Institutional Environments and Open Collaborative Systems

Marcos de Oliveira¹, Vasco Furtado²³, Stephen Cranefield¹, Martin Purvis¹

¹Department of Information Science – University of Otago
P O Box 56 – 9054 – Dunedin – New Zealand

²Computer Science – University of Fortaleza
Fortaleza, CE – Brazil

³Empresa de Tecnologia da Informação do Ceará (ETICE)
Fortaleza, CE – Brazil

{moliveira,scranefield,mpurvis}@infoscience.otago.ac.nz,
vasco@unifor.br

Abstract. A wiki is a typical Web 2.0 system where people from various backgrounds interoperate and/or cooperate, sharing information and, at the same time, building a knowledge base on different areas of interest. Its characteristics are very similar to the ones owned by the open multi-agent systems modeled as electronic institutional environments. This paper proposes an approach to develop open collaborative systems as institutions of agents and demonstrates that suitability through the adaptation of a typical wiki application, known as WikiCrimes, to an institutional environment.

1. Introduction

One can plainly see that we are at a moment in which collaboration is on the rise. That momentum of collaboration is primordially leveraged by the Web 2.0 in which the quantitative difference between producer and consumer decreases significantly, since several applications emphasize the production of information by any Internet User. With the mass production of content, collaboration mechanisms such as Wikipedia [Wikipedia Team 2009] came almost naturally. Wiki-style applications have been used in several different scenarios like business project management and crowd sourcing [Howe, J. 2008].

Despite those initiatives the design task for these systems still faces big challenges. The open and participative characteristics identified in such a kind of system makes it susceptible to abuses or fraud attempts. It is important that as many people as possible collaborate with the system, contributing for the growing of its data records. But it is as important that the information registered in the system be reliable, so that the system can become a trustful source of information.

In saying that, our main claim is that open collaborative system must be viewed as a kind of open multi-agent system where a number of human and/or artificial agents interoperate pursuing their individual or common goals. These agents can contribute positively or negatively to the organization and goals of the system as a whole, and the
agents are free to join and leave the system as they wish, since they obey certain rules that must be observed concerning the management of the system.

The management of interoperation among agents is a complex task and robust techniques and methodologies to the development of reliable and open Multi-Agent Systems (MAS) have been studied in the academia [Aldewereld, H. 2005], [Grossi, D. et al. 2005], [Arcos, J. et al. 2005], [Minsky, N. and Ungureanu, V. 2000], [Colombetti, M. 2002], [Ricci, A. and Omicini, A. 2003], [Singh, M. P. 1999], [Zambonelli, F. et al. 2001]. Such techniques aim at the modeling and implementation of features that give openness to those agents, allowing them to have the ultimate choice of obeying regulations or deal with possible sanctions imposed by the MAS norms. After all agents are autonomous entities and the biggest challenge is to have a coordination system where the agents can be free to decide what to do but at the same time be encouraged or seduced to obey the regulations of the artificial society where they are entering.

In this paper the suitability of an MAS institutional model for designing an open collaborative system known as WikiCrimes [Furtado, V. et al. 2008], [WikiCrimes Team. 2009] is defended. WikiCrimes aims to offer a common interaction space among the public in general so that they are able to notify criminal facts and follow where they occur as well. The goal is to get the collaborative individual participation for generating useful information for everyone.

The institutional model we are going to present support particular features that are essential for open collaborative systems [De Oliveira, M. et al. 2008], [De Oliveira, M. et al. 2007]. In that model, as in ordinary institutional environments, the agents assume predefined roles in the MAS and will be supervised by some authorities, defined as system agents. The system agents are part of the infrastructure built to manage commitments among agents, or groups of agents, and agents’ reputation in the institution. We intend to show the suitability of this institutional model, in particular because the reputation of the agents can be computed through mechanisms to reward and/or punish agents involved in the interoperation process.

Moreover, the agent roles as well as the interactive process in the institution are modeled via Colored Petri Nets (CPN) [Jesen, K. 1997]. CPN brings to this approach a well-defined semantics which builds upon true concurrency, hierarchical representations and an explicit description of both states and actions. Adding to that CPN has an elaborated set of computer tools supporting their drawing, simulation and formal analysis.

2. Institutional Design and Open Systems

How does one constrain an environment as heterogeneous as a human community? In different parts of the world, different rules and norms are created so that the members of that group of people can feel a sense of security and order, and so that they can carry on with their lives in a more predictable way. When a person needs some kind of service, he knows where to go, and if not, he will ask another person or consult some kind of public catalogue. Once the person chooses to go to a place and make use of some sort of service he will make use of his past experiences and knowledge to carry on with actions, autonomously, and in a more or less standard way. Rules are there to be
followed as well as to guide the members of a community so that they will have their rights observed in local and global aspects.

One of the main goals of our approach is to define and maintain an environment where agents will interact observing a set of rules or norms but not necessarily or compulsorily will their actions be restricted via some kind of interface agent, e.g. the governors, in the Electronic Institution Development Environment (EIDE) [Arcos, J. et al 2005], or the controllers in Law-Governed Interaction (LGI) [Minsky, N. and Ungureanu, V. 2000].

Our approach [De Oliveira, M. et al. 2008], [De Oliveira, M. et al 2007] is to offer mechanisms that will encourage the agents to play a cooperative role in the agent society, but ultimately the choice of cooperating or not cooperating will be its own. That choice might generate the employment of sanctions by the system agents on uncooperative agents, and by so doing, we will be able to use a model closer to the human way of organizing societies.

Therefore, among the empirical issues with respect to electronic institutions, we are more concerned with the autonomy the agents will have with respect to their freedom of expression within the institution. By that we mean that we humans autonomous agents have our individual goals and beliefs, and those are installed into the autonomous artificial agents that we build. They are there, because the agent-oriented paradigm is inspired by the more open and dynamic model of human society. For a completely constrained environment, one may employ the more conventional object-oriented paradigm and infrastructure.

Such mechanisms as governors [Arcos, J. et al 2005] and controllers [Minsky, N. and Ungureanu, V. 2000], in our view, might compromise the agents’ autonomy. We prefer to influence agents to behave according to the rules of the environment.

That is the motivation of our work, to develop an environment where agents are influenced to cooperate and follow a predefined set of rules. That environment is organized based on institutional concepts with the defined roles for joining agents. These concepts can be used to create an artificial environment similar to real world institutions, where people can join to obtain or offer access to services.

Autonomous agents can be modeled to copy human reasoning or strategies when interacting with others and deciding their course of action in the electronic environment. Interaction protocols can be used in accordance with the speech acts, institutional actions or illocutions identified in the dialogs executed by agents. Those interaction protocols help to predict actions and model the conversation space before the event, including possible exceptions of expected courses of action by agents in the electronic environment. Authorities in the institutions are available to audit interactions and observe rules in the society.

The rigid control of agents’ interoperation may or may not grant rigid security for MAS, but it definitely interferes with the agents’ autonomy. If agents are only allowed to follow the rules, part of the intrinsic characteristics of the multi-agent model might be compromised and their capability to deal with real world situations diminished. We then advocate for an environment model in which we can define norms, with applicable sanctions if they are violated. The performance of the system as a whole must be observed so that autonomous agents, though free to misbehave, must recognize
the possibility of losing reputation points as well, which will make them less attractive as a partner for future interoperation.

When modeling and implementing open agent systems that allow heterogeneous agents to join the system and perform tasks we use the abstraction of institutional environments and normative spaces. Since the agents that join the institution are heterogeneous the necessity of the insertion of social norms in the system becomes evident. Norms are introduced to balance the functioning of the system and introduce a variable mechanism of control in the environment. With that mechanism the goals of the system as a whole can be met when autonomous agents are seeking their own goals.

The degree of openness observed in normative systems is variable. Norms define what is legal or illegal in the system and at the same time influence the agents to behave in a desired way, much like legal frameworks are developed to guide humans in the real world. It is important to observe the necessity of having this control over the autonomous agents in order to grant a sense of order to MAS, and a degree of openness as variable as the domain modeled needs or the system designer’s desire.

Interaction and coordination are identified as major concerns when designing and deploying MAS, giving a distinct approach toward the modeling and design of distributed intelligent systems. Software engineering agent-oriented methodologies, such as Gaia [Wooldridge, M. 2005], have been developed to observe the interaction between agents as a critical design aspect when building MAS.

The system organization also influences the design of MAS. A social setting is realized in the form of an environment where agents play roles and interact with each other pursuing individual or common goals. The organization has a defined structure, which defines and enforces norms to manage the interoperations among agents. Norms are associated with roles agents assume in the system upon registration and will guide the agent behavior in the system.

The operational use of norms in institutional environments is directly related to the context in which the agents interoperate, and it is defined through ontologies [De Oliveira, M. et al. 2007], [Grossi, D. et al. 2005]. Our approach is to build an institutional environment, as defined in [De Oliveira, M. et al. 2008], [De Oliveira, M. et al. 2007], out of a collaborative system and, with that, demonstrate its suitability to model this category of systems.

3. WikiCrimes as an Open Collaborative System

WikiCrimes is driven by three goals: i) to give more transparency and publicity to criminal information, ii) to provide means to citizen prevention, and iii) to reduce underreporting phenomena (crimes that are not notified to the authorities) so common in that context. These goals have always been in the agenda of several countries around the world, particularly, those who the populace suffers with the high rates of violence.

The idea behind WikiCrimes [Furtado, V. et al 2008], [WikiCrimes Team. 2009], Figure 1, is to provide a common area of interaction among people so that they can make the reports and monitor the locations where crimes are occurring. It is based on the principle that the ones who hold information about crimes are the citizens. If they want to make such information public, they can. Thus, individual participation, in a collaborative manner, can generate knowledge of the masses. In other words, if there is
active participation, crime mapping starts being done collaboratively, and everyone will benefit from having access to information about where crimes occur.

In wiki-style open systems, especially WikiCrimes, there must be mechanisms for rewarding and/or punishing agents involved in the interoperation process. In this kind of systems there is a constant trade-off between diminishing the constrains imposed to the agents with the intention to grow the number of participants in the system, and the rigid control that can be imposed to avoid unwanted behavior, as for example the registering of fake crimes. In WikiCrimes there are not many requisites to become a member of the system, the only personal questions asked are a valid email address and a name, no real identification numbers are asked so that people do not be afraid to post information in the system. But for every criminal fact registered in the system it is requested that there be indication of at least one person that can confirm that the information posted is true.

4. WikiCrimes as an Institutional Environment

In an institutional environment a group of persons agree to follow a set of regulations in order to develop a fruitful relationship with the others participating in the institution. That set of regulations is formed based on the actions each individual can perform in the institutional environment, according to the role they play. In other words, institutional actions [Colombetti, M. 2002] taken by agents in an agent society must follow the rules imposed by a certain set of regulations defined by an institution. Those institutional actions are in fact the speech acts identified in the context of an institution that are used as illocutions in conversations and are used here to predefine courses of actions and build interaction protocols.

4.1. System Elements

A significant characteristic of our approach is its open and distributed nature. Elements are organized in such a way that they are not compelled to report their actions to any
other participating agent in the structure. From [De Oliveira, M. et al. 2008], [De Oliveira, M. et al. 2007] we extract the definition of Institutional Environment (IE) and here we define the WikiCrimes Institutional Environment $W_{IE}$ as:

$$W_{IE} = (O_s, N_s, I_a)$$

Where: $O_s$ stands for system ontology. Here it defines generic terms used in any institutional environment, such as: request protocol, request role, request reputation, and inform reputation; $N_s$ stands for normative space. This will be defined below and $I_a$ stands for institutional actions. In this context we will list all the action identified in the context of WikiCrimes.

The institutional environment represents well defined groups of agents that together form organizations that follow a set of regulations, which specify how agents should undertake activities in a specific domain. Therefore, we use institutional actions to identify standard dialogs that take place in an institutional environment and define CPNs that will manage the interaction protocols necessary to achieve the wanted outcomes. Some institutional actions identified are: register in the WikiCrimes IE, register a crime, search the data base, confirm a crime, denounce abuse and disconfirm a crime.

$$N_s = (R_s, R_e, O_c, W_f, L)$$

Where: $R_s$ stands for system roles; $R_e$ stands for external roles and correspond to roles available to be played by agents that register in the institutional environment; $O_c$ stands for context ontology; $W_f$ stands for workflow, see Figure 2 and $L$ stands for content language and represents the language expressed in the content attribute of the FIPA (Foundation for Intelligent Physical Agents) messages exchanged among agents in the $N_s$.

![Figure 2: Partial view of the WikiCrimes Institutional Environment's workflow (Wf).](image)

In Figure 2 the circles represent states and the squares represent sub-nets as defined in [De Oliveira, M. et al. 2008], [De Oliveira, M. et al. 2007]. The sub-net is actually another CPN [Jesen, K. 1997] which, in this specific case, defines agent roles. So in the $W_f$ level the workflow or the interoperation process as a whole is identified.
The system level infrastructure identifies system agents that assume roles from \( R_a \) and are deployed to manage the interactions in the normative space. Following is the description of such agents.

**Institution Agent:** The first step for an agent to get into the institution is to register with the Institution Agent and assume a role in the artificial society. Upon registering, the agent will gain access to a representation of \( W_f \) in the form of a CPN that represents a suggested path or course of actions the recently registered agent should follow in order to obey the norms that regiment the institutional environment. The \( W_f \) can be as constrained as the MAS developer wants; it depends on the degree of restrictions, security or access he wants to deploy in the system, such as for example, demanding the use of monitored interactions for certain activities. It is important to mention here that the member agents will have their own goals and strategies to obtain them, and even though they have knowledge about the \( W_f \), they still can participate freely in conversations without compulsorily following it, but at least they all know the rules of the game.

**Monitor Agent:** This system agent represents a monitoring authority in the institutional environment. It monitors certain activities, defined as institutional actions, in accordance with the norms defined in the normative space by the \( W_f \). The monitoring is done through the use of commitments and observation by the monitor agent of the commitments’ life cycle. Once in the artificial society, the agent can commit itself to perform tasks and request other agents to make commitments to perform tasks for it. In the case that the agent does not have an acceptable level of trust of the agent with which it is starting the interoperation process, it can ask for a monitored interaction, where the monitor agent will audit the commitment shared by the agents engaged in interaction. That interaction is stored in a database of audited interactions for later examination, if requested to the monitor agent.

**Reputation Agent:** The agents in the open MAS have access to a system agent called the Reputation Agent. This agent is responsible for giving information about other agents that have been present at some time in the society and have developed a reputation. The agents are not obliged to report or to consult the Reputation Agent prior to every interoperation that they take part in. This offers flexibility and the possibility of open MAS implementations with verifying degrees of agent autonomy.

The \( R_a \) is composed by the following roles:

**Browser User:** The user that browses the institutional environment basically seeks information in the system. The institution will try to encourage this kind of user to assume the Registered User Role.

**Registered User:** in this role the agent will a typical user of a collaborative system. In the case of the \( W_{IE} \), it will be able to register crimes, browse the environment, confirm crimes, denounce abuse, disconfirm crimes, leave messages to the Institutional Agent about the \( W_{IE} \), indicate other agents playing the Registered User or Invited User roles to confirm crimes.

**Invited User:** this role is played by the agent that is indicated to confirm a crime. The indication is done by an agent playing the Registered User role. This agent is one step closer than the Browser User to being able to play the Registered User role.
**Certifier Entity:** This is a special kind of role played by agents that own a very respected position in the $W_{IE}$. This agent has the reputation of a System Agent, the difference is that the Registered Users can denounce abuses or disconfirm crimes registered for this category of agent and they can have their position in the $W_{IE}$ reviewed.

**Administrator:** this role works closely with the Institutional Agent, observing the breaking of the norms of the $W_{IE}$ and performing administrative tasks, such as changing Registered Users properties and managing Certifier Entities.

### 4.2. Agent Roles as CPN

In our approach, CPN are used to represent agent roles and agents conversations. Those two concepts together define what we call the Conversation Space of an institution. The overall institution conversations are represented as a CPN, as in Figure 2, that has specific access points, where smaller CPNs, representing roles of individual participating agents, can be plugged in. With that approach, an agent will be involved in the part of the conversation appropriate for its role in the institution. Figure 3 depicts a CPN that represents partially an Invited User role in the $W_{IE}$.

![CPN Diagram](image)

**Figure 3:** Partial view of the Invited User role CPN.

Every agent in the institutional environment will have a local state that concerns its role in the whole system. By distributing the state representation among a set of tokens, it is easier to represent the local state of an individual agent role (in the context of a larger institution), and this can be useful for local management of individual agents.

5. Norms and Commitments

In our approach, a commitment is an object created at the agent level and handled by the Auditor agent involved in the interaction. The Debtor is the agent that makes the commitment, and the Creditor is the agent relative to which the commitment is made.

We have adapted a model defined in [Colombetti, M. et al. 2002] to our needs. The commitment states used are: unset, pending, cancelled, active, violated, fulfilled and not fulfilled. The not fulfilled commitment state is introduced to differentiate a state identified when the commitment’s task does not have a deadline associated with it and the commitment is never violated nor fulfilled, because commitment violation is directly related to a deadline.

The Monitor Agent will observe the normative space when monitoring interactions. Operations on commitments reflect its state and, together with the norms defined in the Wf, a monitored agent can experience a change in its system reputation, for better or worse, or, in an extreme case, could lead to the banishment of the agent from the institutional environment.

The audit process performed by the Monitor Agent can be visualized as the proof given by the debtor agent that he performed some task. That would have the form of a commitment with condition equal to “true,” acknowledging some information about a task the debtor should perform. Basically, the debtor would be committing itself formally to an authority in the MAS (represented by the Monitor Agent) that it did something. If an agent commits itself too often to false statements, the Monitor Agent should receive many complaints about that agent, which would lead the Monitor Agent to use its power to update the reputation of agents in the MAS.

The normative space guides the Monitor Agent during the monitoring process; its role is a commitment management protocol that manages commitment objects according to their life cycle. Participating agents can decide to break the protocol, but the normative space defines the level of tolerance to those actions, and depending on the application domain that the institutional environment is implementing, agent acts can decrease its reputation to such a level that other agents which consult the Reputation Agent will cease to communicate with the untrustworthy agent. The time factor is an important element in the representation of commitments. To express a commitment formally, it is necessary to find a representation mechanism able to handle the time constraints found in the definition of commitments.

6. Reputation and Trust

The Institution Agent does not maintain any control over the registered agents. In fact, it is the reputation that the agent has in the MAS that will determine its useful existence in the institution. Based on the trust the agents have for each other, they will interact or not. With time and the development of a number of interactions, agents can build up trust networks and establish trust relationships with each other.

The reputation the agent develops in the institutional environment, together with restrictions of the normative space may restrict levels of access to certain resources. Such concepts as level-of-access to resources and services can be modeled in the normative space through the context ontology. Therefore, since all the agents are aware
of the rules of the institutional environment they know that their actions might not only affect their reputation, but diminish their level of access to resources in the system.

The reputation update model sometimes needs to express characteristics of the context in which the institution was developed, and different reputation update models can be attached to the Reputation Agent CPN to express that. Another important aspect of that approach is that the external agents can have their own definition of trust and use the one defined by the institution to add information to it, or simply ignore it.

The agents are not compelled to use the Reputation Agent before every transaction. They can have, for example, a history of their conversations internally and their own information about other agents and choose to refuse certain kinds of interactions from some agents, that they might not trust. But in case they are willing to use the information agent, it is available. Being aware of the normative space, they can calculate the risk of losing privileges in the institutional environment and act as they will.

One of the main concerns in WikiCrimes, and in collaborative systems in general, is the number of hoaxes or in our case false criminal facts registration. How to avoid that a person starts to spread false information in the system? Why does that specific area have so many crimes? Is it really true or someone is trying to make a joke, speculate about the real estate market or diminish the image of the local police? The reputation model chosen to regulate the WIE will play a strong role in the matter of identifying this “bad agents” in the institutional environment.

In WikiCrimes there is the concept of reputation for the source of information, i.e. whoever registers the criminal fact. Some entities such as the press and governmental bodies, are labeled certifier entities, and therefore considered very well reputed. But that is not enough, the open characteristic of the system, in the sense that anyone could be a user, does not make easy the task of knowing the reputation of all the users of the system. To solve this problem, for every criminal fact posted in the system, it is asked of the users that they indicate at least another person, trough their email address, who can confirm the information posted. The more the information is confirmed the more it is considered trustful. These indications for the confirmation of information generates a graph where the vertices represent users from WikiCrimes and the edges represent the indication of others to confirm the criminal fact registered. The graph represents a Social Network formed by the WikiCrimes users. It is fundamental for the attribution of reputation to the users when they are not qualified as certifier entities.

We can assume that an agent delegates a commitment to another agent when the first indicates the second to confirm a crime. The act of acceptance of the commitment by the agent indicates a relationship of trust between the agents. If the commitment broken by the second agent it will be penalized in reputation points.

Therefore the WIE will have a Social Network aspect related to it, as well. Agents can indicate other agents to confirm crimes. This indication specifies a social relationship among a group of agents. The agent will indicate another agent that it probably knows, and trust that this agent will confirm the crime that was registered. This relationship suggests the adoption of mechanisms to propagate trust in the social network built in the WIE. The Administrator and Certifier Entity roles have a very good
reputation to start with. This set of agents can serve as a starting point for the propagation of trust, starting from the Administrator and the Certifier Entities, then to the agents indicated by them, then to the indicated by the indicated, and so on, similar to the propagation of trust and distrust for demotion of web spam described in [Wu, B. 2006]. Another approach used is the propagation of trust to build reputation whenever there is a change in the reputation of the agent by means of action they realize in the WIE.

7. Final Remarks and Future Work

This paper demonstrates the suitability of our model for Institutional Environments to model and build open collaborative systems as open multi-agent systems such as the WikiCrimes system. We have adapted the model for the propagation of trust from [Wu, B. 2006] to our WikiCrimes Institutional Environment as part of our first prototype. With that we intend to make our model of Trust and Reputation stronger, and avoid the undesirable behavior in the WIE. This approach brings a high level of agent autonomy, which contributes to the openness of the institutional environment. The agents will be able to behave freely, without too many constraints or having to give too much information about themselves. This leads to a greater number of participants and at the same time identifies the bad behaving ones, so that they will not have the information posted by them confirmed, and therefore will lose useful life in the institutional environment.

At the moment we are refining the mechanisms to calculate the reputation of the agents in the WIE. A set of agents were built with different strategies of crime registration in an attempt to simulate bad behavior by those agents in the WIE. Algorithms for pattern recognition are being implemented to identify the most participative groups of agents in the registration of crimes in some geographic area to see if these groups could be characterized as a group of bad agents by the system [Furtado, V. et al. 2009].

References


