x)$;
• $d^{A:B}(x) = b^A(B) \, d^B(x)$;
• $u^{A:B}(x) = d^A(B) + u^A(B) + b^A(B) \, u^B(x)$.

Consensus: The *consensus* of two possibly conflicting opinions is an opinion that reflects both opinions in a fair and equal way, i.e. when two observers have beliefs about the truth of x, the consensus operator $\bar{\ }$ produces a consensus beliefs that *combines the two separate beliefs into one*:

(i) $w^A(x) = <b^A(x), d^A(x), u^A(x)>$ represents agent A's opinion about x;

(ii) $w^B(x) = <b^B(x), d^B(x), u^B(x)>$ represents agent B's opinion about x;

(iii) $k = u^A(x) + u^B(x) - u^A(x)u^B(x)$;

(iv) $w^{A,B}(x) = w^A(B) \ \bar{\ } \ w^B(x)$ represents the consensus between agent A's opinion about x and agent B's opinion about x. $w^{A,B}(x) = <b^{A,B}(x), d^{A,B}(x), u^{A,B}(x)>$ and is calculated as follows for $k \neq 0$:

• $b^{A,B}(x) = (b^A(x)u^B(x) + b^B(x)u^A(x))/k$;
• $d^{A,B}(x) = (d^A(x)u^B(x) + d^B(x)u^A(x)) / k$;
• $u^{A,B}(x) = (u^A(x)u^B(x))/k$.

3.2.2. APPLYING SUBJECTIVE LOGIC IN OUR APPROACH.

Our goal is to come up with a consensus between the different testimonies and depositions about the violation of a norm considering the reliability of the witnesses. In order to do so, it is important to understand what a testimony/deposition is in the context of subjective logic. The testimony or deposition given by agent A attesting something about a proposition x can be seen as the A's opinion about x, i.e., $w^A(x)$.

Second, it is necessary to state that the testimonies (or the opinions of the agents about facts) will be evaluated by the judge agent according to its own opinion about the agents, for instance, $w^J(a)$ where A is one of the witnesses. Such an opinion is directly influenced by the reputation of the agent.

After evaluating the judge's opinions about the agents that have given their testimonies and depositions, it is necessary to evaluate the judge's opinions about testimonies and depositions given by those agents. In order to do so the *discounting operation* will be used. Finally, after having the judge's opinions about all testimonies and depositions, it is necessary to put them all together to form the judge point of view about the violated norm. The *consensus operator* is therefore used.

Judge's opinions about the agents: The reputation provided by the reputation system reflects how much the judge believes in the agent, i.e. $b^J(a)$, and not its whole opinion about such agent, i.e $w^J(a)$.

Judge's opinions about testimonies and depositions given by the agents: Now that we already have evaluated the belief of the judge agent in a testimony A, it is necessary to determine the judge's opinion about a testimony/deposition x given by an agent, i.e $w^{J:A}(x)$. The discounting operator presented in Section 3.2.1 is, thus, used as described in equation (1):

$w^{J:A}(x) = w^J(a) \ \tilde{\ } \ w^A(x) = < b^{J:A}(x), d^{J:A}(x), u^{J:A}(x)> \ (1)$
• $b^{J:A}(x) = b^J(a)b^A(x)$;
• $d^{J:A}(x) = b^J(a)d^A(x)$;
• $u^{J:A}(x) = d^J(a) + u^J(a) + b^J(a)u^A(x)$;

- $d^J(a)+u^J(a) = 1 - rep(a)$ since $b^J(a)+d^J(a)+u^J(a) = 1$ and $b^J(a) = rep(a)$.

Judge point of view about the violated norm: Given that there may exist more than one agent testifying about the same fact (proposition x), all testimonies and depositions can be combined using the consensus operator to produce the judge's own opinion about the proposition x. The consensus puts together all testimonies and depositions while considering the reputation of the witnesses. For instance, let's suppose that A, B and C are agents that provided their testimonies and depositions, the consensus is formed by using equation (2):

$$w^{J:(A,B,C)}(x) = (w^J(a) \sim w^A(x)) \oplus (w^J(b) \sim w^B(x)) \oplus (w^J(c) \sim w^C(x)) \qquad (2)$$

4. THE REPUTATION SUBSYSTEM

As stated before in Section 2.2, the proposed reputation subsystem is part of the organization infrastructure. Its goal is to evaluate the reputation of the agents based on violated norms. Not only the reputation of the defendant agent being accused of violating a norm can be update but also the reputation of the witness agent that is providing the testimony can be modified.

4.1. EVALUATING DEFENDANTS' REPUTATION

The reputation subsystem evaluates the agents reputations based on the verdicts provided by the judgment subsystem. The judgment subsystem informs the reputation subsystems about the verdicts and the testimonies by stating the witnesses, the defendants and the norms. In case defendants are condemned by the judgment subsystem, their reputations are updated according to the norms that they have violated. The more important is the norm, the more influence it will exert in the agent reputation. Each norm must stipulate how the reputation of the agent should be modified in case the agent violated it. This information is called the *power of the norm*. The *power of a norm* can vary from 0, for norms that do not influence the agent reputation, to 1, for norms that strongly influence the agent reputation in case it is violated.

Since we assume that the judgment subsystem deals with uncertainty, the reputation subsystem may also consider it when evaluating the reputation of the agents. The reputations of two agents considered guilt for violating the same norm cannot be evaluated in the same way if the judgment subsystem is surer of the guiltiness of an agent than another. The same norm cannot influence the reputation of two agents in the same way when one were considered 90% guilt and the other 51% guilt for violating the same norm. Therefore, the reputation subsystem applies the *percentage of blame* informed by the judgment subsystem to the *power of the norms*. On one hand, when the judgment subsystem is quit sure that the agent is guilt its reputation is strongly influenced by the *power of the violated norm*. On the other hand, when the judgment subsystem is not so sure about the violation of the norm the agent's reputation is softly influenced by the *power of the norm*. Expression (3) evaluates the influence of the violated norm n_i on the reputation of agent a_j by considering the *power of the norm* and the *percentage of blame*.

$$defRepInf(a_j,n_i) = normPower(n_i) * blamePercentage(a_j, n_i) \qquad (3)$$

The influence of a violated norm on an agent reputation may change during the agent lifecycle. Frequently, norms recently violated influence more the reputation of an agent than the norms violated longtime ago. In order to overcome such issue, we propose to consider the time during while a norm will influence the agent reputation. This information during while the norm will influence the agent reputation must be part of the norm specification. Such information is used to estimate the remaining days during while the violated norm will influence the reputation. Thus, recently violated norms will strongly influence the reputations of agents, and norms violated longtime ago will weakly influence the reputations or will not influence the reputation at all, in case the time has expired. Expression (4) evaluates the influence of norm n_i on the reputation of agent a_j by considering the *power of the norm*, the *percentage of blame* and the *number of days remaining*. Note that the agent's reputation will automatically increase during the passing days since the *remainingDays* attribute decreases.

$$defRepInf(a_j,n_i) = normPower(n_i) * blamePercentage(a_j, n_i) * remainingDays(a_j,n_i) \qquad (4)$$
$$where\ remainingDays(a_j,n_i) = [\ totalTime(n_i) - passedDays(a_j,n_i)]\ /\ totalTime\ (n_i)$$

Although a violated norm may no more be influencing the reputation of an agent, the information about its violation can still be stored by the reputation subsystem. This is important while considering relapses. The influence of a norm on the reputation of an agent may increase in case of relapses. The relapse factor varies from 1 (representing no relapse at all) to a value near zero (representing many relapses) according to the importance of the norm for the system. Note that the result value must not overflow the maximum value of the *norm power* that is 1.

$normPower^r(n_i) = normPower(n_i) * (1/relapse(n_i))$ but $0 <= normPower^r(n_i) <= 1$ *(5)*

The influence of a violated norm on the reputation of an agent may decrease in case of *confession*. If the agent confesses, the *power of the norm* in the reputation of the agent decreases. Equation (6) modifies the power of the norm by considering *confession*. This factor may vary according to the importance of the norm. The more important is a norm, the minor influence this factor will imply in the agents' reputations.

$normPower^{rc}(n_i) = normPower^r(n_i) * confession(n_i)$ *(6)*

To evaluate the reputation of a defendant agent it is necessary to consider all the norms that the agent has violated, where each one is evaluated according to equation (7). Its reputation is evaluated by putting together all the partial influences, as stated in equation (8). It exemplifies the reputation of a defendant agent a_j by considering that it has violated k norms. Note that it may be the case that the defendant reputation is equal to zero if the sum of its partial influences is equal to (or grater than) one. It may also be the case that the defendant reputation is equal to 1 if its reputation is no more being influenced by the violations. Thus, a reputation may vary from 1 to 0 and we consider reputations greater than 0.5 good reputations and lower than (or equal to) 0.5 bad reputations.

$defRepInf(a_{j,} \quad n_i) = normPower^{rc}(n_i) *$
$blamePercentage(a_j,n_i) * remainingDays(a_j,n_i)$ *(7)*

$defendantRep(a_j) = 1 - {}_{0<i<=k}[defRepInf(a_j,n_i)]$,
if ${}_{0<i<=k}[defRepInf(a_j,n_i)] <= 1$ **and**
$defendantRep(a_j) = 0$, if ${}_{0<i<=k}[defRepInf(a_j,n_i)] > 1$
(8)

4.2. EVALUATING WITNESSES' REPUTATION

In case the defendant agents are absolved by the judgment subsystem, the reputation of the defendant agent is not modified. It is the reputation of the witness agent that should decreased since it has told a lie. The witness reputation is also evaluated by using the *power of the norm*. However, violating a norm is usually considered more dangerous than accusing another one of violating it. Therefore, we have defined a factor for adapting the power of the norm for witnesses that lie. Such factor, called witness factor, must be less than 1 but higher than (or equal to) 0, in order to decrement the power of the norm. Equation (9) modifies the power of the norm by considering relapses and a liar witness.

$normPower^{rw}(n_i) = normPower^r(n_i) * witnessFactor(n_i)$ but $0 <= normPower^{rw}(n_i) <= 1$ *(9)*

Equation (11) evaluates the reputation of the witness agent a_j by considering that it has provided k false testimonies. It puts together the partial influences of the lies it has told, as stated in equation (10).

$witRepInf(a_{j,} \quad n_i) = normPower^{rw}(n_i) *$
$blamePercentage(a_j,n_i) * remainingDays(a_j,n_i)$ *(10)*

$witnessRep(a_j) = 1 - {}_{0<i<=k}[witRepInf(a_j,n_i)]$,
if ${}_{0<i<=k}[witRepInf(a_j,n_i)] <= 1$ **and**
$witnessRep(a_j) = 0$, if ${}_{0<i<=k}[witRepInf(a_j,n_i)] > 1$
(11)

4.3. REPUTATION TYPES

Trust and reputation are context dependent [17]. If we trust a person when he is driving a car it does not means that we will trust him when he is piloting an airplane. In addition, if we trust a taxi driver when driving in New York it does not means that we will trust him when informing about any New York address.

In order to take on account the context while evaluating the reputation of agents, we consider two perspectives: the role played by the agent and the service being provided. A person may have a good reputation being a taxi driver but a terrible reputation being a pilot. Moreover, although a person has a very good reputation driving his taxi, he may have a not so good one when giving information about addresses.

To deal with the distinct contexts, three different kinds of reputations were defined*: local reputation*, *role reputation* and *norm reputation*. The *local reputation* of an agent, equation (12), is the one evaluated by the average of the results provided by (8) and (11). The *local* reputation of an agent considers all violated norms and all told lies in a given organization Org_n.

$localRep_{Orgn}(a_j) = [defendantRep_{Orgn}(a_j) + witnessRep_{Orgn}(a_j)] / 2$ *(12)*

Role reputations only consider norms that were violated while playing a specified role or lies that were told while playing this role. Our proposed reputation model is capable of identifying social structures and evaluating the reputation of the agents according to those structures. For each role played, the agent has an associated role reputation. The equation used to evaluate a role reputation is similar to the one used for local reputations, but now we consider only the norms violated while the agent is playing a given role r, as

depicted in equation (13). By using the information provided by role reputations it is possible to know if the agent is trustful to play a role. For instance, it is possible to know if the reputation of an agent is good while considering to pilot airplanes.

$$roleRep(a_j^r) = [defendantRep(a_j^r) + witnessRep(a_j^r)] / 2 \qquad (13)$$

Norm reputations focus on the violation of a norm and on the lies told while considering a norm. Norm reputations are independently of the role being played. For each system norm, the agents have one norm reputation that is evaluated by the average of equations (7) and (10), as illustrated in equation (14). N_i is the norm being considered. By using the information provided by norm reputations it is possible to know if the agent can be trusted for providing a service. It is possible to know it a taxi driver can be trusted while providing information about the New York addresses.

$$normRep(a_j,n_i) = [defRepInf(a_j,n_i) + witRepInf(a_j,n_i)]/2 \qquad (14)$$

4.4. PUTTING TOGETHER THE AGENT REPUTATIONS

As stated before, we assume that large-scale multi-agent systems are composed of sets of organizations grouped in a hierarchy structure. In such systems, an organization can define several sub-organizations but a sub-organization can only be part of one super-organization. Each organization defines its own norms that must be obeyed by agents playing roles in it and also by agents playing roles in any of its sub-organizations. Norms defined in organizations are also valid in their sub-organizations. Moreover, a norm defined in a sub-organization cannot contradict a norm defined in its super-organization. Norms of sub-organizations can only be more restrictive than norms of their super-organizations.

Figure 3 illustrates norms defined in different levels of an organization hierarchy. Norms 1, 2, 3 and 4 are defined in the first level of the hierarchy represented by organization Org 1. These four norms must be obeyed not only by agents playing roles in Org 1 but also by agents playing roles in all its sub-organizations, i.e., Org 1.1, Org1.2 and Org 1.2.1. Norm 5 illustrates that sub-organizations can define their own norms. Norms 6, 7, 8 and 9 exemplifies that sub-organizations can refine norms defined in their super-organizations. As a consequence, agents playing roles in Org 1.2 must obey in fact norms 1, 7, 3 and 8.
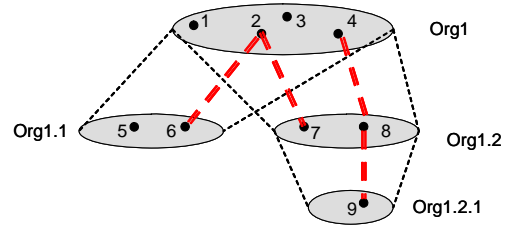

Figure 3: Organization hierarchy

The reputations of the agents are evaluated according to the norms violated in the organizations where they are playing roles. The three reputation kinds defined in Section 4.3 (*local*, *role* and *norm* reputations) are used to evaluate the reputations of the agents in each organization. Each organization evaluates the three reputation kinds considering its own norms and the norms defined in their super-organizations. Those reputations do not include the violations performed in their sub-organizations. In order to consider those violations while evaluating the reputation of an agent, three others reputation kinds are available:

(i) $globalRep_{Orgx}(a_j)$ represents the average of the reputations evaluated in Org_x and in all its sub-organizations, as stated in equation (15);

(ii) $globalRoleRep_{Orgx}(a_j^r)$ represents the average of the reputations evaluated while the agent is playing a given role r in Org_x and in all its sub-organizations (if it is the case), as depicted in equation (16);

(iii) $globalNormRep_{Orgx}(a_j, n_i)$ represents the average of the reputations evaluated according to the violation of a given norm in Org_x and the same norm[1] in all its sub-organizations, as stated in equation (17).

For organizations that do not have sub-organizations, for instance Org 1.1, the global reputations are equal to the local reputations.

$$globalRep_{Orgx}(a_j) = \{ localRep_{Orgx}(a_j) + [\ _{0<m<=n}globalRep_{Orgm}(a_j) / i] \} / 2 \qquad (15)$$

$$globalRoleRep_{Orgx}(a_j^r) = \{roleRep_{Orgx}(a_j^r) + [\ _{0<m<=n}globalRoleRep_{Orgm}(a_j^r) / i] \} / 2 \qquad (16)$$

$$globalNormRep_{Orgx}(a_j,n_i) = \{ normRep_{Orgx}(a_j,n_i) + [\ _{0<m<=n} globalNormRep_{Orgm}(a_j,n_i) / i] \} / 2 \qquad (17)$$

Norms defined in organizations that are not in the same hierarchy do not influence the reputation of agents playing roles in those organizations. For instance, while evaluating the reputation of an agent in Org 1.2 the violations that this agent may have done in Org 1.1 do not influence its reputation in Org 1.2 but will influence its reputation in the Org 1 point of view.

---

[1] Such norm is of course a norm defined in Org_m or in its (super-…)super-organization.

5. CASE STUDY: CARGO CONSOLIDATION AND TRANSPORTATION

In order to validate our approach we present in this section a case study. Our purpose is to illustrate how the governance mechanism (the judgment system together with the reputation system) can be used to regulate application norms. In this context, we will present some aspects of the cargo consolidation and transportation domain and exemplify how two application norms together with the associated strategies are used (by the judgment system) to verify when agents have violated the norms and how the reputation of those agents are being affected by the violation (evaluation provided by the reputation system).

Cargo consolidation is the act of grouping together small shipments of goods (often from different shippers) into a larger unique unit that is sent to a single destination point (and often to different consignees), in order to obtain reduced rate of shipping. Importers and exporters that want to ship small cargos may look for consolidator agents that provide cargo consolidation services to ship their goods.

An open multi-agent system approach is entirely adequate for developing applications on this domain because such applications mostly involve interactions between different autonomous partners playing different roles in order to accomplish similar objectives. Such an approach may also provide support for the automation of the negotiation between the agents looking for reducing the prices and the delivery time. In addition, such applications are particular governed by several rules that are used to regulate the behavior of the heterogeneous and independently designed entities that reinforce the open characteristic of the systems. Since such rules are adequately modeled as norms in governance multi-agent systems, this paper focuses on presenting a governance system for regulate norms that not only are related to the interactions between the agents (that can also be accomplished by other approaches) but also to actions that can modify the application environment (for instance, by updating the state of a resource, changing the position of an agent in the environment or in an organization.)

In a cargo consolidation system there are three main groups of agents: importers, exporters and cargo consolidator agents. The system is supposed to give support to several activities that are regulated by numeral norms, since the system is being implemented as an open multi-agent system. The system was designed to contemplate three organizations: the main organization (Org 1), the sub-organization where importers are responsible for contract the consolidators (Org1.1) and the sub-organization where exporters are responsible for contract the consolidators (Org1.2). The norms we will use to exemplify our approach are both defined in Org1 and, thus, must be obeyed by the agents playing the consolidator role in all (sub-) organizations.

5.1. NORM 01

The consolidator agent must not change its shipment schedule once it has been presented.

The violation of this norm may be testified by importers or exporters that have been injured. On one hand, bringing the delivery date forward may be favorable to some importers whose assembly line is waiting for some supply but will be prejudiced to exporters that will have to delivery their goods to the consolidator agents before the deadline agreed. On the other hand, the postponement of the delivery time may be favorable to exporters that will be able to delivery their goods after the deadline agreed but will be prejudice to importers that will receive their goods later on.

5.1.1. THE JUDGMENT SUBSYSTEM

We are supposing that the same agent has violated two times norm 01 while playing the consolidator role in Org 1.2 and that its global reputation in Org 1.2 is 0.94 when the judgment system receives the testimony about the second violation. In this section, we present the judgment process that judges this testimony. We detail the two application dependent steps (Steps IV and V) and also the application independent Step VI that makes a consensus between the testimonies. Let's suppose that a testimony was provided by one of the application agents (an importer, for instance) stating that an agent (the agent consolidator) has violated norm 01 (Step I). After checking that norm 01 really applies to the defendant agent (Step II) and that the defendant did not confess that has violated it (Step III), it is necessary to judge the testimony according to the particular characteristics of norm 01 (application dependent Step IV).

In order to judge testimonies stating violation of norm 01, such testimonies must inform shipment schedule firstly defined by the consolidator agent and the actual shipment schedule. The strategy used supposes that there is a system's resource that stores the shipment schedules. The resource is analyzed with the aim to compare the information provided in the testimony with the stored information. If the schedule provided by the resource is equal to the first schedule available in the testimony, the schedule was not changed and the testimony is discarded. If the schedule provided by the resource is different to the actual schedule provided by the testimony, the testimony is also discarded because the testimony describes a fact that

cannot be confirmed. In both cases the witness is providing a false testimony. The judgment process is finished and the defendant is considered 100% innocent (Step VII).

Nevertheless, if the schedule provided by the resource is equal to the actual schedule provided by the testimony, the judgment process should continues in order to try to find out if the schedule was really changed. Since the application does not have logs to inform when resources are updated, the alternative to find out if the consolidator agent has really changed the schedule is to ask other agents about their opinions (application dependent Step V). The information provided by the witness is confronted with the information provided by other agents, in this case, with the opinion of two others importers and two exporters, that participate in the system, about the violation of norm 01.

The decision (Step VI) is established based on the information provided by the testimony, the defendant statement and the importers' and exporters' depositions by using subjective logic. Such testimonies and depositions are analyzed from the point of view of the judge and, therefore, there is a need for evaluating how much the judge believes in each agent, i.e $b^J(a)$. As stated before, the reputation of the agent (provided by the reputation system) reflects how much the judge believes in the agent; $b^J(a) = rep(a)$. The reputations of the agents at the moment of the judgment are illustrated in Table 1. Such reputations are the global reputations of the agents in Org 1.2, where the norm was violated.

Table 1: The judge's beliefs

	Reputation	$b^J(a)$
Consolidator Agent	0.95	$b^J(c) = 0.95$
Importer1	0.75	$b^J(i1) = 0.75$
Importer2	0.97	$b^J(i2) = 0.97$
Importer3	0.74	$b^J(w) = 0.74$
Exporter1	0.57	$b^J(e3) = 0.57$
Exporter2	0.64	$b^J(e3) = 0.64$

The judge's beliefs are used to evaluate the judge's opinion about the testimonies and depositions provided by the agents. Such opinions ($w^{J:W}(x)$, $w^{J:C}(x)$, $w^{J:I1}(x)$, $w^{J:I2}(x)$, $w^{J:E1}(x)$ and $w^{J:E2}(x)$), evaluated by using equation (1), are depicted in Table 2 and Table 3. We are supposing that the two importers and the two exporters, together with the witness, have stated that the defendant is guilt ($w_A(x)$). The verdict, i.e the judge point of view about the violated norm, can be provided by applying the consensus operator illustrated in equation (2). In this example the

verdict (equation (18)) states that the probability of the consolidator agent has violated norm 01 is 67%.

$$w^J = w^{J:W}(x) ⁻ w^{J:C}(x) ⁻ w^{J:I1}(x) ⁻ w^{J:I2}(x) ⁻ w^{J:E1}(x) ⁻ w^{J:E2}(x) = <0.67, 0.31, 0.02> \qquad (18)$$

Table 2: Judge's beliefs according to norm 01

	Statement	$w^A(x)$	$b^J(a)$
Witness (Importer3)	Guilty	<1,0,0>	0,74
Consolidator	Innocent	<0,1,0>	0.95
Importer1	Guilty	<1,0,0>	0.75
Importer2	Guilty	<1,0,0>	0.97
Exporter1	Guilty	<1,0,0>	0.57
Exporter2	Guilty	<1,0,0>	0.64

Table 3: Judge's opinion about the violation of norm 01

	$w^{J:A}(x)=w^J(a)~ w^A(x)$
Witness (Importer3)	$w^{J:W}(x) = <0.74,0,0.26>$
Consolidator	$w^{J:C}(x) = <0,0.95,0.05>$
Importer1	$w^{J:I1}(x) = <0.75,0,0.25>$
Importer2	$w^{J:I2}(x) = <0.97,0,0.03>$
Exporter1	$w^{J:E1}(x) = <0.57,0,0.43>$
Exporter2	$w^{J:E2}(x) = <0.64,0,0.36>$

5.1.2. THE REPUTATION SUBSYSTEM

We are supposing that the second violation of norm 01 in Org 1.2 has occurred 5 days after the first violation and, at that moment, the global reputation of the consolidator agent in Org 1.2 was 0.95, as illustrated in Table 1. The judgment of the second violation, exemplified in Section 5.1.1, states that the probability of the consolidator agent has violated norm 01 five days after has violated the same norm is 67%. The goal of this is to evaluate: (i) the norm reputation of the consolidator agent by considering the two violations of norm 01 in Org 1.2, and (ii) the global reputation of the consolidator agent in the point of view of Org 1.2 and Org 1. Those reputations are evaluated in the same day of the second violation. In addition, we are considering that the agent has not confessed the second violation, the power of norm 01 is 0.3, the period during while the agent's reputation may be influenced by the norm is 15 days and the relapse factor is 0.9.

The norm reputation of consolidator agent a^c in the point of view of Org 1.2. Since this agent has never provide a false testimony (witnessRep$_{Org1.2}(a^c)$=1), the *normRep$_{Org1.2}(a^c,n_1)$* only takes into account the two violations of norm 01.

$$normRep_{Org1.2}(a^c,n_1) = \{defendantRep_{Org1.2}(a^c,n_1) +1\}/2 = \{(1 - [0,198 + 0,22]) + 1\} / 2 = 0,79$$

$defendantRep_{Org1.2}(a^c, n_1) = 1-$
$[defRepInf_{Org1.2}(a^c, n_1)_{v1} + defRepInf_{Org1.2}(a^c, n_1)_{v2}]$

$defRepInf_{Org1.2}(a^c, n_1)_{v1} = normPower^{rc}(n_1) *$
$* blamePercentage(a^c, n_1) * remainingDays(a^c, n_1)$
$= 0,3*1,0*0,66 = 0,198$

$normPower^{rc}(n_1) = normPower(n_1) * (1/relapse(n_1)) *$
$* confession(n_1) = 0,3 * (1/1) * 1 = 0,3$

$remainingDays(a^c, n_1) =$
$[totalTime(n_1) - passedDays(a^c, n_1) / totalTime(n_1)]$
$= (15 - 5)/15 = 0,66$

$defRepInf_{Org1.2}(a^c, n_1)_{v2} = normPower^{rc}(n_1) *$
$* blamePercentage(a^c, n_1) * remainingDays(a^c, n_1)$
$= 0,33*0,67*1,0 = 0,22$

$normPower^{rc}(n_1)_{v2} = normPower(n_1) * (1/relapse(n_1)) *$
$* confession(n_1) = 0,3 * (1/0,9) * 1 = 0,33$

$remainingDays(a^c, n_1)_{v1} =$
$[totalTime(n_1) - passedDays(a^c, n_1) / totalTime(n_1)]$
$= (15 - 0)/15 = 1,0$

***The global reputation of consolidator agent in the Org 1.2 point of* view.** Given that this is the second violation of the agent for the same norm (norm 01), playing the same role (consolidator), in the same organization Org 1.2, the norm reputation, role reputation and local reputation in the Org 1.2 point of view are the same.

$globalRep_{Org1.2}(a^c) = localRep_{Org1.2}(a^c) =$
$= roleRep_{Org1.2}(a^c) = normRep_{Org1.2}(a^c, n_1) = 0,79$

The global reputation of consolidator agent in the Org 1 point of view. Note that the consolidator agent has never violated a norm or told a lie in the organizations Org1 and Org1.1 ($globalRep_{Org1}(a^c) = globalRep_{Org1.1}(a^c) = 1$).

$globalRep_{Org1}(a^c) = \{ localRep_{Org1}(a^c) + [(globalRep_{Org1.1}(a^c) + globalRep_{Org1.2}(a^c)) / 2]\} / 2 = 0.94$

5.2. NORM 02
Consolidator needs to delivery the cargo to the importer(s) in the determined location and in the established deadline

The violation of this norm may be testified by the importers that are prejudiced by it.

5.2.1. THE JUDGMENT SUBSYSTEM
In this section, we suppose that the consolidator agent has violated norm 02 while playing role in Org 1. Similar to the judgment presented Section 5.1.1, we focus on the two application dependent steps (Steps IV and V) and on Step VI while illustrating the judgment process of norm 02. As in Section 5.1, we assume that the judge system could not provide a verdict before executing Step IV.

In order to judge testimonies stating violations of norm 02, such testimonies must contain the transportation documents called House Bill of Landing (HBL) and Master Bill of Landing (MBL). A bill of landing is a document issued by the carrier (the consolidator agent, in this case) that describes the goods, the details of the intended transportation, and the conditions of the transportation. The difference between HBL and MBL is that the MBL describes several small cargos consolidated in a single shipment and the HBL describes each small cargo.

Therefore, in step IV, the judge must first ensure that the exporter has really delivered the cargo at the place designated by the consolidator on the appropriated date. When this task is accomplished, the consolidator gives a copy of the HBL (related to the cargo delivered by the exporter) to the exporter. The judge can, therefore, ask to the exporter about his copy of the HBL. If the exporter does not have this document, the judgment process is finished, the witness' testimony is considered false and the defendant is considered 100% innocent (Step VII). The consolidator agent has not delivered the cargo because the exporter has not delivered its cargo to the consolidator agent.

On the other hand, if the exporter has its copy of the HBL the judge must execute step V, continuing the judgment process to come to a verdict. Since, the witness' cargo has been consolidated with others cargos, the judge may ask all other importers mentioned in the MBL for their HBL in order to find out if their cargos have been delivered in the correct date and place. After receiving the importers depositions, the judge needs to execute step VI, where it puts together all statements while considering the reputations of consolidator agent and all importers of the mentioned shipment. We are supposing that there were three cargos consolidated in this shipment. Table 4 and Table 5 depicts the judge's opinion about the testimony and depositions provided by the witness, the consolidator agent and the two importers ($w^{J:C}(x)$, $w^{J:W}(x)$, $w^{J:I1}(x)$ and $w^{J:I2}(x)$). Note that the reputations of the agents ($b^J(a)$) being used in this evaluation are their global reputations in the Org 1 point of view, since norm 02 was violated in Org 1. Therefore, they are different from the one presented in Table 1. According to Section 5.1.2, the global reputation of the consolidator agent in the Org 1 point of view is 0.94.

The verdict, i.e judge point of view about the violated norm, can be provided by applying the consensus operator as show in equation (19). In this example the verdict states that the probability of the consolidator agent has violated norm 02 is 54%.

$w^J = w^{J:W}(x) ^- w^{J:C}(x) ^- w^{J:I1}(x) ^- w^{J:I2}(x) =$
$= <0.54, 0.43, 0.03>$ *(19)*

Table 4: Judge's beliefs according to norm 02

	Statement	$wA(x)$	$b^J(a)$
Witness (Importer1)	Guilty	<1,0,0>	0.85
Consolidator	Innocent	<0,1,0>	**0.94**
Importer2	Guilty	<1,0,0>	0.80
Importer3	Guilty	<1,0,0>	0.91

Table 5: Judge's opinion about the violation of norm 02

	$w^J(a) \sim w^A(x) = w^{J:A}(x)$
Witness (Importer1)	$w^{J:W}(x) = <0.85,0,0.15>$
Consolidator	$w^{J:C}(x) = <0,0.94,0.06>$
Importer2	$w^{J:I1}(x) = <0.80,0,0.20>$
Importer3	$w^{J:I2}(x) = <0.91,0,0.09>$

Note that both strategies presented in section 5.1 and 5.2 are simple examples that can be used to judge the testimonies related to norms 01 and 02. Other more complex and completely different strategies could have been implemented to judge the same testimonies. Our intention while presenting such strategies was to illustrate how simple strategies can be sufficiently good to provide the verdicts based on reputations.

5.2.2. THE REPUTATION SUBSYSTEM

We are considering that the power of norm 01 is 0.8, the period during while the agent's reputation may be influenced by the norm is 30 days. In this section we are interested in evaluate the global reputation of the consolidator agent in the Org 1 point of view that may take into account the three violations.

$globalRep_{Org1}(a^c) = \{ localRep_{Org1}(a^c)+$
$[(globalRep_{Org1.1}(a^c) + globalRep_{Org1.2}(a^c)) / 2]\}/ 2$
$= 0.83$
$localRep_{Org1}(a^c) = [defendantRep_{Org1}(a^c)+1]/2 =$
$[(1- defRepInf_{Org1}(a^c)) +1] /2 = [(1 - 0.43)+ 1] /2$
$= 0.78$
$defRepInf_{Org1}(a^c) = normPower^{rc}(n_2) *$
$* blamePercentage (a^c,n_2) * remainingDays(a^c, n_2)$
$= 0.8*0.54*1=0.43$
$normPower^{rc}(n_1) = normPower(n_1) * (1/relapse(n_1))$
$* confession(n_1) = 0.8 * (1/1) * 1 = 0.8$
$remainingDays(a^c,n_2) =$
$[totalTime(n_2) - passedDays(a^c,n_2) / totalTime(n_2)]$
$= (30 - 0)/30 = 1$

6. RELATED WORK

Since in this paper we present a governance system that makes use a reputation subsystem, the related work should contemplate not only the available governance or law-enforcement systems but also the already published reputation systems.

6.1. GOVERNANCE SYSTEMS

Different enforcement systems have been proposed in the literature. The majority, such as [14][15][10], focuses on regulating the interaction between agents. They usually provide governors [9] or law-governed interaction (LGI) [14] mechanisms that mediate the interaction between agents in order to regulate agent messages and make them comply with the set of norms. Every message that an agent wants to send is analyzed by the mechanism. If the message violates an application norm, the message is not sent to the receiver. The main disadvantages of such approaches are (i) they influence the agents' privacy since those mechanisms interfere in every interaction between agents and (ii) they do not govern non-dialogical actions since they only concern about the compliance of messages with the system norm [20]. Therefore, they can not govern norm 01 described in the above section.

Other approaches provide support for the enforcement of norms that regulate not only the interactions between agents but also the access to resources [5] and the execution of actions [20]. TuCSoN [5] provides a coordination mechanism to manage the interaction between agents and also an access control mechanism to handle communication events, in other words, to control the access to resources. In TuCSoN agents interact through a multiplicity of independent coordination media, called tuple centres. The access control mechanism controls agent access to resources by making the tuple centres visible or invisible to them. Although in TuCSoN norms can be described to govern the access to resources, the governance is restricted and only applied to resources that are inserted in tuple centre environments.

In [20] the authors claim that the governance system enforces the observable behavior of agents in terms of public messages and visible actions. They introduce a classification of norms and, according to such classification, they provide some implementation guidelines to enforce them. The main drawback of this approach is that it does not provide support for the enforcement of messages and actions that are not directly accessed by the governance system. Such an approach assumes that the governance system can enforce every norm since it can access all messages and actions regulated by a norm. But in open MAS with heterogeneous and independently designed agents, there will be private messages that can only be perceived by senders and receivers and execution of actions that can only be noticed by the agents that are executing them or by a group of agents that suffers from their violations [2]. For instance, norm 02 cannot be regulated by such approaches since the messages exchange between the involved agents are not public ones.

6.2. REPUTATION SYSTEMS

Centralized reputation systems used by eBay [8], Amazon Auctions [3] and Sporas [22] were developed in order to inform buyers about the performance of sellers in previous negotiations. Such systems represent the sellers' performance by attributing to each one a single global reputation value. The system receives from buyers their personal evaluations about the performance of the sellers during the interactions. The system puts together all these information to update the reputation of the sellers.

The main differences between our centralized approach and those ones are: (i) the reputations provided by the centralized system are not biased by the agents' point of view. The centralized system does not receive the evaluations made by one agent about the performance of its partner. Our centralized reputation system receives testimonies about norm violations. The judgment subsystem judges these testimonies and informs the reputation subsystem about their veracity. The reputation subsystem uses the verdict provided by the judgment subsystem to update the reputation of the sellers; (ii) the system can provide different reputations for each agent according to different contexts. Three distinct contexts were defined, so far: global, norm and role context. The different reputations help agents to foresee the behavior of agents in different situations; and (iii) our approach provides the possibility to group the agents in several subsystems that can together provide a more stable support for the evaluation of the reputations.

As stated before, our reputation model is a hybrid one. To implement the decentralized reputation part any published decentralized reputation model can be used. The implementation of the decentralized reputation part does not affect the centralized one. In decentralized models such as [1], FIRE [11], and Regret [17] agents are endowed with the capacity to evaluate the interactions and to store then individually. In such models the agents themselves use different information sources to evaluate the trust and reputation of others. Socio-cognitive models such as [6] where trust is considered an agent mental state can also be used to implement the decentralized subsystem.

The most important advantage of our approach is the use of a centralized reputation mechanism, implemented by the organizations, together with the decentralized one. The organizations can provide trustful and unbiased agents' reputations that are accessible by any system agent. This is extremely important when considering two situations: (i) agents that want to know the reputation of other agents with whom they have never interacted; (ii) agents that want to update the reputation of partners with whom they have not been interacting for

long time. In both cases agents can use the reputations provided by the centralized system. Based on theses reputations, an agent decides if it may interact with another agent.

7. CONCLUSIONS

In this paper we have presented a governance system that provides the regulation of multi-agent systems' norms by judging testimonies pointed out by agents about norm violations. The judgment process takes into account the reputations of the involved agents (agents that provided the testimonies and the agents being accused of violating norms) while judging the testimonies.

As illustrated by the case study, the governance system is able to regulate private dialogical actions (exemplified by the violation of norm 02 related to the sending of a private message) and also private non-dialogical ones (characterized by the violation of norm 01 that is related to a resource modification) without influence the agents' privacy, since it is based on testimonies. The judgment system uses the reputations provided by the reputation system while judging the testimonies. Those reputations are evaluated according to the verdicts supplied by the judgment of previous testimonies and according to characteristics of the involved norms. Therefore, the reputations are not simply evaluated by agents' opinions about other agents behavior. The reputations are available not only to the judgment system but also to any application agent. Thus, the agents neither need to meet frequently to have consistence reputations of other agents nor to look for other agents that may have the reputations of the desired ones. They can ask to the reputation systems implemented in the organizations.

Whereas we believe that the advantages of our proposed mechanism are really important, it has some potential weaknesses. First, it may be difficult to distinguish if a testimony is true or false and, therefore, to provide a good verdict. We proposed to solve this problem by using probability based on subjective logic while providing the verdicts. By using subjective logic it is possible to put together different agents' point of view while considering their reputations to determine a verdict. Another important drawback is that violations that go without testimonies will not be punished. This could lead to an undesired system state. One way to overcome this issue is motivating the agents to give their testimonies, using, besides the mechanism for punishing them based on their reputations, an agent rewards program, for instance. Finally, we intent to investigate the possibility of using argumentations during steps IV and V of the judgment system. It would be interesting to

compare the approach using subjective logic and the one using argumentation in terms of flexibility, efficiency and feasibility.

REFERENCES

[1] Abdul-Rahman, A. and S. Hailes. Supporting Trust in Virtual Communities. In: Proceedings of the 33rd Hawaii International Conference on System Sciences, Vol. 6. (2000)

[2] Aldewereld, H., Dignum, F., García-Camino, A., Noriega, P., Rodríguez-Aguilar, J. A., Sierra, C.: Operationalisation of Norms for Usage in Electronic Institutions. In Proceedings of the AAMAS06 Workshop on Coordination, Organization, Institutions and Norms in agent systems (COIN), Hakodate, (2006)

[3] Amazon Site. http://www.amazon.com World Wide Web (2006)

[4] Boella, G.; van der Torre, L.: Regulative and Constitutive Norms in Normative Multi-Agent Systems. In Proceeding of 9th Int. Conference on the Principles of Knowledge Representation and Reasoning. California (2004)

[5] Cremonini, M., Omicini, A., Zambonelli, F.: Coordination and Access Control in Open Distributed Agent Systems: The TuCSoN Approach. In Proc. of the 4th Int. Conf. on Coordination Languages and Models, LNCS 1906, (2000) pp 99-114

[6] Castelfranchi, C. and R. Falcone: Social Trust: A Cognitive Approach. In: C. Castelfranchi and Y. Tan (eds.): Trust and Deception in Virtual Societies. Kluwer Academic Publishers (2001) pp. 55–90

[7] Daskalopulu, A., Dimitrakos T., Maibaum T.: E-Contract Fulfilment and Agents' Attitudes. In Proc. ERCIM WG E-Commerce Workshop on The Role of Trust in e-Business, Zurich (2001)

[8] eBay Site. http://www.ebay.com World Wide Web (2006)

[9] Esteva, M., Rodriguez-Aguilar, J. A., Rosell, B., Arcos, J. L.: AMELI: An Agent-based Middleware for Electronic Institutions. In Proc. of the 3rd Int. Joint Conf. on Autonomous Agents and MAS, USA (2004) pp. 236-243

[10] Esteva, M., de la Cruz, D., Sierra, C.: Islander: An Electronic Institutions Editor. In Proc. of Int. Conf. on Autonomous Agents and Multi-Agent Systems (2002) pp. 1045–1052

[11] Huynh, T. D., Jennings N. R., Shadbolt, N. R.: FIRE: An Integrated Trust and Reputation Model for Open Multi-Agent Systems. In: Proceedings of the 16th European Conference on Artificial Intelligence (2004) pp. 18–22

[12] Jøsang A.: An Algebra for Assessing Trust in Certification Chains. In Proc. Network and Distributed Systems Security Symposium, (1999)

[13] López, F.: Social Powers and Norms: Impact on Agent Behavior. PhD thesis. University of Southampton. UK (2003)

[14] Minsky, N., Ungureanu, V.: Law-Governed Interaction: A Coordination & Control Mechanism for Heterogeneous Distributed Systems. In ACM Trans. on Software Eng. and Methodology, vol 9, no 3 (2000) pp. 273-305

[15] Paes, R.: Regulating the Interaction Between Agents in Open Systems – a Law Approach. Master's thesis, Pontificia Univeridade Catolica do Rio de Janeiro, PUC-Rio, Rio de Janeiro, BR (2005)

[16] Patel, J., Teacy, W., Jennings, N., Luck, M., Chalmers, S., Oren, N., Norman, T., Preece, A., Gray, P., Shercliff, G., Stockreisser, P., Shao, J., Gray, W., Fiddian, N., Thompson, S.: Monitoring, Policing and Trust for Grid-Based Virtual Organizations. In Proc. of the UK e-Science All Hands Meeting 2005 UK (2005)

[17] Sabater, J., Sierra, C.: Reputation and Social Network Analysis in Multi-Agent Systems. In: Proceedings of First International Conference on Autonomous Agents and Multiagent Systems, Bologna, Italy (2002) pp. 475-482

[18] Silva, V., Cortês, M., Lucena, C. J. P.: An Object-Oriented Framework for Implementing Agent Societies, MCC32/04. Technical Report, Computer Science Department, PUC-Rio. Rio de Janeiro, BR (2004)

[19] Singh, M.: An Ontology for Commitments in Multiagent Systems: Toward a Unification of Normative Concepts. Artificial Inteligente and Law v. 7 (1) (1999) pp. 97-113

[20] Vázquez-Salceda, J., Aldewereld, H., Dignum, F.: Implementing Norms in Multi Agent Systems. Lecture Notes in Computer Science, v. 3187 (2004) pp. 313 – 327

[21] Yu, B., Singh, M. P.: Distributed Reputation Management for Electronic Commerce. Computational Intelligence 18(4), (2002) pp. 535-549

[22] Zacharia, G. and P. Maes. Trust management through reputation mechanisms. Applied Artificial Intelligence 14(9), (2000) pp. 881–908