Chapter 1

ENABLING OPEN AGENT INSTITUTIONS*

Juan A. Rodríguez-Aguilar*
Institut d'Investigació en Intel.ligència Artificial (IIIA)
Spanish Scientific Research Council (CSIC)

Carles Sierra
Institut d'Investigació en Intel.ligència Artificial (IIIA)
Spanish Scientific Research Council (CSIC)

Abstract In this paper we argue that open multi-agent systems can be effectively designed and implemented as *electronic institutions* composed of a vast amount of heterogeneous (human and software) agents playing different roles and interacting by means of speech acts. Thus taking inspiration from traditional human institutions, we offer a general agent-mediated computational model of institutions that serves to realise an actual agent-mediated electronic auction house where heterogeneous agents can trade.

Introduction

Up to date most of the work produced by multi-agent systems(MAS) research has focused on systems developed and enacted under centralised control. Thus, MAS researchers have bargained for well-behaved agents immersed in reliable infrastructures in relatively simple domains. Such assumptions are not valid when considering *open systems* ([3]) whose components are unknown beforehand, can change over time and can be both human and software agents developed by different parties. Examples of open agent systems include open electronic marketplaces and virtual supply chains, disaster recovery operations, collaborative design and international coalition military forces.

*This work has been partially funded by research grant FI-PG/96-8.490.
*On leave at iSOCO, Intelligent Software Components, S. A.

Although open systems have recently started to be considered by MAS researchers as one the most important application of multi-agent systems, their inherent issues (agent heterogeneity, reliability, accountability, legitimacy, societal change, etc.) have not been conveniently addressed yet. And then how to approach their design and construction? Although there has been a surge of interest in agent-oriented methodologies and modelling techniques in the last few years motivated and spurred by the first generation of agent developments ([11, 7, 8]), at present most agent applications lack a principled methodology underpinning their development, and so they are produced in an ad hoc fashion.

Another fundamental aspect is to opt for either a micro (agent-centered) view or a macro (organisation-centered) view of MAS. Although early work in DAI identified the advantages of organisational structuring as one of the main issues in order to cope with the complexity inherent to designing DAI systems (f.i. [2]) MAS research has traditionally kept an individualistic character, evolving patterned on a strong agent-centered flavour. And yet, there is an increasing interest in incorporating organisational concepts into MAS as well as in shifting from agent-centered to organisation-centered designs ([5, 1, 7]) that consider the organisation as a first-class citizen. Nonetheless, in general the introduction of social concepts into multi-agent systems has been undertaken in a rather informal way.

In this paper we adopt a macro perspective in order to effectively construct open multi-agent systems. Thus we argue on the need for deploying normative environments similar to those provided by human institutions following the pioneering work in ([5]). Institutions ([6]) represent the rules of the game in a society, including any (formal or informal) form of constraint that human beings devise to shape human interaction. They are the framework within which human interaction takes place, defining what individuals are forbidden and permitted and under what conditions. Furthermore, institutions are responsible for ascertaining violations and the severity of the punishment to be enacted. We uphold that open multi-agent systems can be successfully designed and implemented as institutionalised agent organisations (henceforth *electronic institutions*).

In Section 1 we present a case study of human institution in order to subsequently identify its components in Section 2. Next, in Section 3 we describe the two types of agents on which we found a computational model of electronic institution which successfully served to realise an actual agent-mediated electronic auction house. Finally, Section 4 contains some conclusions.

1. The Fish Market. An Actual-world Human Institution

As a starting point for the study of institutions we choose the fish market as a paradigm of traditional human institutions. The actual fish market can be described as a place where several *scenes* take place simultaneously, at different places, but with some causal continuity. Each scene involves various agents who at that moment perform well-defined functions. These agents are subject to the accepted market conventions, but they also have to adapt to whatever has happened and is happening at the auction house at that time. The principal scene is the auction itself, in which buyers bid for boxes of fish that are presented by an auctioneer who calls prices in descending order —the downward bidding protocol. However, before those boxes of fish may be sold, fishermen have to deliver the fish to the fish market (in the *sellers' admission scene*) and buyers need to register for the market (at the *buyers' admission scene*). Likewise, once a box of fish is sold, the buyer should take it away by passing through a *buyers' settlements scene*, while sellers may collect their payments at the *sellers' settlements scene* once their lot has been sold.

2. Institution Components

In order to engineer open agent multi-agent systems as electronic institutions we must firstly identify the core notions and components of electronic institutions, the computational counterpart of institutions, taking inspiration on the case study presented above. Thus our conception of electronic institution shall be founded on the following concepts:

Agents and Roles. Agents are the players in an electronic institution, interacting by the exchange of illocutions (speech acts), whereas roles are standardised patterns of behaviour. Any agent within an electronic institution is required to adopt some role(s). We fundamentally distinguish two classes of roles: *institutional*, and *non-institutional*.

Dialogical framework. In a dialogical institution, agents interact through illocutions. Institutions establish the ontology and the common language for communication and knowledge representation, which are bundled in what we call dialogical framework. By sharing a dialogical framework, we enable heterogeneous agents to exchange knowledge with other agents.

Scene. Interactions between agents are articulated through agent group meetings, which we call *scenes*, with a well-defined communication protocol. We consider the protocol of a scene to be the specification of the possible dialogues agents may have to articulate a multi-agent activity. A scene defines a role-based framework of interaction for agents. A distinguishing feature of scenes is that agents may join in or leave during the activity.

Performative structure. Scenes can be connected, composing a network of scenes, the so-called *performative structure*, which captures the existing relationships among scenes. A performative structure specifies how agents can legally move from scene to scene by defining both the pre-conditions to join in and leave scenes. Considering the fish market, while some activities like the admission of buyers and sellers are completely independent, others are tightly related. For instance, a buyer cannot bid for any good unless he has previously and successfully set up a credit line.

Normative Rules. Agent actions in the context of an institution have consequences, usually in the shape of compromises which impose obligations or restrictions on dialogic actions of agents in the scenes wherein they are acting or will be acting in the future. For instance, after winning a bidding round the bidder is committed to subsequently pay for the acquired good. Obligations and prohibitions are captured by means of normative rules.

Based on the institution components introduced above, in ([8]) we offer a formal specification of electronic institutions that founds the computational model presented in Section 3.

3. Agent-mediated Institutions

The workings of an electronic institution can be fully realised by means of the articulation of two types of agents: *institutional agents* and *interagents*. Institutional agents are those to which the institution delegates its services, whereas interagents are a special type of facilitators that mediate all the interactions of external agents within an electronic institution and enforce institutional rules. Our agent-mediated computational model (thoroughly detailed in [8].) has proven its usefulness in the development of FM96.5, the computational counterpart of the fish market ([10]), which served as the basis for the subsequent development of FM, an agent-mediated test-bed for auction-based markets([9]).

3.1 Institutional Agents

An institution delegates part of its tasks to agents adopting institutional roles (in the fish market the auctioneer is responsible for auctioning goods, the sellers' admitter for registering goods, and the accountant for the accounts' bookkeeping). We refer to this type of agents as institutional agents. An institutional agent can possibly adopt multiple institutional roles. In order to fully specify an institutional role we must specify its *life-cycle* within an institution in terms of its responsibilities along with the policy of responsibilities' management.

More concretely, we specify an institutional role's life-cycle as a regular expression built by combining the following operations: $x.y$ (*x* followed by *y*), x^* (*x* occurs 0 or more times), $x||y$ (*x* and *y* interleaved), $x|y$ (*x* or *y* occurs), x^+ (*x* occurs 1 or more times), $[x]$ (*x* is optional); where x and y stand for scene (activity) names. Table 1.1 contains the specification of the *buyer_admitter*, *auctioneer* and *seller_accountant* roles in FM96.5, the computational counterpart of the fish market. For instance, an institutional agent playing the *buyer_admitter* role must firstly enter at the *registry* scene with the *boss* of the market. Next, it is expected to meet other institutional agents (auctioneer, buyers' accountant, sellers' admitter and sellers' accountant) at the *opening* scene. Afterwards it can start processing buyers' requests for admission.

buyer_admitter	$= registry.((opening.(buyer\_admission)^*)\|\|closing)$
auctioneer	$= registry.((opening.(request\_goods.(auction\|\|credit\_line))^*)\|\|closing)$
seller_accountant	$= registry.(opening.(good\_adjudication^*\|\|seller\_settlements^*)\|\|closing)$

Table 1.1. Institutional agents' responsibilities specification.

Institutional agents might be required to comply with several responsibilities at the same time. In such a case, an institutional agent must know how to prioritise (schedule) simultaneous responsibilities. For this purpose, responsibilities are ranked according to their relevance. As an example, the responsibilities for the auctioneer are ranked as follows: *(registry,High) (opening,High) (closing,High) (request_goods,Medium) (credit_line,Medium) (auction,Low)*; where *High*, *Medium* and *Low* denote different priority degrees..

And yet there remains the matter of deciding how to behave within each scene in which an institutional agent will get involved. When participating in a scene, at some states an institutional agent will be expected to *act* by uttering an illocution as a result of an inner decision-making process. For instance, an auctioneer must know how to select the winner of a bidding round, a buyers' admitter must decide whether to admit a buyer or not, and a sellers' admitter must know how to tag the incoming goods to be put at auction. These inner activities yield illocutions to be uttered by the institutional agent. Since an institutional agent must know which method to fire at those scene states at which it is expected to act, his *behaviour specification* is provided as a collection of methods to be fired at particular states of the scene. For instance, Figure 1.1 contains a specification of an auction scene protocol (the graph nodes denote scene states connected by arcs labeled by illocution schemes. Transitions occur when illocutions uttered by agents match illocution schemes.). The auctioneer is instructed to run the *declareWinner* method at ω_7.

In ([8]) we propose a general model of institutional agent in order to ease development. Thus, the very same institutional agent model (architecture) can be employed to deploy several institutional agents playing different roles.

3.2 Interagents

Interagents ([4]) constitute the sole and exclusive means through which agents interact with the rest of agents within the institution. They become the only channel through which illocutions can pass between external agents and institutional agents. Notice that interagents are all owned by the institution

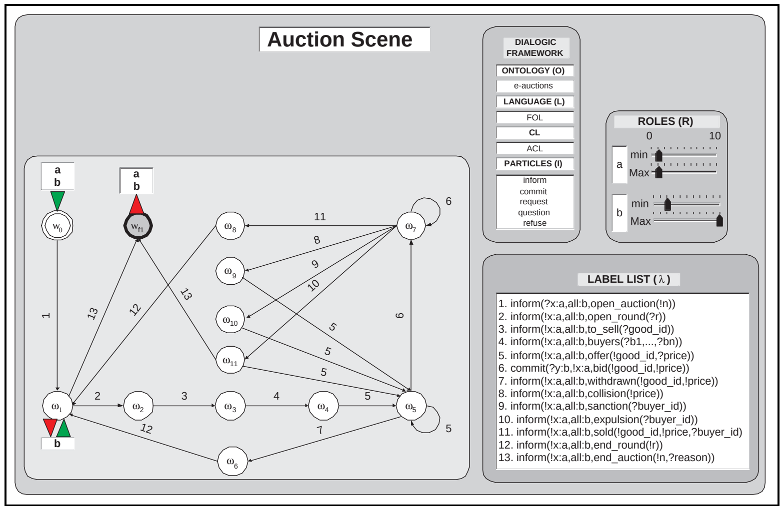

Figure 1.1. Graphical Specification of an Auction Scene

but used by external agents. The mediation of interagents is key in order to guarantee: the legal exchange of illocutions among agents within scenes; the sound transition of external agents from activity to activity within the institution's performative structure; the enforcement of institutional rules; and the accountability of external agents' interactions.

One of the fundamental tasks of interagents is to ensure the legal exchange of illocutions among the agents taking part in some scene: what can be said, to whom and when. For this purpose, interagents employ *conversation protocols*(CP) ([4]). CPs define coordination patterns that constrain the sequencing of illocutions within a scene and allow to store, and subsequently retrieve, the contextual information (illocutions previously sent or heard) of ongoing scenes. We can think of CPs as scenes extended with the necessary actions to keep contextual information. Based on contextual information, when receiving some illocution from an external agent to be transmitted, an interagent can assess whether the illocution is legal or else whether it must be rejected or some enforcement rule activated.

Consider the auction scene. A buyer agent receives the prices called by the auctioneer through his interagent, which keeps track of the latest price called. When the buyer agent submits a bid, his interagent collects it and verifies whether the buyer is bidding for the latest offer price. If so, the interagent posts the bid to the auctioneer, otherwise it's rejected. Once the bid has been submitted, the buyer is not allowed to re-bid. If he tries, their bids are disallowed, and if he compulsively tries his interagent unplugs him from the institution.

Then his interagent autonomously follow the required procedures to log the buyer out from the auction house.

Interagents also constrain external agents' behaviour in their transition between scenes. Figure 1.2 depicts the specification of the performative structure projection for buyer agents in FM96.5, the computational counterpart of the fish market. If some buyer requests his interagent for leaving the institution after making some acquisitions in the auction scene, his interagent will refuse the request because the agent has pending obligations: the payment of the acquired goods, as stated by the institutional normative rules.

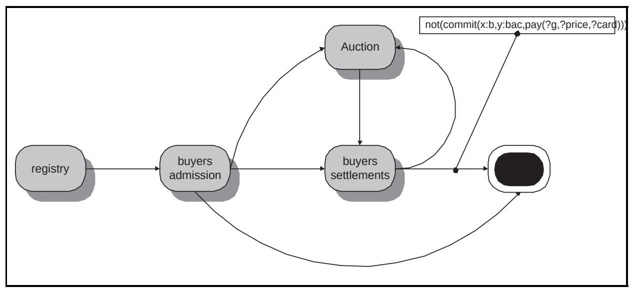

Figure 1.2. Performative structure projection for buying agents.

In general, based on external agents' actions, the facts deriving from their participation in scenes and the institutional normative rules, interagents are capable of determining which obligations and prohibitions to trigger.

Finally, interagents handle transparently to external agents their incorporation into ongoing scenes, their exit from ongoing scenes, their migration between scenes, and the joint creation of new scenes with other agents by means of their coordinated activity with institutional agents, as fully accounted by the computational model detailed in ([8]).

4. Conclusions

Organisational and social concepts can enormously help reduce the complexity inherent to the deployment of open multi-agent systems. In particular, institutions are tremendously valuable to help solve the many inherent issues to open multi-agent systems. The conception of open multi-agent systems as electronic institutions lead us to a general computational model based on two types of agents: institutional agents and interagents. Although our computational model proved to be valuable in the development of the computational counterpart of the fish market, we claim that such a computational model is general enough to found the development of other agent institutions.

References

[1] Ferber, J. and Gutknetch, O. (1998). A meta-model for the analysis of organizations in multi-agent systems. In *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98)*, pages 128–135.

[2] Gasser, L., Braganza, C., and Herman, N. (1987). *Distributed Artificial Intelligence*, chapter MACE: A flexible test-bed for distributed AI research, pages 119–152. Pitman Publishers.

[3] Hewitt, C. (1986). Offices are open systems. *ACM Transactions of Office Automation Systems*, 4(3):271–287.

[4] Martín, F. J., Plaza, E., and Rodríguez-Aguilar, J. A. (2000). An infrastructure for agent-based systems: An interagent approach. *International Journal of Intelligent Systems*, 15(3):217–240.

[5] Noriega, P. (1997). *Agent-Mediated Auctions: The Fishmarket Metaphor*. Number 8 in IIIA Monograph Series. Institut d'Investigació en Intel.ligència Artificial (IIIA). PhD Thesis.

[6] North, D. (1990). *Institutions, Institutional Change and Economics Perfomance*. Cambridge U. P.

[7] Parunak, H. V. D. and Odell, J. (2001). Representing social structures in uml. In *Proceedings of the Agent-Oriented Software Engineering Workshop*. Held at the Agents 2001 Conference.

[8] Rodríguez-Aguilar, J. A. (2001). *On the Design and Construction of Agent-mediated Institutions*. PhD thesis, Autonomous University of Barcelona.

[9] Rodríguez-Aguilar, J. A., Martín, F. J., Noriega, P., Garcia, P., and Sierra, C. (1998). Competitive scenarios for heterogeneous trading agents. In *Proceedings of the Second International Conference on Autonomous Agents (AGENTS'98)*, pages 293–300.

[10] Rodríguez-Aguilar, J. A., Noriega, P., Sierra, C., and Padget, J. (1997). Fm96.5 a java-based electronic auction house. In *Second International Conference on The Practical Application of Intelligent Agents and Multi-Agent Technology(PAAM'97)*, pages 207–224.

[11] Wooldridge, M., Jennings, N. R., and Kinny, D. (1999). A methodology for agent-oriented analysis and design. In *Proceedings of the Third International Conference on Autonomous Agents (AGENTS'99)*.