V. Dignum, J-J. Meyer, H. Weigand, F. Dignum. An Organizational-oriented Model for Agent Societies. In: Proceedings of International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'02), at AAMAS, Bologna, Italy, 16 July 2002.

An Organization-oriented Model for Agent Societies

Virginia Dignum [1,2], John-Jules Meyer[2], Hans Weigand[3], Frank Dignum[2]

[1]Achmea, The Netherlands
virginia.dignum@achmea.nl

[2]University Utrecht, The Netherlands
{virginia, jj, dignum}@cs.uu.nl

[3]Infolab, Tilburg University, The Netherlands
weigand@kub.nl

Abstract. We present a conceptual framework for agent societies that incorporates the specification of global organizational characteristics with individual aims and capabilities. The framework consists of three interrelated models each describing different aspects of the society. The organizational model describes agent societies in terms of roles, constraints and interactions rules. The social model populates the society with actual agents that realize the objectives of the society by enacting role(s) described in the organizational model. Finally, the interaction model describes the interaction commitments between those agents. The interdependencies between models are described by means of contracts. Contract rules specify commitments between agents and society concerning role enactment, and commitments between agents concerning interaction.

1. INTRODUCTION

In general, an organization can be defined as a specific solution created by more or less autonomous actors to achieve common goals. Organizational structure is often viewed as a means to manage complex dynamics in (human) societies. This implies that approaches to organizational modeling must incorporate both the structural and the dynamic aspects of such society. In structured organizations, social interaction emerges from a set of (negotiated) social norms and is regulated by mechanisms of social control.

Software agents are autonomous entities with reasoning and communicative capabilities, and therefore utmost suitable to implement, simulate or represent real-life entities displaying the same autonomy. The power of the agent paradigm lies in its capability of modeling complex, open systems in a way that reflects natural interactions and relationships in domain (cf. [2, 17]). Multi-agent systems, or agent societies, represent the interactions between agents and are as such the virtual counterpart of real-life societies and organizations. Agents take specific roles in the system and interact with others as a means to accomplish their roles. This perspective makes the design of the system less complex since it reduces the conceptual distance between the system and the real-world application it models.

The development of multi-agent systems calls for models, languages and methodologies to represent interaction, roles and other concepts that characterize societies. Such modeling primitives are usually not provided by single-agent languages. Moreover, traditional multi-agent models often assume an individualistic perspective in which agents are taken as autonomous entities pursuing their own individual goals based on their own beliefs and capabilities. In this perspective, global behavior emerges from individual interactions and cannot easily be managed or specified externally. However, in business environments the behavior of the global system must be taken in account and structural characteristics of the domain have to be incorporated. That is, society design must consider organizational characteristics such as stability over time, some level of predictability, and commitment to aims and strategies, etc. These are the rules and global objectives that govern the activity of an enterprise, organization or nation. Such characteristics are often specified top-down and imposed onto the participants. Global characteristics are external to each individual agent and independent from the own goals and behavior of the agent, and therefore cannot easily be incorporated in a multi-agent architecture that starts from an individualistic perspective. On the other hand, the volatility of business environments stresses the need for models and systems that accommodate changes required by new or different organizational aims with a minimum impact on the already existing services. From the organizational point of view this creates a need to check conformance of the actual behavior of the society to the behavior desired by the organization.

We assume furthermore that individual agents participating in an organizational agent society are designed outside the scope of the society design and are heterogeneous in architecture and modeling language. Therefore, models for agent societies cannot be based on assumptions on the internal components, beliefs or actions of individual agents. From the above considerations, we have identified the following requirements for design formalisms of agent societies [8]:

- Support and direct the analysis of the organizational structure of the domain in order to determine society norms and facilitation roles
- Explicitly specify the organizational and normative elements of the society since an open society cannot rely on its embedding in the intentions, desires and beliefs of each agent.
- Include formalisms for the description, construction and control of the organizational and normative elements of a society (roles, norms and goals) instead of agent beliefs and states.
- Provide mechanisms to describe the environment of the society and the interactions between agents and the society, and to formalize the expected outcome of roles in order to verify the overall animation of the society.
- Provide methods and tools to verify whether the design of an agent society satisfies its design requirements and objectives.
- Provide building directives concerning the communication capability and ability to conform to the expected role behavior of agents in the society.

In this paper, we present a conceptual framework for agent societies that distinguishes between the mechanisms though which the structure and global behavior of the model is described and coordinated, and the service-providers (agents) that populate the model. The framework represents interaction between agents in a way that (1) is independent of the internal design of the agents, and (2) is able to

integrate organizational characteristics and demands with the agent's own goals in a dynamic way that preserves the autonomy of the participating agents. The model makes use of the concept of contract as a means to conjugate the top-down specification of organizational structures with the autonomy of participating agents.

The remaining of this paper is organized as follows: Section 2 introduces the concept of agent societies and discusses the organizational and individual perspectives on agent societies. Our framework for agent societies is described in more detail in section 3. A language for specification of contracts is introduced in section 4 . Section 5 discusses related work and finally, in section 6 we draw some conclusions and indicate directions for further work.

2.   AGENT SOCIETIES

The role of any society is to allow its members to coexist in a shared environment and pursue their respective goals in the presence and/or in co-operation with others. A collection of agents interacting with each other for some purpose and/or inhabiting a specific locality can be regarded as a society. Societies usually specify mechanisms of social order in terms of common norms and rules that members are expected to adhere to [5]. Main aspects in the definition of society are purpose, structure and norms. The purpose, or global goals of the society, is an high level description of the aims of the society, as desired by the society 'owners'. The structure of the society can described by roles, interactions and communication language. Roles represent the different entities or activities that are needed to fulfil the society purpose. The global goals of the society are therefore, the starting point to specify the goals and actions assigned to roles. Social norms describe the desired behavior of members (from the point of view of the society design) and what sanctions must be applied in case of undesirable action. Norms are often established and enforced by institutions that have a legal standing and thus lend legitimacy and security to members of the society.

Societies can be considered from different perspectives. We distinguish between the collective, or **organizational** perspective, and the individualist, or **agent** perspective. Means of structuring the society and some level of global management and direction are needed in order to reduce the system's complexity, increase the system's efficiency, and more accurately model the problem being tackled. Depending on the view taken, the objectives and structure of the society and its effect on the society inhabitants will seem different. In our approach development of an agent society starts from an organizational perspective. This has in our opinion two strong advantages: in one hand, by taking an organizational view on the problem domain we are bridging the distance between agent developers and domain practitioners as the approach taken directly relates to the organizational perception of the problem. On the other hand, it has been pointed out by Wooldridge and Jennings that it is a common misconception to think that multi-agent systems require no real structuring [23]. That is, agents will not interact with each other just because they have the possibility to do so or because they happen to share the same environment [5]. However, individual agents seeking access to the society have their own individual perspectives. This means that frameworks for agent societies need to provide a bridge between both perspectives by making explicit the commitments and

obligations of both parties. Both perspectives are illustrated in figure 1 and described in more detail in the following subsections.
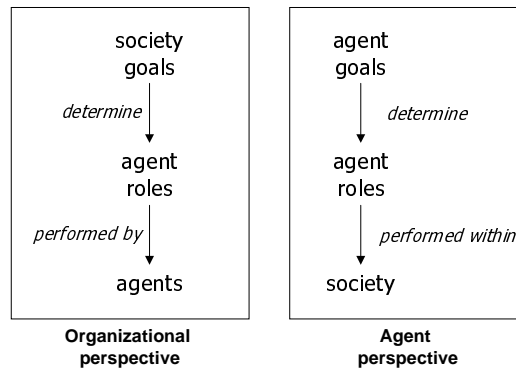

Fig. 1. Collective and individual perspective on agent roles

2.1.   Agent perspective

From an individual, or agent perspective, societies are viewed as the environment in which individual action takes place. This view considers MAS as mere aggregations of agents (autonomous entities with individualistic behavior) interacting together. Coordination emerges from patterns of actions of agents as a result of interaction. Agent goals determine agent role(s) in a society. Societies can impose interaction and constrains the agent's actions. Usually, agents are not particularly concerned with the overall objectives of the society or of the specific role being played, as long as its own, individual, objectives get achieved by performing that role. Each agent pursues its own goals according to its own internal state and values (internal architecture, e.g. BDI). In order to reach those goals it will possibly need to play certain roles in a society (or in several societies). However, it cannot be assumed that the agent is aware of or concerned with the fulfillment of the society objectives or are even aware of those objectives. From this perspective, the agent view is that individual agents happen to be together in an environment and thus may interact and/or collaborate with others in order to achieve their own goals.

Existing architectures, behavioral strategies and models for group formation often assume this individualist perspective in which agents are taken as autonomous entities pursuing their own individual goals based on their own beliefs and capabilities. In this perspective global behavior emerges from individual interactions and cannot easily be managed or specified externally. However, in business environments the behavior of the global system needs to be considered and designed collectively from a top-down perspective. That is, the multi-agent model must be able to consider the global characteristics of the organization, such as stability over time, some level of predictability, clear commitment to aims and strategies.

2.2. Organizational perspective

From an organizational perspective, social design is concerned with the definition of the aims of the society as a whole as well as the roles and responsibilities for participants. Global aims are leading and roles are specified in such a way that these goals can be realized. From an organizational perspective the main function of individual agent is the enactment of a role. That is, society goals determine the agent roles, as well as the norms and interaction rules holding in the society. Agents are actors that perform role(s) described by the society design. However, the agent's own capabilities and aims determine the specific way an agent enacts its role(s). From this perspective, the society is not concerned about which individual agent will actually play a specific role as long it gets performed. In [11] we describe how society goals can be used to determine specific roles based on the identification of the coordination mechanisms and stakeholders in the domain.

From the society perspective agents are seen as black boxes which internal structure is not known. This implies that the society cannot make assumptions or demands on the internal capabilities of an agent. Issues such as efficiency in how to solve a local problem are often not a concern of the society in itself, but of the agent assigned to solve it. Verification of society activity relies therefore on the explicit specification of the agreements made by the agents concerning their performance. An agent will enter an agreement with a society for the performance of a certain role, and as such assume the responsibility to perform the actions associated with the role and achieve the role's goals. How and whether the agent does it, is often not the concern of the society, except that the society can impose sanctions and penalties when behavior does not follow the commitments. However, the society can indicate in the preconditions of the role that the problem should be solved in one or the other way. From the collectivist approach (society view) roles have a clear and well-defined reason for existing in the system and will enable the achievement of the global goals of the society. The society is not concerned about who performs the role and why, as long as the role is being played and the services (objectives) associated with the role are being offered.

3. CONCEPTUAL MODEL FOR AGENT SOCIETIES

In our work, societies are viewed as a structural relationship between a collection of agents, that is, as a framework for the activity and interaction of individual agents. We describe societies in terms of their structure, without being concerned with the way individual agents are designed. In order to achieve its objectives, a society starts from the assumption that actors performing roles behave according to the social norms holding. Both roles and norms are specified by the society design. Which agent will in a specific situation (actual state of the society at a given moment) perform the role is not of importance to the society as a whole. The agent must however (1) fulfil the preconditions specified by the role, and (2) commit itself to achieve the goals corresponding to the role. Therefore, the society needs to describe its roles to agents, check the execution of the role and maintain and enforce a list of commitments (who does what in which conditions).

Conceptually, different types of roles can be identified in a society. We distinguish between social, or **facilitation** roles, that is roles needed in order to keep the society going, and **operational** roles, which will provide the actual objectives of the society [11]. Facilitation roles are usually played by mutually trusted agents, whereas the type of society organization will determine how trust between agents playing operational roles is realized. Organization theory shows that organizations with different objectives exhibit different requirements for coordination. Coordination models (market, hierarchy and network) are determined by transaction costs and reflect the balance between organizational objectives and activities. For example, the market model fits well in an exchange situation whereas the hierarchical model is better suited for production environment. Coordination models describe issues such as communication forms, desired social order and co-operation possibilities between partners. In [10] we argued that trust and assignment of facilitation roles to agents depend on the coordination model chosen and a methodology for analysis and design of agent societies based on different coordination types is presented. Standard society types as market, hierarchy and network, can be used as starting point for development and can be extended where needed and determine the basic norms and facilitation roles necessary for the society. That is, reusable frameworks for the facilitation layer of agent societies can be developed based on the coordination characteristics of the application domain. Such frameworks are then extended with a domain-specific operational layer.

The framework integrates a top-down specification of society objectives and global structure, with a dynamic fulfillment of roles and interactions bet and separates the description of the structure and global behavior of the domain from the specification of the individual entities (service-providers) that populate the domain. This separation provides several advantages to our framework above traditional MAS models. On one hand, coordination and interaction in MAS are usually described in the context of the actions and mental states of individual agents [14]. In open societies such approach is not possible because agents are developed independently from the society and there is therefore no knowledge about the internal architecture of agents nor possibilities to directly control or guide it. Furthermore, conceptual modeling of agent societies (based on the social interactions) requires that interaction between agents be described at a higher, more abstract level, that is, in terms of roles and institutional rules. On the other hand, society models designed from an organizational perspective, reflect the desired behavior of an agent society, as determined by the society 'owners'. However, once 'real' agents populate the society, their own goals and behavior will affect the overall society behavior, that is, such social order as envisioned by the society designer is in reality a conceptual, fictive behavior. Castelfranchi distinguishes between social order, the non-accidental, non-chaotic pattern of interaction in a given system of interacting agents and social control, agent action aimed at enforcing the conformity of behavior of other agents to some social norm [3]. He argues that due to the autonomous behavior of agents, social control is not enough to deal with the challenge of social order, but agent societies must be able to cope with unintended, emergent behavior of its members.

Our society framework provides different, integrated models to specify desired and actual behavior of the society. Central to the framework is the **organizational model** that describes the structure and global characteristics of a domain from an organizational perspective from the premise that it is the society goals that determine

agent roles and interaction norms. The organizational model is based on the analysis of the domain in terms of the coordination and normative elements and describes the expected behavior of the society. The framework does not specify the internal architecture of individual agents. Active entities are described as roles specified in terms of externally perceived actions and behavior. Other components of the model are constraints, interaction rules, and communicative and ontological frameworks.

We assume that individual agents are designed independently from the society to model the goals and capabilities of a given entity[1]. Individual agents are seen as actors that perform role(s) described by the society design and, as such interact with others as a means to accomplish their goals. In order to realize their own goals, individual agents will join the society as enactors of role(s) described in the organizational model. This means that several populations are possible for each organizational model. Agent populations of the organizational model are described in the **social model** in terms of commitments regulating the enactment of roles by individual agents. In the framework, agents are seen as autonomous communicative entities that will perform the society role(s) according to its own internal aims and architecture. Because the society designer does not control agent design and behavior the actual behavior of the society instance might differ from the intended behavior. The only means the society designer has for enforcing the intended behavior is by norms, rules and sanctions. That is, when an agent applies and is accepted for a role, it will commit itself to the realization of the role goals and it will function within the society according to the constraints applicable to its role(s). These commitments are specified as social contracts that can be compared to labor contracts between employees and companies. The society can sanction undesirable (wrong) behavior as a means to control how an agent will do its 'job'.

Finally, interaction between agents populating a society are described in the **interaction model** by means of interaction contracts. This model accounts for the actual (emergent) behavior of the society at a given moment. Interaction agreements between agents are described in interaction contracts. Usually interaction contracts will 'follow' the intended interaction possibilities specified in the organizational model. However, because of the autonomous behavior of agents, the interaction model must be able to accommodate other interaction contracts describing new, emergent, interaction paths.

These models are described in more detail in the following sections. Figure 2 depicts the interrelation between the different models. In the organizational model the structure of the society is described in terms of roles and interaction patterns between roles. Those roles are, in the social model, assigned to specific agents. This is a many-to-many relationship, that is, it is possible that one agent plays more than one role and that one role is played by more than one agent. Finally, the interaction model describes the actual interactions between agents playing roles, which, depending on the terms of the social contracts between the agents and the organization, may or may not be restricted to the structural interaction patterns specified in the organizational model.

---

[1] In [4] we present preliminary work on how to map the social aspects of an agent organization into the description of the behavior of the individual agents using as example agents developed using the 3APL language.
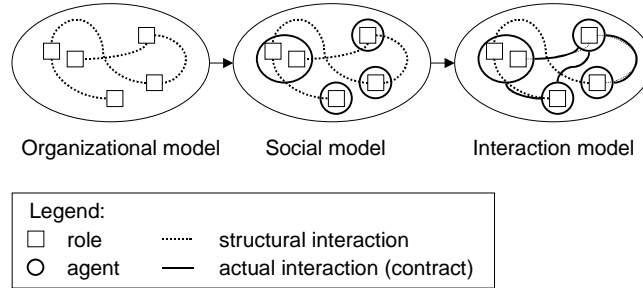
Organizational model   Social model   Interaction model

Legend:
- □ role   ········ structural interaction
- ○ agent   ——— actual interaction (contract)

Fig. 2. Organizational framework for agent societies

3.1. Organizational model

The organizational model specifies the structure of an agent society in terms of externally observed behavior and components. We define an **organizational model** as a tuple **OM** = (R, CF, I, N), where R is the set of role descriptions, CF is the communicative framework, I is the set of scripts for interaction scene and N is the set of society norms or rules. The elements of **OM** can be referred to by:

$$roles(OM) = R$$
$$communication\_framework(OM) = CF$$
$$interactions(OM) = I$$
$$norms(OM) = N$$

Global society goals are not part of the organization model but form the background to the society definition and are represented by the goals of the roles. These global goals are specified in terms of roles that correspond to the different stakeholders in the domain. That is, overall goals of a society can be represented as an hierarchy of sub-goals that correspond to the goals of the different roles. Furthermore, the organizational model is split into two parts: facilitation and operation. The facilitation layer provides the backbone of the society and consists of institutional agent roles, which are designed to enforce the social behavior of agents in the society and assure the global activity of the society. The operational layer models the overall objectives and intended action of the society and consists of domain related roles. The operational layer is always domain and application dependent whereas the facilitation layer depends on the cooperation characteristics of the environment and can be reused across domains. Typically, facilitation issues are suitable for outsourcing to an external party.

We illustrate the different components of a society using the example of a trading society. The overall goal of this society is to generate transactions. This goal can be split into a facilitation component that aims at the regulation of those transactions, and an operational part where transactions are generated. Typically, domain related and operational stakeholders in such a society are sellers and buyers, which exhibit autonomous behavior in the society. These will be specified as roles in the operational component of the model. Finally the activity of the facilitation layer can be described in terms of a registrar role that regulates the participants and a market master role that

regulates the transactions and supports the matching between sellers and buyers. This hierarchy of goals is depicted in figure 3. The roles at the lowest level can be described as follows. The Seller represents an entity that wants to exchange its goods for money; the Buyer represents an entity that wants to exchange its money for goods; the Market Master takes care of introducing potential sellers to potential buyers; and the Registrar keeps track of who are the sellers and buyers at each moment.
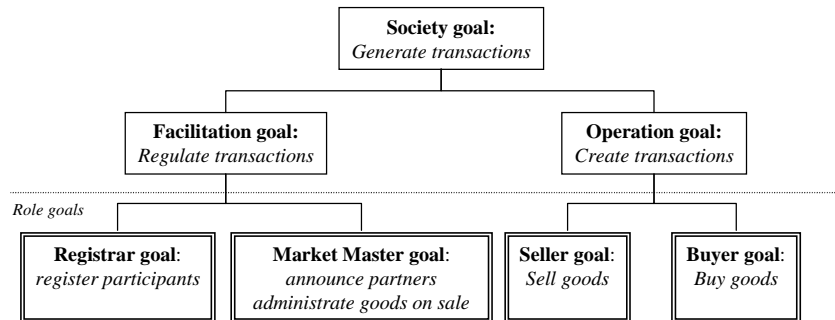

Fig. 3. Goal hierarchy of a trading society

3.1.1. Roles

A role is the abstract representation of a policy, service or function. Role descriptions in the organizational model identify necessary activities and services necessary to achieve society goals. Furthermore, roles must describe the necessary capabilities that must be enacted by any agent pretending to play that role. As in human societies, roles are organized into a role-relationship network. A role is defined as a triple $r = (G, N, R)$, where G, N, and R stand respectively for the goals, the norms, and the interaction rules that are associated with the role. Relations between roles determine duties for agents playing the role. The role(s) it plays at any given moment determine the current actions of an agent, and will influence the possibilities of further action of the agent in the society. We identify the following relationships between roles in a society:

Implication: $r > r'$ means that an agent playing r also plays role r' (e.g.: car-seller $>$ seller). As a consequence, the role r has to fulfill all the duties of role r'.

Conflict: $r \otimes r'$ means that an agent playing r cannot play r', at the same time. (e.g.: banker $\otimes$ seller). Defines separation of duties for agents.

Transformation: $r \Rightarrow r'$ means that an agent playing r may eventually be enabled to play r', or 'r will become r'. (e.g.: buyer $\Rightarrow$ preferred-client). Defines possible transformations of duties for agents.

3.1.2. Communicative framework

A consequence of assuming that agents are heterogeneous and built using different architectures and languages, is that agents will be participating that are endowed with

different internal languages and ontologies. Therefore, models for agent societies must provide means to relate different agent languages and ontologies in order to allow agents to communicate successfully. The organizational model must also provide a description of the illocutions, or communication primitives, that hold in the society. As postulated in speech act theory [22] and used in most agent communication languages (e.g. KQML [15]), agent illocutions are not just prepositions that may be true or false, but speech acts (that is, assertions, commands, requests, suggestions, promises, threats, etc.) that may succeed or fail. Necessary for communication are a common language to represent concepts and relationships in the domain (an ontology), a common language for communication and means to represent beliefs.

Based on [13], we propose the following definition of ontological framework that identifies the communication language, the domain representation language and the ontology used in the society. A **communicative framework** is a tuple CF = (O, IL, ACL, DR) where O is the domain ontology (that is, the vocabulary used and understood within the society), IL the set of illocutions, ACL is the agent communication language (intentions and vocabulary of utterance) and DR is the language for representation of domain content.

For example, the following is the abstract representation of a possible communication message (ignoring syntactic aspects determined by ACL):

```
accept(seller, buyer, bid(€500, bicycle))
```

meaning that the actor `seller` tells the actor `buyer` (that is, the agents that fulfil those roles), that it agrees with `buyer`'s bid of €500 for `bicycle`. The illocution `accept` is an element of IL and the expression `bid(€500, bicycle)` must be defined in DR. Concepts such as `bid`, €500 and `bicycle` must be defined in ontology O.

3.1.3. Interaction structures

The possible actions of a role determine interactions with other agents. These interactions are articulated into (organizational) interaction patterns between agents that establish the possible dialogues and interchanges between agents. Such patterns are designed in such a way that the society aims are achieved. Examples are the negotiation dialogue needed to reach an agreed contract, or the workflow representing a sales process. We define an **interaction structure** as a pair I = (S, T) where S is a set of scenes (represented by interaction scripts) and relation T: S×S gives the scene transformation matrix. For s, s' ∈ S, (s, s') ∈ T, means that s' can be reached from s.

The scene transformation induces a partial ordering T^* on the set of scenes. Scene transformation T has consequences for *roles(OM)* in the sense that actors will follow different paths according to their roles. Furthermore, the performance of an actor in a given scene will have consequences to the roles that actors will play in other states. For example, the interaction structure of the trading society is depicted in Fig. 4. Note that the graphical representation is very similar to (and in fact inspired by) the work of Sierra [13]. However, an important difference is that in their work interaction is fixed in protocols, while our framework uses landmarks that indicate boundaries for interaction without completely fixing it. Our approach, based on the principles of

refinement and collaboration autonomy, enables agents to influence the specific ways interaction will actually occur.

Informally this interaction structure says that participants must start by registering with the Registrar, after which Sellers can register their goods with the Marketmaster which will then announce the goods. Buyers register their requests with the Marketmaster. Then follows the negotiation of terms and conditions for exchange. Finally the exchange can be settled. (This is a very simplistic example, meant purely as an illustration for the representation of Interaction Structures; no realistic trade procedure is meant to be portrayed here.)
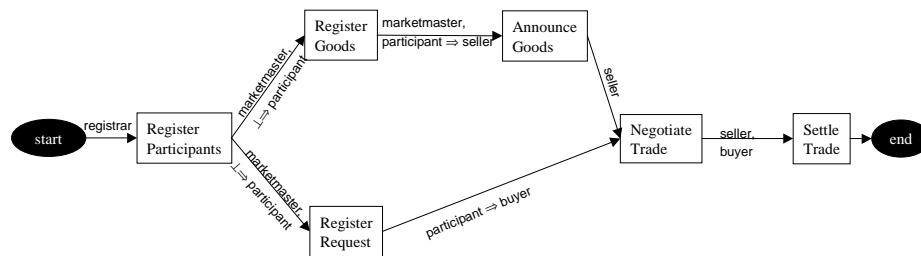

Fig. 4. Interaction Structure for 'Trading Society'

An **interaction script** describes a scenario of activity, that is, how roles interact and evolve in the context of a scene from the point of view of the society. Interaction scripts serve as a blueprint for the actual interactions between actors (the interaction play, that will be specified in the interaction model). Landmarks are logical expressions that describe the characteristics (for instance, goals and action plans) of the scene. The level of specification of landmarks determines the degree of freedom the actors have about their performance. The more landmarks are given, the less the actor can decide for itself. Norms are deontic expressions that regulate the activity of actors. An interaction script, $IS \in S$, is defined by $IS = (P, CF', L)$ where:

- P is the set of participating roles, $P \subseteq roles(OM)$,
- CF' is the communicative framework used in the scene, possibly a specialization of *comm-framework* (OM),
- L are the landmarks describing the interaction (propositions that hold for the scene, including the norms)

Landmarks can be more specific than just goals; for example, the organizational design can make explicit to a certain extent how a certain interaction are supposed to be done. Decisions on how to further realize an interaction are to be fixed as contracts between the intervening actors in the scene play agreements in the Interaction Model.

3.1.4.   Society norms and rules

The actions of agents when performing roles have consequences for the future actions and possibilities of the actor. Society norms may be related to the actions of an agent within an interaction structure or to the connections between scenes, determining the possible performative paths that agents can take within the society depending on their roles. For instance, a possible rule on a trading house society is that a trading agent

that wins a bidding round is consequently obliged to pay for the goods. Another rule can be that an agent assuming a given rule must take up contracts accepted by the previous enactor of the role.

We represent society norms as deontic logic expressions [19]. For instance, the obligation of the buyer to pay for the goods bought in an interaction structure representing an exchange is represented as:

```
OBL(buyer, pay-goods(Prod, Price, Seller), settle-trade)
```

where `buyer` is an agent playing the buyer role, `pay-goods(Prod, Price, Seller)` a possible expression in the communication language (defined in CF), and `settle-trade` a interaction scene in I.

Norms can be related to roles, scenes or transitions between scenes. Roles norms indicate the rules of behavior for actors performing a role irrespective of interaction scene. Scene norms are specified in an interaction script and describe the expected behavior of actors within an interaction scene. Finally, arcs between scenes in an interaction structure define the possible paths agents may follow depending on their roles. It is possible to impose constraints over outgoing arcs of an interaction script that impose additional limitations to actors attempting to follow that arc.

3.2. Social model

The interaction structure as specified in the organizational model indicates the interaction possibilities for agents at each stage and the consequences of their choices. However, further representation of commitments between agents is necessary in order to verify the 'run-time' behavior of the society. Commitments enable actors to know how to proceed if the norm is violated but first of all indicate to the agents what behavior they can expect from the other agents. In this way, coordination can become possible. Therefore, a formal language for specification of commitments is needed in order to be able to evaluate and verify agent interaction. We propose a logical framework, the Logic for Contract Representation (LCR), as possible formalism for representation of commitments. This logic is introduced in section 4. A first example of application of such explicit commitments, or contracts, appears in the social model, where it is used to describe the agreements between an agent an the society concerning the agent's behavior.

We define a social model as a tuple SM = (OM, A, M), where OM is an organizational model of an agent society, A is a set of agents and M is the set of social contracts mapping agents in A to roles. M represents the accepted agreements for role-playing or contracts between agents and the society as of their presence in the society.

A **social contract** SC: A × *roles*(OM)* → C describes the conditions and rules applying to an agent enacting role(s) in the agent society. C is the set of contract clauses, defined as a tuple (P, N, V, S, T) where P is the set of **pre-conditions**, N is a logical expression defining a **deontic-modality** describing the role to be played by the agent, V is a time expression defining the **validity** of the clause, S defines the possible **states** of the contract clause, and T gives the **transition rules** between states. Contract clauses are described in more detail in section 4.1.

Informally, social contracts must determine the operational roles and social norms applicable to an agent that is going to play a role in the society. Social contracts must describe:
- Role(s) to be played by agent
- Rules and interaction structures involved
- Time period
- Price and/or Conditions for agent action as enactor of role
- Sanctions

A consequence of the organizational perspective taken is that agents are viewed in our framework as black boxes and therefore the model above makes no demands on the architecture of individual agents participating in the society. In order to support design and evolution of these models, it is necessary to describe mechanisms that prescribe how social contracts can be set up and negotiated. This means that it is needed to find a way to describe agent capabilities and constraints in order to be able to describe the process of mapping an agent a to a role description r. This meta description of agent capabilities must be independent from agent architectures, languages and models. Further research is needed on this.

3.3. Interaction model

The previous specification models for agent societies describe the organizational structure and desired behavior (OM) and the commitments between society and participating agents (SM). Such models still do not completely specify the actual interactions between agents. That is, there is possible still some room for agents to decide on how to fulfil its organizational role(s). Commitments between actors on how to realize a needed interaction must also be specified in order to be able to verify the actual behavior of a society. As in the SM, such commitments are represented as contracts, formally specified using LCR.

We define the interaction model of an agent society as a tuple IM = (SM, IC), where SM is a social model and IC the set of interaction contracts between agents in SM. An **interaction contract** IC: *agents*(SM)* $\rightarrow$ C describes the conditions and rules applying to interaction between agents in the agent society. C is the set of contract clauses as above. In the some way social contracts describe the roles and norms applicable to an agent as enactor of a role in the society, interaction contracts describe the operational roles and social norms applicable to the interaction between agents. Interaction contracts have two or more contractors and must describe:
- Description of the agreement(s)
- Rules and interaction structures involved
- Time period
- Price and/or Conditions for action of each agent
- Sanctions

We have argued before that the type of coordination structure applicable to an agent society has consequences in terms of freedom of participation and interaction of agents [11]. That is, for example in an hierarchical society, interaction between agents follows pre-defined paths and rules, and agents have less freedom to determine their own contacts. In the other hand, a network setting leaves enough room for agents to

determine their own partners and decide on rules for the interaction. Therefore, the coordination type of the society determines how interaction contracts between agents are negotiated and how much must be described in the society. Due to the differences in autonomy and interaction, social contracts in different society types will specify more or fewer aspects of the agent actions and interactions in the society. In the following we consider the most generic type of contracts. We are currently working on the modification of this generic contract type to the different society types.

4. CONTRACTS

A contract can be defined as a statement of intent that regulates behavior among organizations and individuals [20]. Current models and implementations of multi-agent systems often make use of implicit (hard-coded) information to represent shared context which makes interoperability of heterogeneous agents difficult [6]. Contracts have been proposed as means to make explicit the way agents interact with and within the society. Contracts provide a means to specify and eventually verify that the required global properties of the society will indeed emerge from the interactions between agents. This view is similar to the architecture proposed by [1] for the coordination of software components.

Contractual agent societies are inspired by the work of a number of organizational theorists, economists and interactionist sociologists, who model organization and social systems after contracts. From a contractual perspective, organizations can be seen as sets of agreements for satisfying diverse interests of self-interested individuals [6]. Social order emerges from the negotiation of the rights and duties of participants. As described above, our framework uses two types of contracts:

- **Social contracts** are explicit specifications of commitments between an agent and the society, such as which roles the agent will play, what are the rules it must adhere to, what can the agent expect from the society, etc.
- **Interaction contracts** specify agreements between individual agents.

This use of the term social contract differs from [6] where both the above types are called social contracts.

4.1. Logic for Contract Representation

We are developing a language for contract description to be used in agent society frameworks. Because the primitives of both social and interaction contracts are similar (cf. 3.2. and 3.3.), in this language both type of contracts are represented in the same way. This language is build over a Logic for Contract Representation (LCR) based on branching-time logic. In the following we briefly introduce the most relevant aspects of LCR. A more detailed description can be found in [9].

Formulae in branching-time logics are interpreted over tree-type branching structures that represent all conceivable ways the system can evolve to. Nodes represent *states* and arcs correspond to the occurrence of *events*. A *path* represents a course of events and links states in the time structure according to the choices and possibilities available to agents at each moment. Our proposal extends the formalism

based on Temporal and Deontic Logic, BTLcont, proposed by Dignum and Kuiper [7]. BLTcont is in itself an extension to the well known branching-time temporal logic (CTL*) proposed by Emerson and Halpern [12]. It is beyond the scope of this paper to describe LCR formally, we will just describe its most relevant features.

We further extend branching time logic with a *stit* operator, E_a ('agent a sees to it that'). This allows us to refer to the externally 'observable' consequences of an action instead of the action itself. Remember that agent internals are not visible from the organizational perspective, and therefore it is not possible to refer to specific actions of an agent. In our use of E_a (based on Pörn) we draw from the logic proposed by Wooldridge for the combination of a *stit* operator with a temporal logic [14].

Moreover, clauses in a contract (deontic expressions in LCR) indicate that something must happen (it is desirable that something happens) but in fact if may never happen at all! A logic for contract representation must therefore be able to reason about states in which an obligation has been violated. Obligations have to do with the preference of individuals (or societies) to be in a certain state. $O_a\varphi$ (the obligation for agent a to see to it that φ holds) indicates that, in the holding society, it is preferable for a to be in a state where φ holds than in any other state. This does not mean that agent a cannot be in other states either by choice or necessity. A violation, viol(a,φ), is interpreted as 'agent a is in a violation situation concerning φ'. The basic idea is that worlds in which a violation proposition holds are less preferred by the agent concerned. In order to make it possible that agents can be redeemed from a violation, sanctions can be defined. A sanction is a (conditional) obligation whose realization will remove the violation.

4.2.   Language for contracts

In the following, the language for contracts is introduced. Social contracts describe a possible instantiation of role descriptions for a specific agents. Interaction contracts describe commitments between actors that specify a possible instantiation of an interaction script. Besides the parties directly involved in the contract, we assume that contracts can also indicate third parties, that is, actors(s) responsible for monitoring the execution of the contract and apply sanctions whenever needed. A contract is specified by defining:
-    The set of participating actors.
-    A possibly empty set of monitoring agents, responsible for the control of the contract fulfillment and the application of sanctions (the action of monitoring agents follows a kind of 'meta'-contract, described bellow)
-    The set of contract clauses. Contract clauses are deontic expressions of LRC and may as such indicate deadlines and/or conditions. Deontic modalities express the deontic expectation over an action. Actions are either speech acts (inform, confirm, request, etc.) or material acts (pay, deliver, write, etc.) Deontic actions express explicitly the semantics of the intended result.

In order to describe and verify contract execution, possible states for a contract must be indicated as well as how to move between states. A contract transition graph represents the possible evolution(s) of the contract and the consequences of the changes in state to the different parties. Transitions between states represent events in the society.

The following contract represents a possible instantiation for the interaction scene 'Settle Trade' in the Trading Society: S is a seller actor and B is a buyer actor. S and B agreed before that S will sell a bicycle to B for 500. Landmarks defined in the interaction script for this scene indicate that the buyer is obliged to pay for the goods (cf. Example in 3.1.4) but do not define any specific procedure for the trade. The following contract expresses the exchange commitment agreed between S and B: S has 2 days to give the bicycle to B after which B must pay S within 1 day. If S does not provide the bicycle on time, then the exchange does not go through. If B does not pay on time then an extra 10 is due. A very informal representation of this contract is as follows:

```
CONTRACT:
    CONTRACTING PARTIES: S (seller), B (buyer)

    CONTRACT CLAUSE 1:
        IF exchange(S,B, bicycle, €500)
        THEN OBLIGED(S, get-bike(B, bicycle) BEFORE 2 days

    CONTRACT CLAUSE 2:
        IF get-bike(B)
        THEN OBLIGED(B, get-money(S, €500) BEFORE 1 day

    CONTRACT CLAUSE 3:
        IF VIOLATION(S, get-bike(B, bicycle) BEFORE 2 days)
        THEN cancel-exchange(S,B, bicycle, €500)

    CONTRACT CLAUSE 4:
        IF VIOLATION(B, get-money(S, €500) BEFORE 1 day)
        THEN OBLIGED(B, get-money(S, €510) BEFORE 2 days
END CONTRACT
```

Note that this exchange is tipycally possible in the case S and B trust each other. A possible instantiation of 'Settle Trade' for the case that S and B do not trust each other is to use a trusted third party T (S gives the bicycle to T which requests and received the money from B after which T will give the bike to B and the money to S).
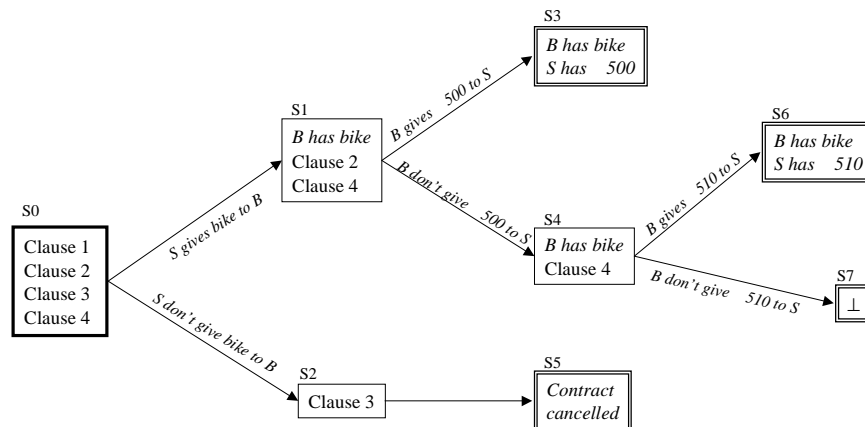

Fig. 5. Contract transition graph

Figure 5 displays the contract transition graph for this contract. S0 is the initial state where all contract clauses are valid and S5, S6 and S7 are final states. Final state

S7 represents a situation not specified in the contract. Since contracts do not have to be exhaustive, these cases may happen often. Consequences of reaching such state can be dealt with by society norms (for instance, in the case above, B could be expelled from the society).

5. RELATED WORK

Our model combines ideas from a behavioral design of agent societies with contracts, as formalism for the representation of interaction agreements. In the following we describe some related work on both areas. To our knowledge there is no other work being done on the use of contracts as means to integrate the specification of organizational structures with agent autonomy.

5.1. Work on organizational frameworks for MAS

The main challenge in multi-agent society design is that of coordination. This challenge has been recognized by many authors and several approaches have been developed and advocated. Such approaches take either a bottom-up (eg. goal management in which members of the group take control of the definition of their work [18]) or a top-down view of coordination (eg. shared ontologies [16] and the hierarchical assignment of responsibilities used in many human organizations).

Behavioral approaches to the design of multi-agent systems are gaining terrain in agent research and several research groups have presented models similar to our proposal. Recent developments recognize that the modeling of interaction in MAS cannot simply rely on the agent's own (communicative) capabilities. Furthermore, organizational engineering of MAS cannot assume that participating agents will act according to the needs and expectations of the system design. Concepts as organizational rules [25], norms and institutions [13] and social structures [21] all start from the idea that the effective engineering of MAS needs high-level, agent-independent concepts and abstractions that explicitly define the organization in which agents live [26].

AALAADIN was one of the first architectures for multi-agent systems [14] that is based on organizational principles. This architecture is based on three elements: *agent*, an active communicative entity, *group*, a set of agents, and *role*, the abstract representation of an agent function or service within a group. However expressive, AALAADIN does not offer primitives to describe interaction and coordination within and between groups and agents and the environment. This model was used as basis for a proposal for representation of social structures in AUML that does describe interaction between roles [21]. The model developed by the Alfebiite consortium is meant for the design of open agent societies and considers aspects of security and legal consequences of agent action in agent societies [2]. The model includes representation primitives for agents, constraints, communication language, roles, social states and agent owners. In our opinion, this model lacks primitives for the representation of groups and complex interaction and coordination in a society. Finally, the approach proposed in [13] focuses on the normative aspects of societies.

This proposal aims at the modeling of institutionalized electronic organizations (institutions).

5.2.   Work on Coordination Contracts

The use of contracts as means of formalize and specify coordination between autonomous entities has been proposed both for open information systems as well as for software engineering. Dellarocas introduces Contractual Agent Societies as open information systems where independently developed agents configure themselves automatically through a set of dynamically negotiated contracts [6]. A Contractual Agent Society (CAS) is a multi agent system where coordinated activity emerges from a set of negotiated social contracts (social norms) enforced through mechanisms of social control. In a CAS independently developed agents configure themselves automatically and coordinate their behavior through a set of dynamically negotiated social contracts. The intended behavior of the system as a whole can be derived from the existing contracts.

Andrade et.al see coordination contracts as semantic primitives that provide mechanisms for establishing interactions between software components [1]. Coordination contracts are put in place among a set of components to ensure that required global properties of the system will emerge from the functionality of the components and the interaction established between them through the contract.

6.   CONCLUSIONS AND FUTURE WORK

This paper presents our current research that aims at the specification of a framework to describe agent societies. In order to be applicable to organizational domains, the model takes a collectivist view on agent societies, that considers the global characteristics of the domain. The framework consists of three models what allows for a conceptual separation between organizational and individual perspectives. The organizational model describes agents societies in terms of roles, constraints and interactions rules. The organizational structure is specified in two layers that separate the representation of organizational and operational aspects of the domain. The social model describes the active community in the domain as a population of agents that realize the objectives of the society by enacting role(s) described in the organizational model. The relation between roles in the organizational model and the agents is described by contracts. Contract rules specify commitments between agents and society concerning role enactment. Finally, the activity and interaction between agents is specified in the interaction model. As in the social model, contracts are used to represent commitments between agents concerning interaction. Our work combines a society model similar to [2] and [13] with contracts, as a formalism to specify cooperation between agents and society, as proposed by [6] and [1]. Our aim is to provide a model for agent societies that incorporates the concept of contracts as first order objects that describe and verify coordination between agents. Current work concentrates on the formalization of the concepts introduced and on the semantics of contracts. Besides several aspects already mentioned in the different sections, future work includes the refinement of the model presented in this paper to a degree that

admits verification of global society requirements in one hand, and the goals of individual agents on the other hand. Furthermore, we are investigating the implications of the framework to the specification of individual agents.

7. REFERENCES

[1]   Andrade, L., Fiadeiro, L, Gouveia, J., Koutsoukos, G. Lopes, A. and Wermelinger, M.: Coordination Patterns for Component-Based Systems. Musicante M., Haeusler H. (Eds.): Proc. of Brazilian Symposium on Programming Languages, UFPR 2001, (2001), 29-39.

[2]   Artikis, A., Pitt, J.: A Formal Model of Open Agent Societies. Proc. 5th Conference on Autonomous Agents, (2001) 192-193.

[3]   Castelfranchi, C.: Engineering Social Order. In: Omicini, A. Tolksdorf, and Zambonelli, F. (Eds.): Engineering Societies in the Agent World, LNAI 1972, Springer (2000).

[4]   Dastani, M., Dignum, F., Dignum, V.: Organizations and Normative Agents. (2002). Proc. First Eurasian Conference on Advances in Information and Communication Technology – Eurasia ICT, Tehran, Iran, (2002).

[5]   Davidsson, P.: Emergent Societies of Information Agents. Klusch, M, Kerschberg, L. (Eds.): Cooperative Information Agents IV, LNAI 1860, Springer, (2000), 143–153.

[6]   Dellarocas, C.: Contractual Agent Societies: Negotiated shared context and social control in open multi-agent systems. Proc. WS on Norms and Institutions in Multi-Agent Systems, Autonomous Agents, Barcelona, (2000).

[7]   Dignum,F., Kuiper, R.: Specifying Deadlines with Dense Time using Deontic and Temporal Logic. In: International Journal of Electronic Commerce, 3(2), (1999), 67 – 86.

[8]   Dignum, V., Dignum, F.: Modeling agent societies: coordination frameworks and institutions. In: Brazdil, P, Jorge, A.: Progress in Artificial Intelligence. LNAI 2258, Springer-Verlag, (2001), 191 - 204.

[9]   Dignum, V., Meyer, J.-J., Dignum, F., Weigand, H.: Formal Specification of Interaction in Agent Societies. 2nd Goddard Workshop on Formal Approaches to Agent-Based Systems (FAABS), Maryland, Oct, (2002)

[10] Dignum, V., Weigand, H., Xu L.: Agent Societies: Towards framework-based design. Wooldridge, M., Weiss, G., Ciancarini P. (Eds.): Agent-Oriented Software Engineering II, LNCS 2222, Springer-Verlag, (2002) 33-49.

[11] Dignum, V., Weigand, H.: Towards an organization-oriented design methodology for agent societies. In: V. Plekhanova (Ed.): Intelligent Agent Software Engineering. Idea Publishers. (2003). In press.

[12] Emerson, E.: Temporal and Modal Logic. In: J. van Leeuwen (Ed.): Handbook of Theoretical Computer Science, Elsevier Science Publishers (1990).

[13] Esteva, M., Padget, J., Sierra, C.: Formalizing a language for Institutions and Norms. Proceedings of the 8th International Workshop on Agent Theories, Architectures and Languages, ATAL-2001, Seattle, (2001)

[14] Ferber, J., Gutknecht, O.: A meta-model for the analysis and design of organizations in multi-agent systems. Proc. of ICMAS'98, IEEE Press, (1998).

[15] Finin, T., Labrou, Y., Mayfield, J.: KQML as an agent communication language, in Bradshaw, J. (Ed.): Software Agents, MIT Press, Cambridge, (1997).

[16] Fox, M., Gruninger, M.: Enterprise Modeling. AI Magazine, AAAI Press, Fall, (1998), 109-121.

[17] Lind, J.: Issues in Agent-Oriented Software Engineering. In: Ciancarini, P., Wooldridge, M. (Eds.): Agent-Oriented Software Engineering, LNCS 1957, Springer, (2000), 45 – 58.

[18] Malone, T., Crowston, K.: The Interdisciplinary Study of Coordination. ACM Computing Surveys, 26(1), March, (1994).

[19] Meyer. J-J. and Wieringa, R.: Deontic Logic in Computer Science, Wiley, (1993).

[20] Morciniec, M., Salle, M., Monahan, B.: Towards Regulating Electronic Communities with Contracts, Proceedings of 2nd Workshop on Norms and Institutions in MAS, Autonomous Agents 2001, Montreal, (2001).

[21] Parunak, H. V. D. and Odell, J.: Representing Social Structures in UML. In: Wooldridge, M., Weiss, G., Ciancarini P. (Eds.): Agent-Oriented Software Engineering II, LNCS 2222, Springer-Verlag, (2002) 1 - 16.

[22] Searle, J.: Speech Acts. Cambridge Univ. Press, (1969).

[23] Wooldridge, M., Jennings. N.: Software Engineering with Agents: Pitfalls and Pratfalls, IEEE Internet Computing 3 (3) (1999), 20-27.

[24] Wooldridge, M.: Time, Knowledge, and Choice . In: M. Wooldridge, J. P. Mueller, M. Tambe, (Eds.): Intelligent Agents II, Springer-Verlag, (1996).

[25] Zambonelli, F.: Abstractions and Infrastructures for the Design and Development of Mobile Agent Organizations. In: Wooldridge, M., Weiss, G., Ciancarini P. (Eds.): Agent-Oriented Software Engineering II, LNCS 2222, Springer-Verlag, (2002), 245 – 262.

[26] Zambonelli, F., Jennngs, N., Wooldridge, M.: Organizational Abstractions for the Analysis and Design of Multi-agent Systems. In: Ciancarini, P., Wooldridge, M. (Eds.): Agent-Oriented Software Engineering, LNCS 1957, Springer, (2000), 235 – 251.