Adaptive Regulation of Open MAS: an Incentive Mechanism based on Modifications of the Environment*

(Extended Abstract)

Roberto Centeno
Centre for Intelligent Information Technology
University Rey Juan Carlos
Madrid, Spain
roberto.centeno@urjc.es

Holger Billhardt
Centre for Intelligent Information Technology
University Rey Juan Carlos
Madrid, Spain
holger.billhardt@urjc.es

ABSTRACT

The global objective of open multiagent systems might be in conflict with individual preferences of rational agents participating in such systems. Addressing this problem, we propose a mechanism able to attach incentives to agent actions such that the global utility of the system is improved. Such incentives are dynamically adjusted to each agent's preferences by using institutional agents called *incentivators*.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence - Multiagent Systems

General Terms

Algorithms

Keywords

Environments, organisations and institutions, self-organisation

1. INTRODUCTION

The main problem in Open MultiAgent Systems (OMAS) is to deal with situations in which the global objective of the system is in conflict with the individual objectives of its population of agents. Due to their open nature, such a population is usually unknown at design time. Thus, the task of assuring that agents behave according to the preferences of the system becomes even more complicated. The MAS community (e.g. [2], [1]) has dealt with this problem by endowing systems with organisational models based usually on normative mechanisms in charge of regulating agents' behaviour. However, those approaches have weaknesses due to: i) they are usually defined at design time, thus, they have less flexibility in certain unforeseen situations; ii) their population may still have a certain degree of freedom, which may lead to inefficiency evolutions of the system; and iii) the population could be not sensitive to the defined penalties/rewards established as consequence of norm violations.

Addressing the aforementioned problems, we propose to endow OMAS with an adaptive incentive mechanism able to induce agents to act in the desired way by modifying the consequences of their actions.

2. EFFECTIVE INCENTIVE MECHANISM

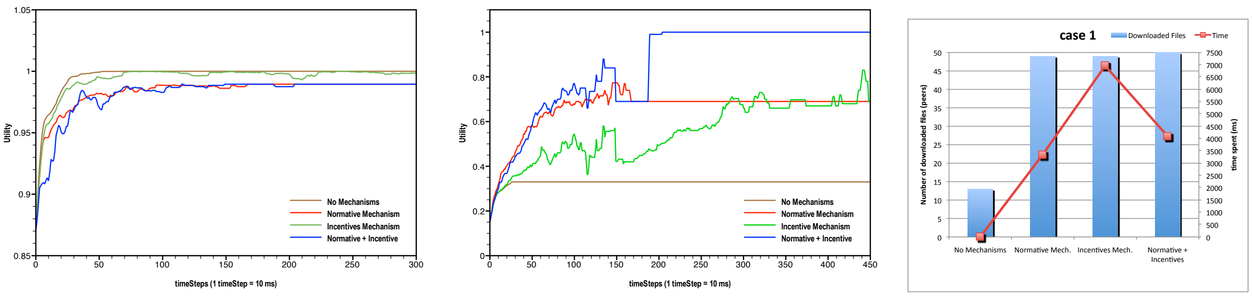
From the point of view of the designer of an OMAS, the problem consists of how to optimise the global utility of the system assuming that participants (rational agents) will try to optimise their own individual utilities. In order to do this, we focus on influencing agents' behaviour by means of incentive mechanisms[4]. We consider that incentives are modifications of the environment that have the aim to make a particular action more attractive than other alternatives, such that a rational agent would decide to take that action. Besides, an incentive mechanism is *effective* if its implementation implies an improvement of the utility of the system.

An incentive mechanism has to accomplish two tasks: i) to select the actions that should be promoted in order to improve the utility of the system; and ii) to establish the required changes so as to make the desired actions more attractive for agents. Both tasks are accomplished at runtime. The incentive mechanism is deployed as an infrastructure (similar to AMELI in Electronic Institutions[3]) endowed with institutional agents (*incentivators*). Each agent is assigned to an incentivator aiming to discover its preferences. Furthermore, incentivators can communicate with each other, allowing them to coordinate their actions.

In order to make actions more attractive, from an agent point of view, it is necessary to know in which attributes of the environment it is interested. Since in OMAS such preferences are unknown, they need to be discovered. We propose to use a non-intrusive approach where each incentivator discovers the preferences by observing its agent's behaviour in response to given incentives. The characteristics of the discovering process are: i) it is a learning process; ii) it is independent; and iii) the incentivator receives an immediate local reward. With this in mind, Q-learning with immediate rewards and ϵ-greedy action selection has been chosen. In each step, each incentivator selects the most promising attribute to modify and a value for this attribute, applies the changes, observes its agents reaction and modifies the q-values for attributes and values accordingly.

*Research supported by the projects "AT" CONSOLIDER CSD2007-0022 and TIN2009-13839-C03.

Cite as: Adaptive regulation of open MAS: an incentive mechanism based on online modifications of the environment (Extended Abstract), R. Centeno, H. Billhardt, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Tumer, Yolum, Sonenberg and Stone (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. 1243-1244.



(a) Agents utility (b) System utility (c) Downloaded files and time

Figure 1: Experimental results

The second task is to decide which actions should be promoted so as to improve the system's utility. Incentivators are endowed with a reinforcement multiagent cooperative learning algorithm (Q-learning combined with a gossip-based algorithm) so as to learn the desired joint actions in a cooperative way. In particular, they exchange information that allows to calculate a global reward for the learning process.

As case study we have chosen a p2p scenario where peers share a file by using a simplification of the BitTorrent protocol. We focus on the communication phase carried out to obtain each block belonging to a file. In this phase, a peer has to decide which neighbours will ask for the next block to download; and to which requests it will answer by uploading the requested block.

The systems' preferences have been captured by a multi-attribute utility function based on the following attributes: i) peers should download/upload as many blocks as possible; ii) the usage of the network should be as low as possible; and iii) the time spent on downloading files should be as short as possible. Peers have to pay a regular fee in order to connect to the network with a certain bandwidth. Besides, they have a file (partially or completely downloaded) they are sharing. Thus, peers' preferences are based on the bandwidth, fee, number of downloading/uploading blocks and time spent.

We compare our incentive mechanism with a standard normative system. The normative system is based on three norms that have been designed before knowing the population: **N1**: "It is prohibited to use more bandwidth than 85%"; **N2**: "A peer is obliged to upload a block when at least 25% of the bandwidth is available"; and **N3**: "It is prohibited to request a block to more than the 85% of neighbours". Norm violations – detected with a 100% of efficiency – are penalised with an increase on the fee in 5 units. Regarding the incentive mechanism, incentivators are authorized to modify the bandwidths and the fees.

We have specifically chosen a peer population that is sensitive to changes in the fee they are paying. Therefore, the designed norms will be quite effective for the given population of agents. Figure 1(a) plots the average utility obtained by all peers. Agents obtain the highest utility when there is no mechanism regulating the system, because nothing restricts their freedom. The second best performance is provided by our proposal due to agents may be incentivized by giving them a reduction on the fee. On the other hand, the normative system and a combination of both, normative and incentive, perform similarly. Figure 1(b) plots the utility of the system. As it was expected, the worst performance is when no regulation at all is working in the system. It improves when norms are working because with the chosen population the norms are effective. The incentive mechanism performs similar to the normative but it is slower due to the learning algorithms. The best performance is obtained when both mechanisms are combined. Finally, figure 1(c) shows the number of peers that are able to download the whole file. In the case of the normative and incentive systems 49 out of 50 peers download the whole file (spend more time when using incentives). With the combination of incentive and normative all peers (50) download the whole file, spending only slightly more time than in the normative system. We have also conducted experiments where the population is less sensitive to the defined penalties in norms (e.g., simulating "bad" norm design). In this case the incentive mechanism clearly outperforms the normative mechanism.

3. CONCLUSION

In this paper we propose an effective incentive mechanism that is able to induce desirable behaviour by providing incentives to agents. It is deployed by using an infrastructure based on institutional agents called *incentivators*. By means of Q-learning algorithms agents' preferences are discovered, by observing how agents react to modification in the environment. Moreover, incentivators learn – in a cooperative way – which joint action should be incentivized in order to increase the utility of the system. The proposed mechanism has been tested in a p2p file sharing scenario, showing that it is a valid alternative to standard normative systems.

4. REFERENCES

[1] V. Dignum, J. Vazquez-Salceda, and F. Dignum. OMNI: Introducing social structure, norms and ontologies into agent organizations. In *PROMAS'04*, volume LNCS 3346, pages 181–198, 2004.

[2] M. Esteva, J. Rodriguez, C. Sierra, P. Garcia, and J. Arcos. On the formal specification of electronic institutions. *Agent Mediated Electronic Commerce*, LNAI 1991:126–147, 2001.

[3] M. Esteva, B. Rosell, J. Rodríguez-Aguilar, and J. Arcos. AMELI: An agent-based middleware for electronic institutions. In *AAMAS'04*, volume 1, pages 236–243, 2004.

[4] R.Centeno, H.Billhardt, R.Hermoso, and S.Ossowski. Organising mas: A formal model based on organisational mechanisms. In *SAC'09*, pages 740–746, 2009.