World Academy of Science, Engineering and Technology
International Journal of Humanities and Social Sciences
Vol:6, No:11, 2012

A Cognitive Architectural Approach to the Institutional Roles of Agent Societies

Antônio Carlos da Rocha Costa

***Abstract*—**This paper concerns a formal model to help the simulation of agent societies where institutional roles and institutional links can be specified operationally. That is, this paper concerns institutional roles that can be specified in terms of a minimal behavioral capability that an agent should have in order to enact that role and, thus, to perform the set of institutional functions that role is responsible for. Correspondingly, the paper concerns institutional links that can be specified in terms of a minimal interactional capability that two agents should have in order to, while enacting the two institutional roles that are linked by that institutional link, perform for each other the institutional functions supported by that institutional link. The paper proposes a cognitive architecture approach to institutional roles and institutional links, that is, an approach in which a institutional role is seen as an abstract cognitive architecture that should be implemented by any concrete agent (or set of concrete agents) that enacts the institutional role, and in which institutional links are seen as interactions between the two abstract cognitive agents that model the two linked institutional roles. We introduce a cognitive architecture for such purpose, called the Institutional BCC (IBCC) model, which lifts Yoav Shoham's BCC (Beliefs-Capabilities-Commitments) agent architecture to social contexts. We show how the resulting model can be taken as a means for a cognitive architecture account of institutional roles and institutional links of agent societies. Finally, we present an example of a generic scheme for certain fragments of the social organization of agent societies, where institutional roles and institutional links are given in terms of the model.

***Keywords*—**Simulation of agent societies, institutional roles, cognitive architecture of institutional roles.

I. Introduction

INSTITUTIONAL roles and institutional links of agent societies have been defined either in terms of sets of norms that constraint the possible behaviors and interactions that the agents that implement those roles may perform, or else directly in terms of that set of possible behaviors and interactions (cf. the various approaches presented in, e.g., [12]).

In this paper, we propose a formal model to support the simulation of agent societies. The proposal adopts a cognitive architecture approach to the institutional roles and institutional links, that is, an account in which institutional roles can be defined in terms of the internal architecture of abstract cognitive agents, and in which institutional links are seen as interactions between the abstract cognitive agents that model institutional roles.

We introduce a cognitive architecture for such purpose, namely, one that adapts Yoav Shoham's BCC (Beliefs Capabilities-Commitments) [22] agent architecture to social contexts, through the introduction of the notions of institutional capabilities and institutional commitments. It is showed how the resulting Institutional BCC (IBCC) model can be taken as a means for an architectural account of institutional roles and institutional links of agent societies.

Finally, we present a generic institutional segment of a simple agent society, where institutional roles and social links are computationally specified by abstract cognitive agents.

The paper is structured as follows. Section II summarizes the concepts of social roles and institutional links. Section III briefly presents the main features of Shoham's BCC agent model. Section IV discusses the institutional concepts that are necessary to introduce in the BCC model to turn it into the abstract cognitive model required by its lifting to social contexts, specially institutionalized social contexts. Section V formally introduces the IBCC model of institutional roles and institutional links. Section VI presents an example of a schematic institutional segment of agent society. Section VII comments on related works. Section VIII is the Conclusion.

II. Social Roles in Agent Societies

In this section, we place the concept of agent society in perspective, explaining how the account of institutional roles that we present in the following sections fits into that concept. We take that an agent society may be conceived as a system endowed with at least the following components [1]:

- a population of agents whose behaviors generate the dynamics of the system;
- a network of interactions between the agents, the interactions implying the performance of institutional functions for each other (e.g., exchanges of services, objects or information);
- an institutional structure that both regulates the behaviors and interactions of the agents, and provide for their integration into a persistent social system.

The notion of institutional role is crucial to any

A. C. R. Costa is with Centro de Ciências Computacionais, Universidade Federal do Rio Grande, RS, Brazil (e-mail: ac.rocha.costa@gmail.com).

[1] In previous works, like [8]-[10], we have been elaborating the PopOrg model of *organization* of agent societies. The PopOrg model is now being extended to become a more general *institutional model* of agent societies. Both the relation between the notions of organization and institution, and the functionalist basis of the extended model can not be fully discussed here (cf., however, [16]). The part already obtained of this more general model is used here to support the proposed cognitive architecture approach to institutional roles.

interactional approach to agent societies, since institutional roles can be taken as abstract specifications of sets of agents behaviors that the institutional structure of the society specifies as adequate for the agents, depending on the way they are inserted in the society.

Also, the notion of institutional role implies the complementary notion of institutional link. We take a social link between two or more institutional roles to be an abstract operational specification of sets of possible interactions that are expected to be performed, in various situations, by the agents enacting the institutional roles linked by the institutional link.

In combination, institutional roles and institutional links imply the two kinds of elements of the institutional structure: regulative elements (like norms, customs, etc.), and integrative elements (such as social values and institutional functions), which the agents accept, adopt, and become responsible for, when enacting those institutional roles.

III. The BCC Agent Model

The Beliefs-Capabilities-Commitments (BCC) agent model introduced here is essentially the agent model that underlies the Agent0 agent oriented programming language introduced by Yoav Shoham in [22] [2].

The BCC agent model was defined to occupy an intermediate position in the hierarchy of refinements of agent models that goes from the higher-level, more general, Actor model [14], and the more refined, more detailed, agent models like BDI [19] and SOAR [15].

Contrary to the Actor model, the BCC agent model is a cognitive model, accounting for the internal information processing of the agents in terms of cognitive concepts, like beliefs and behavioral commitments and capabilities. On the other hand, contrary to models like BDI and SOAR, the BCC agent model is not a motivational model, that is, it abstracts away mental concepts like goals, desires, intentions, and the like.

In other terms, the BCC agent model may be characterized as an architecture for agents that perform reactive behaviors (better: non-deliberative behaviors), endowed with several special features, such as:

- reactions need not be immediately produced: they may be committed to certain time in the future, determined by a given reaction delay (we call such mechanism behavioral commitments);
- agents may interact with each other through speech acts, specially in the form of request, unrequest and inform messages:
  - `req(ag,act,t)`, when received by agent ag at a time t' < t, behaviorally commits `ag` to perform action act at time t;
  - `unreq(ag,act,t)`, when received by agent `ag` at a time t' < t, cancels the previously established behavioral commitment;
  - `inform(ag,i)`, when received by agent `ag` at any time, provides `ag` with the information i, which `ag` may choose to believe or not;
- agents may have perceptions of their environments, which are independent of the messages they exchange.

The control of the agent behavior is performed, in a BCC agent, through sets of behavioral commitment rules, which determine the conditions (actual beliefs and behavioral capabilities) under which the agent should or is allowed to commit to new behavioral commitments. Discharge of behavioral commitments is automatically controlled by the agent architecture, by taking into account the passage of time.

We remark that the behavioral capabilities and commitments involved in the BCC agent model are of an internal nature, that is, they concern the internal operation of the agents. In the IBCC model of institutional roles, introduced below, the capabilities and commitments involved are of an external (or better, institutional) nature, that is, they concern social interactions performed in the context of an institutional structure.

IV. From the BCC Model to the IBCC Model

Two main modifications were introduced in the BCC model, in order to turn it into the IBCC model. The first modification concerns the notion of capability employed in the model, while the second is concerned with the notion of commitment.

A. Institutional Capabilities

While the BCC agent model is concerned with behavioral capabilities of individual agents that inhabit the population structure of agent society, the IBCC model is concerned with the so-called institutional capabilities of the institutional roles that exist in the institutional structure of the agent society.

The notion of institutional capability adopted in the model is implied by that of institutional right: to say that a institutional role has the institutional capability of performing an action or behavior is the same as saying that it has the institutional right to do it.

A full discussion of the notion of (institutional) right is not possible at this point. It suffices to say that (institutional) rights are institutional facts and, as such (cf., e.g., [21]):

- are created by the common beliefs of the agents of the society;
- are hold by the entities that have them only as long as those common beliefs are maintained;
- are attributed to the entities that have them by entities specially entitled to allocate them;
- are attributed through particular institutional acts expressed through special kinds of speech acts.

We remark that we take institutional rights (and so, institutional capabilities) to be held, in the first place, by institutional roles, not by the agents that enact those roles.

[2] However, since the agent model that underlies Agent0 was not defined in a fully formal way in [22], we do not claim that the BCC model is a faithful presentation of that model. We content ourselves with the BCC model introduced here having an intuitive similarity with Shoham's agent model.

V. Institutional Commitments

While behavioral commitments in the BCC model are established between agents, essentially due to their interindividual interactions, in the IBCC model institutional commitments are established between agents due to their enaction of institutional roles in the institutional structure of the agent society.

So, in the IBCC model, institutional commitments concern, in the first place, not the individual agents as agents, but concern them as role enactors. That is, an institutional commitment ic, taken by an agent a_1, which was enacting role r_1, toward an agent a_2, which was implementing role r_2, should be kept and fulfilled independently of which is the agent that is enacting role r_2 at the time the commitment is to be fulfilled: if agent a_1 (resp., a_2) is substituted by agent a_1' (resp., a_2') in the enaction of role r_1 (resp., r_2), the institutional commitment ic is to be fulfilled even if a_1 (resp., a_2) has never interacted with a_2' (resp., a_1'), previously.

However, we find that, institutional commitments do not differ formally from the social commitments that concern interindividual interactions. So, we model institutional commitments on the formal concept of social commitment, as that concept was defined both in [3, 4] and in [23] 3.

Thus, we take an institutional commitment to be a relation that is established between two institutional roles (the source and the target of the institutional commitment), before a third institutional role (the authority), about the performance of a given institutional function by the source institutional role to the target institutional role.

We remark that we model institutional commitments on the notion of persistent mutual commitments: we take them to be persistent because we take institutional functions to be persistent, and we take them to be mutual, because an institutional function can be realized only if both the source and the target institutional roles participate in the exchange process that performs that institutional function.

Symbolically, institutional commitments could then be denoted by expressions of the form `ICmt(x,y,z,f)`, where:

- `x` is the source institutional role, i.e., the one which commits itself to the performance of the institutional function;
- `y` is the the target institutional role, i.e., the one on behalf of which the institutional function is to be performed;
- `z` is the authority, i.e., the institutional role before which the institutional commitment is made;
- `f` is the institutional function that `x` commits itself, before `z`, to perform to `y`.

However, we remark two issues concerning institutional commitments. First, that the institutional functions are taken to be persistent in the operation of the agent society (cf. [11]) and, so, concern persistent services performed by the source institutional role that committed to the institutional function. They do not concern episodic services (episodic tasks) that agents perform to each other, as is often assumed in the general concept of social commitment.

Second, we note that a formal analyzes of the relation between institutional functions, persistent services, and the persistent exchange processes through which institutional roles perform persistent services to each other and, through that means, the institutional functions to which they are committed, is out of the scope of this paper (cf. [11]).

So, we simply assume here that, when formally defining an institutional commitment, we will refer not directly to the institutional function involved in the institutional commitment, but to the persistent exchange process through which the source institutional role performs the institutional function to the target institutional role.

That is, institutional commitments will be formally denoted, below, by expressions of the form `ICmt(x,y,z,e)`, where `e` is the persistent exchange process that realizes the (implicitly indicated) institutional function to which the institutional commitment refers.

VI. Performatives for the Control of Institutional Commitments and Capabilities

The institutional commitments of an institutional role vary in time as the institution is functioning. In this, the IBCC model is like the BCC model, where behavioral commitments vary in time.

Differently from the BCC model, however, where the set of behavioral capabilities of an agent was assumed to be fixed, in the IBCC model the set of institutional capabilities of an institutional role is assumed to be variant, according to the temporal evolution of the institution.

To allow for time-variant institutional capabilities in the IBCC institutional roles, we make use of institutional messages exchanged between institutional roles. Through such an institutional message, an institutional role can attribute an institutional capability to, or withdraw it from, another institutional role.

We leave out of the IBCC model, however, the concern with the verification of the appropriateness or legality of such attributions or withdrawals of institutional capabilities. This concern is clearly a macro-level concern in the institutional structure of an agent society and is, thus, external to the characterization of what is an institutional role.

Also, we leave out of the rationale for the IBCC model, the analysis of the speech acts and the declarative performatives that characterize the attribution and withdrawal of institutional capabilities(cf., e.g., [21]). The intuitive content of the expressions below, we think, should be enough for the understanding of operation of the institutional messages.

So, in the formal introduction of the model that we give

[3] We notice that the concept of social commitment defined in [23] is an extension of that in [3]. So we take the latter as the basis for our work.

below, we adopt without further discussion the following expressions to represent such speech acts:

- `attr_icap(r,f,r')` meaning: institutional role `r` is attributed the institutional capability to perform function `f` to the institutional role `r'`;
- `wthdrw_icap(r,f,r')` meaning: the institutional capability to perform function `f` to institutional role `r'` is withdrawn from the institutional role `r`;

As explained above, however, institutional functions will not be directly referred to in the formalization introduced below. Instead of that, institutional functions will be indirectly referred through the permanent exchange process that realizes it. Since the institutional capability to perform an exchange process requires the institutional capability to perform the behaviors that implement such exchange process, the following will be the declarative expressions that will be effectively used:

- `attr_icap(r,b)` meaning: the institutional capability to perform behavior `b` is attributed to institutional role `r`;
- `wthdrw_icap(r,b)` meaning: the institutional capability to perform behavior `b` is withdrawn from institutional role `r`;
- `attr_icap(r,e,r')` meaning: the institutional capability to perform exchange process `e` is attributed to institutional roles `r` and `r'`;
- `wthdrw_icap(r,e,r')` meaning: the institutional capability to perform exchange process `e` is withdrawn from institutional roles `r` and `r'`.

Also, for simplicity, we allow institutional exchange processes to be represented by the pair of institutional behaviors that implement them. That is, each institutional exchange process `e` performed by institutional roles `r` and `r'` may be represented by the pair of behaviors `(b,b')` that are such that, when `r` performs `b` and `r'` performs `b'`, the exchange process `e` (and, thus, the corresponding institutional function) is performed between `r` and `r'`. In such case, then, one would write `attrb-icap(r,(b,b'),r')`, instead of `attrb-icap(r,e,r')`, for denoting `attrb-icap(r,f,r')`.

Similar remarks apply to messages of the form `req(r,f,r',t)`, among others.

VII. The IBCC Model of Institutional Roles and Institutional Links

The following are the main features of the IBCC model:

- the IBCC model is designed to express, in a cognitive architectural way, the institutional roles that exist in agent societies;
- in the IBCC model, role capabilities are institutional capabilities, concerning the performance of institutional behaviors, institutional exchange processes, and institutional functions;
- the repertoire of speech acts involved in the operation of the model includes speech acts concerning the attribution and withdrawal of institutional capabilities to institutional roles;
- in the IBCC model, commitments are persistent institutional commitments, which directly concern institutional roles, and which concern agents only in so far as they enact institutional roles.

We remark that the cognitive architecture approach to institutional roles implies that institutional roles are seen in terms of the beliefs and behavioral rules that guide the agents that enact them, when dealing with the institutional capabilities and commitments of those roles.

To allow for that, the notion of an institutional role in the IBCC model is that of an abstract agent endowed with an institutional interface and a cognitive structure.

The institutional interface of an institutional role accounts for the way an agent that enacts the role presents itself to agents that also implement institutional roles in the agent society. In particular, the institutional interface specifies the institutional capabilities, institutional commitments and institutional behaviors that the agent is supposed to perform and deal with, when enacting the institutional role.

The cognitive structure of an institutional role accounts for the cognitive requirements that should be met by any agent that purports to enact the role, in order to able to perform the institutional reasoning (reasoning about institutional capabilities, institutional interactions, institutional commitments, etc.) that is demanded from that institutional role. In other words, any agent supposed to enact the institutional role is supposed to implement such cognitive structure.

In the example IBCC model presented in Sect. VI, both the institutional interface and the cognitive structure of example types of institutional roles are illustrated.

A. The IBCC Model of Institutional Roles

For simplicity, let time have a discrete structure, T=(0,1,…). We define the following sets, regarding a given agent society:

- **IRo**: the universe of institutional roles (whose internal structure is specified below) that agents may enact.
- **IAct**: the universe of institutional actions (purposefully left unspecified) that institutional roles may perform, possibly including non-institutional (behavioral) actions performable by the agents that implement the institutional roles;
- **IMsg**: the universe of institutional messages (internal structure purposefully left unspecified) that agents may exchange when enacting institutional roles;
- **IProp**: the universe of institutional propositions (syntax and semantics purposefully left unspecified) with which institutional roles may both represent their institutional beliefs and construe the institutional messages they exchange with each other;

• **IBel** $\subseteq$ **IProp**: the universe of institutional beliefs about the agent society that agents may have when enacting institutional roles;

• **IBh** $\subseteq$ [T $\rightarrow$ $\wp$(**IAct**)]: the universe of institutional behaviors that institutional roles can perform, with the constraint that, for each institutional behavior b∈IBh and each time t∈T, there exist:

- a subset of actions in b(t) that are the input messages received, at time t, by the institutional role that performs b; that is, there exists ims⊆b such that ims∈ $\wp$(**IMsg**);

- a subset of actions in b(t) that are the output messages sent, at time t, by the institutional role that performs b; that is, there exists oms⊆b such that oms∈ $\wp$(**IMsg**);

• **IEp**=[T→ $\wp$(**IAct**)× $\wp$(**IAct**)]: the universe of institutional exchange processes (or, institutional interactions) that institutional roles can perform between each other, such that if iep∈**IEp** is an institutional exchange process performed by institutional roles r_1 and r_2 then, at each time t∈T it happens that iep(t)=(a_1, a_2), with a_1 the set of actions with which r_1 participates in iep at time t, and a_2 the set of actions with which r_2 participates in iep at time t;

• **ICmt=IRo×IRo×IRo×IEp**: the universe of institutional commitments that institutional roles may have between them, each (r_1,r_2,r_3,e) ∈**ICmt**, meaning that the institutional roles r_1 and r_2 are mutually committed before the authority r_3 to perform the institutional exchange process e;

• **IBelRl**=[**IMsg**× $\wp$(**IBel**)× $\wp$(**ICap**)× $\wp$(**ICmt**)→ **IBel**]: the universe of institutional belief rules through which the institutional roles update their beliefs;

•**ICapRl**=[**IMsg**× $\wp$(**IBel**)× $\wp$(**ICap**)× $\wp$(**ICmt**)→ **ICap**]:the universe of institutional capability rules through which the institutional roles update their institutional capabilities;

•**ICmtRl**=[**IMsg**× $\wp$(**IBel**)× $\wp$(**ICap**)× $\wp$(**ICmt**)→ **ICmt**]the universe of institutional commitment rules through which the institutional roles update their institutional commitments.

We define an institutional role as a time-variant structure, that is, as a time-indexed set of elements, each element representing the state of the institutional role at a given time.

Definition 1. An institutional role is a time-indexed set of states of institutional roles r^t = ($CgnStr^t$;$InstIntrf^t$) where:

– $CgnStr^t$ = (IBh^t; $IBel^t$; $IMsg^t$) is the cognitive structure of the institutional role;

– $InstIntrf^t$ = ($ICap^t$; $ICmt^t$; ICmtRl) is the institutional interface of the institutional role; and:

– IBh^t∈ $\wp$(**IBh**) is the set of institutional behaviors the institutional role is performing at time t, so that for all b∈IBh one has that b(t) is the set of actions that the institutional role performs at time t, due to the performance of the behavior b;

– $IBel^t$∈ $\wp$ (**IBel**) is the set of beliefs the institutional role has at time t;

– $IMsg^t$ ∈ $\wp$ (**IMsg**) is the set of messages the institutional role has received at time t;

– $ICap^t$ ∈ $\wp$ (**ICap**) is the set of institutional capabilities the institutional role has at time t;

– $ICmt^t$ ∈ $\wp$ (**ICmt**) is the set of institutional commitments the institutional role has at time t;

– ICmtRl ∈ $\wp$ (**ICmtRl**) is the set of institutional commitment rules followed by the institutional role.

We remark that institutional roles are essentially time-varying structures, due to temporal changes not only in their sets of institutional behaviors, institutional beliefs and received institutional messages, but also in their sets of institutional capabilities and commitments.

However, for simplicity, we are considering here institutional roles whose institutional interface comprises a set of institutional commitment rules that does not change in time. So, to architecturally characterize such institutional roles it is enough to take IBCC definitions whose sets of institutional commitment rules are time-invariant.

The internal control of the workings of the IBCC institutional roles can be defined as a direct adaptation of the internal control of the workings of the BCC model [22], and can be given by the following basic control loop:

1. Read the current institutional messages and update the sets of institutional beliefs, institutional capabilities and institutional commitments, according to the set of institutional commitment rules.
2. Execute the institutional commitments for the current time, possibly resulting in further changes in the sets of institutional beliefs, institutional capabilities and institutional commitments.
3. Go back to 1.

We remark that the intended aim of this informal definition of the control loop is that its operation satisfies the following constraints, when acting on the set of states of an institutional role.

Definition 2. The control loop of the internal functioning of the institutional rolesworks properly if and only if for any time t∈T the following constraints are satisfied:

1. For any institutional message $imsg_1$∈ $IMsg_1^t$ in any institutional role r_1∈**IRo**, there is an institutional role r_2∈IRo such that at some time t' < t it happened that for some institutional behavior ibh∈IBh^t_2 there was an institutional action α∈ibh(t0) which was the action of r_2 sending the message $imsg_1$ to r_1;

2. For any any institutional commitment (r1; r2; r3; e) ∈ $ICmt_1^t$ in any institutional role r_1∈**IRo** there is an institutional action α∈ibh1(t) of some institutional behavior ibh_1∈IBh_1^t such that α∈prj_1(e(t)); that is, for any institutional role r_1 committed to (continue to) perform an institutional function f

at a time t there is an institutional action that belongs to the exchange process e that realizes f, and that the institutional role performs at that time;

3. For any institutional capability $icap_1 \in ICap_1^t$ in any institutional role $r_1 \in$ **IRo** there was a time $t' < t$ and an action $\alpha_2 \in ibh_2(t')$ of some institutional behavior $ibh_2 \in IBh_2^{t'}$ of some institutional role $r_2 \in$ **IRo** such that α_2 was an action through which the institutional role r_2 sent a capability attribution institutional message with content $icap_1$ to r_1, and such that there was no time t'', with $t' < t'' < t$, for which there was an action $\alpha_3 \in b_3(t'')$ of some institutional behavior $b_3 \in IBh_3^{t'}$ of some institutional role $r_3 \in$ **IRo**, which was an action of sending to r_1 a capability withdrawal institutional message with content $icap_1$;

4. For any institutional belief $ibel_1 \in IBel_1^t$ in any institutional role $r_1 \in$ **IRo** there is an institutional message $imsg_1 \in IMsg_1^{t-1}$ with content inform(r_1; $ibel_1$), such that for some institutional belief rule $ibelrl_1 \in IBelRl$ it happens that $ibel_1 \in ibelrl_1(IMsg_1^{t-1}; IBel_1^{t-1}; ICap_1^{t-1}; ICmt_1^{t-1})$, and so that $imsg_1 \notin IMsg_1^t$;

5. For any institutional capability $icap_1 \in ICap_1^t$ in any institutional role $r_1 \in$ **IRo** there is a message $imsg_1 \in IMsg_1^{t-1}$ with content attr_icap(r_1; $icapt_1$), such that for some institutional capability rule $icaprl_1 \in ICapRl$, it happens that respectively, $icap_1 \in icaprl_1(IMsg_1^{t-1}; IBel_1^{t-1}; ICap_1^{t-1}; ICmt_1^{t-1})$, and so that $imsg_1 \notin IMsg_1^t$;

6. For any institutional commitment $icmt_1 \in ICmt_1^t$ in any institutional role $r_1 \in$ **IRo** there is an institutional commitment rule $icmtrl_1 \in ICmtRl$ such that $icmt_1 \in icmtrl_1(IMsg_1^{t-1}; IBel_1^{t-1}; ICap_1^{t-1}; ICmt_1^{t-1})$, and so that $imsg_1 \notin IMsg_1^t$.

Constraint 1 aims to guarantee that institutional messages are received by an institutional role only if they were previously sent by another institutional role. Constraint 2 aims to guarantee that every institutional action to which an institutional role is committed is effectively performed. Constraint 3 aims to guarantee that every institutional capability attributed to an institutional role is attributed by another institutional role. The only acceptable exceptions for this constraint concern the set of initial institutional capabilities of the institutional roles, that is, institutional capabilities that are attributed to the institutional roles by the system designer, at the initial time t=0. Constraints 4 to 6 concern the execution of (belief, capability, commitment) update rules.

B. The IBCC Model of Social Links

Let $r_1, r_2 \in$ **IRo** be two institutional roles interacting through some exchange process. One may classify the set of all the actions they perform during such exchange as follows:

- institutional input actions ii_1 (resp., ii_2), corresponding to perceptions about the environment, made by r_1 (resp., r_2), and to the arrival of institutional messages sent to r_1 (resp., r_2) by institutional roles other than r_2 (resp., r_1);
- institutional output actions io_1 (resp., io_2), corresponding to actions on the environment made by r_1 (resp., r_2), and to the sending of messages by r_1 (resp., r_2) to roles other than r_2 (resp., r_1);
- institutional exchange actions ie_1 (resp., ie_2), corresponding to institutional actions made by r_1 (resp., r_2) to r_2 (resp., r_1), often the sending of messages between them, with declarative performatives.

Given the above categorization of the actions of the interacting institutional roles r_1 and r_2, the set of actions $\cup\{ibh_i \mid ibh \in IBh_i^t\}$ that each role r_i performs at a given time t can be partitioned into its three subsets, namely, $ibh_1(t)=ii_1(t) \cup ie_1(t) \cup io_1(t)$ and $ibh_2(t)=ii_2(t) \cup ie_2(t) \cup io_2(t)$.

Since an interaction between r_1 and r_2 is an exchange process $e_{r_1r_2} \in IEp$, at each time t the interacting institutional roles r_1 and r_2 perform a pair of sets of institutional actions $e_{r1r2(t)}=(e_1(t),e_2(t))$, with $e_1(t) \subseteq ibh_1(t)$ and $e_2(t) \subseteq ibh_2(t)$.

We note that this structuring of the interaction between two institutional roles amounts to an operational cross-connection between them, as pictured in Fig. 1.

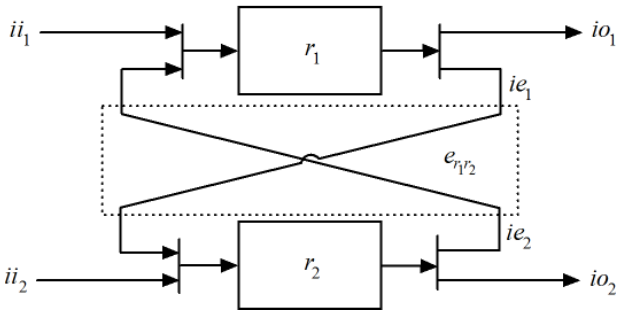

Fig. 1 Cross-connection of two institutional roles, to support their exchange processes

C. Coordination Schemes and the Support of Institutional Commitments

To guarantee the permanence of the social exchange, and thus the full accomplishment of the obligations, the agents that implement the institutional roles involved in an institutional commitment should be able to coordinate and maintain repeated performances of the social routines involved in the permanent social exchanges required by the institutional commitment. For such purpose, coordination schemes can be used [7], as illustrated below, in Section VI.

When the exchange process that supports the institutional commitment of two institutional roles is structured by a given coordination scheme, we say that those institutional roles are committed to the coordination scheme.

VIII. An Elementary Schematic Building Block for Agent Institutions

We present here, in terms of the IBCC model, an example of an elementary schematic building block for institutions of agent societies [4].

The elementary schematic building block is based on a very simple interaction coordination scheme, namely, the Production-Consumer (ProdCons) scheme, which has as large an applicability as the well-know Client-Server coordination scheme, which we could also have taken here, alternatively, for the example.

The ProdCons scheme is introduced below in an abstract way, as a place-transition Petri net [18]. It illustrates, in general terms, a way to coordinate a particular kind of interaction between two institutional roles, namely, an interaction in which one institutional role (the Producer) can be seen as producing something for the other institutional role (the Consumer).

First, we analyze the coordination scheme in terms of the patterns of permanent interactions that it induces between the two institutional roles. Then we show an IBCC presentation of the scheme, that is, a way in which the Producer and Consumer roles, and their interaction, can be described in terms of the IBCC model.

A. The ProdCons Coordination Scheme: a Petri Net Representation

We introduce the simplest version of the ProdCons scheme through the place-transition Petri net pictured in Fig. 2.

To the left of the figure, the net describes the observable behavior of the Producer: it can be in one of two conditions (Idle or Producing). It can perform two condition transitions (Start Production and End Production).

The Start Production transition can happen only when the Producer is in the Idle condition and the Consumer has performed the transition End Consumption, and it results in the Producer going from the Idle to the Producing condition.

The End Production transition can happen only when the Producer is in the Producing condition, and it results in the Producer going from the Producing condition to the Idle condition. It also results in the Consumer being enabled to perform the Start Consumption transition (if it is in the Idle condition).

To the right of the figure, the net describes the observable behavior of the Consumer: it can be in one of two conditions (Idle or Consuming). It can perform two condition transitions (Start Consumption and End Consumption). The effects of the transitions of the Consumer can be analyzed in a way analogous to that of the Producer.

The ProdCons scheme can be simultaneously instantiated for different types of products, by correspondingly instantiating the variable P. The state of the net pictured in Fig. 2 is one where the Consumer has just performed the End Consumption transition concerning the item of a certain type P and the Producer is enabled to perform the Start Production transition concerning the item of the same type P.

Notice that the Producer-Consumer coordination scheme induces the institutional roles involved in it to behave in a persistent and periodic (cyclic) way. Notice also that the interaction process between Producer and Consumer is characterized by an alternating production-consumption cycle, where only one role performs a set of actions at each time.

B. The ProdCons Coordination Scheme: an IBCC Presentation

There are two implicit steps in the behaviors of the two types of institutional roles presented in Fig. 2, corresponding to their institutional commitments to the coordination scheme.

On one hand, the Producer (resp., the Consumer) should only produce (resp., consume) when it has been given the institutional capability (i.e., institutional right) to do so. On the other hand, the Producer (resp., the Consumer) should produce (resp., consume) when it has the institutional obligation to do so [5].

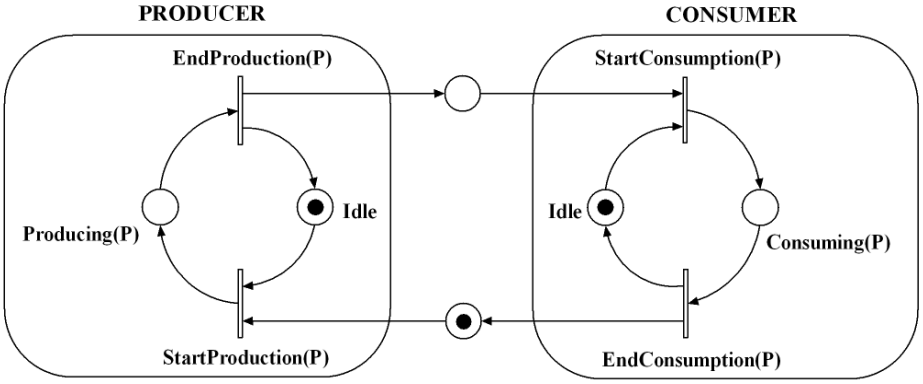

Fig. 2 The Producer-Consumer scheme for coordinating institutional roles

[4] For the sake of space, we do not discuss the overall idea of *schematic institutional building blocks* for agent societies, which will be presented in future work.

[5] We note, about this issue, that we are taking here the term *institutional capability* to mean a *long term* (persistent) feature of institutional roles, while we are taking the term *institutional obligation* to mean a *short term* (transient) feature, referring to momentary obligations.

This distinction reflects in the way institutional capabilities and obligations are dealt with by the institutional commitment rules, in the IBCC model, as shown below.

We remark, however, that any presentation of the ProdCons coordination scheme in terms of the IBCC model of institutional roles should guarantee the correct structuring of the institutional interactions, in order to correctly support the implied institutional commitments.

```
-------------------------------------
Coordination Scheme PRODUCER-CONSUMER
-------------------------------------

INSTITUTIONAL CAPABILITIES:

#Social Role PRODUCER

-- get assignment of capability to produce P
if received msg =
attr_icap(producer,production(P))
then add icap(production(P)) to ICap.

-- get withdrawal of capability to consume P
if received msg =
wthdrw_icap(producer,production(P))
then remove icap(production(P)) from ICap.

#Social Role CONSUMER

-- get assignment of capability to consume P
if received msg =
attr_icap(consumer,consumption(P))
then add icap(consumption(P)) to ICap.

-- get withdrawal of capability to consume P
if received msg =
wthdrw_icap(consumer,consumption(P))
then remove icap(consumption(P)) from ICap.

INSTITUTIONAL COMMITMENTS:

#Social Role PRODUCER

-- establish commitment to producer-consumer
scheme for P
if received msg = req(producer,prod-cons-
scheme(P),consumer)
then add icmt(producer,prod-cons-
scheme(P),consumer) to ICmt.

-- discharge commitment to producer-consumer
scheme for P
if received msg = disch(producer,prod-cons-
scheme(P),consumer)
then remove ocmt(producer,prod-cons-
scheme(P),consumer) from ICmt.

#Social Role CONSUMER

-- establish commitment to producer-consumer
scheme for P
if received msg = req(producer,prod-cons-
scheme(P),consumer)
then add icmt(producer,prod-cons-
scheme(P),consumer) to ICmt.

-- discharge commitment to producer-consumer
scheme for P
if received msg = disch(producer,prod-cons-
scheme(P),consumer)
then remove icmt(producer,prod-cons-
scheme(P),consumer) from ICmt.
```

Fig. 3 An IBCC presentation of the ProdCons institutional coordination scheme

Fig. 3 and Fig. 4 show one such IBCC presentation of the Producer-Consumer coordination scheme defined by the Petri net pictured in Fig. 2.

```
-------------------------------------
Coordination Scheme PRODUCER-CONSUMER
-------------------------------------

INSTITUTIONAL BEHAVIORS:

#Institutional Role PRODUCER

-- start production of an item of P
if icmt(producer,prod-cons-scheme(P),consumer) in
ICmt
   and icap(production(P)) in ICap
   and not producing(P) in IBel
   and received msg = cons_finished(P)
then add producing(P) to IBel.

-- end production of an item of P
if prod_finished(P) in IBel then:
     send consumer msg = prod_finished(P)
     remove producing(P) from IBel.

#Insitutional Role CONSUMER

-- start consumption of an item of P
if icmt(producer,prod-cons-scheme(P),consumer) in
Icmt
   and icap(consumption(P)) in ICap
   and not consuming(P) in IBel
   and received msg = prod_finished(P)
then add consuming(P) to IBel.

-- end consumption of an item of P
if cons_finished(P) in IBel then:
    send producer msg = cons_finished(P)
    remove consuming(P) from IBel.
```

Fig. 4 An IBCC presentation of the ProdCons institutional coordination scheme (cont'd).

Notice that each program command is a commitment rule (see Section III). Notice also that institutional structuring operations that establish the institutional capabilities and institutional commitments of the institutional roles, are treated in the Institutional Capabilities and Institutional Commitments parts of the program, respectively, while institutional behaviors (and the consequent institutional interactions) are treated in the Institutional behaviors part of the program. Finally, notice that in this particular example, both Producer and Consumer roles deal with the institutional structuring operations in the same way.

IX. Related Work

In [5], Castelfranchi and Falcone discuss the many ways in which role may be analyzed, e.g., as power positions, as sets of obligations, or as abstract agents, and explore the idea of a role resulting from a contract, that is, from the acts of delegation and adoption of a task.

From the standpoint we take here, roles are seen as stable components of the organization of an agent society, and formally established. Thus, in the situations we try to formalize, roles pre-exist contracts and, in fact, contracts can only be established between pre-existent roles.

Such notion of role seems to us to be similar to that introduced in [26], where a role was defined as an abstract cognitive structure (with an information and intentional structure) that an agent had to instantiate when implementing the role.

In [26], however, the organizational interfaces of roles were left implicit, since in that respect that paper focused on the pragmatics of the organizational directives that agents could say to each other.

On the other hand, in our architectural view, as in [26], institutional roles only have obligations to perform organizational tasks when institutional commitments are required from them (and are explicitly accepted in view of the existence of the corresponding organizational capabilities).

More recently, a set of multiagent system organizational models have been proposed, where the notion of role has a central position, but each model adopting its own definition of role. In MOISE+ [13], roles as sets of behavior constraints and cooperation patterns to be imposed on the agents that adopt those roles.

In Electronic Institutions [17], roles define both the valid sequences of dialogic interactions that agents can perform to achieve their goals when playing their roles in scenes of the performative structure, and are attached to rule-based norms that specify a dependency relation between actions performed in different situations of the performative structure. [25] treated preliminary issues regarding the implementation of norms and the connection between roles and norms in electronic institutions.

The PopOrg model [8, 9, 10] takes institutional roles as the building blocks of the formal organization structure of agent societies. An institutional role is defined as the set of all behaviors that an agent can possibly perform (i.e., can legitimately perform in the context of the organization) while implementing that role.

Behaviors in the PopOrg model are defined to be traces of actions. Thus, the architectural approach to institutional roles introduced here may prove to fit well with the PopOrg approach, if one succeeds in correlating the set of all possible behaviors that a institutional role may perform at a given time with the the set of all possible performances of commitment rules that the IBCC model of the institutional role allows at that instant.

One particular work that bears some similarity to the work introduced here is that in [2, 1] where an organizational notion of role was adopted, in the powerJava programming language, to allow for an operational notion of pattern modeling relationships between objects. Such operational notion of role supports an architectural view of role as the one introduced here.

X. Conclusion

From the point of view adopted in this paper, institutional roles should be taken as power positions within the organization of agent societies, whose obligations (institutional commitments) and rights (organizational capabilities) vary according to the dynamics of operations of requirement/discharge and assignment/withdrawal, respectively.

From such point of view, the paper introduces an architectural approach to support the modeling of the cognitive structure and organizational interface that such power positions require from the agents that implement them.

This paper was not the place to discuss the whole set of issues involved either in the creation, attribution and withdrawal of rights and obligations (cf. [21]), and the companion issues like legitimacy, disputes, etc. (cf. [6]), or in the establishment and fulfillment of commitments (cf. [4]). We remark, however, that the two notions of right and commitment that we have essayed to embed in the model should be thought of as formally institutionalized at the organization structure level of the agent societies through institutional processes (cf. [20] and [24]) that we still have to explore in our work, mainly within the framework of the PopOrg model [8, 9, 10].

The correlation between the PopOrg model of agent societies and the model of institutional roles happens to be the work we intend to do in the immediate future in connection to the results of the present paper.

The primary aim of the whole work, of course, is the development of concepts and methodologies to support the modeling and simulation of multiagente systems organizations.

Acknowledgment

This work partially supported by CNPq and FAPERGS. The author thanks G. P. Dimuro for comments and suggestions.

References

[1] Matteo Baldoni, Guido Boella, and Leendert van der Torre. Relationships meet their roles in object oriented programming. In F. Arbab and M. Sirjani, editors, FSEN 2007 - International Symposium on Fundamentals of Software Engineering. Springer (LNCS.

[2] Matteo Baldoni, Guido Boella, and Leendert van der Torre. Interaction between objects in powerjava. Journal of Object Tecnhology, 2(2), 2003.

[3] Cristiano Castelfranchi. Commitments: From individual intentions to groups and organizations. In Victor Lesser and Les Gasser, editors, Proceedings of the First International Conference on Multiagent Systems - ICMAS 95, pages 41–48, Cambridge, 1995. MIT Press.

[4] Cristiano Castelfranchi. Formalising the informal? dynamic social order, bottom-up control, and spontaneous normative relations. Journal of Applied Logic, 1(1):47–92, 2003.

[5] Cristiano Castelfranchi and Rino Falcone. From task delegation to role delegation. In AI*IA 97: Advances in Artificial Intelligence, LNCS 1321, pages 278–289. Springer, Berlin, 1997.

[6] James S. Coleman. Foundations of Social Theory. Havard Univ. Press, Cambridge, 1994.

[7] Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro. Introducing service schemes and systems organization in the theory of interactive computation. In Arnold Beckmann, Costas Dimitracopoulos, and Benedikt Löwe, editors, Logic and Theory of Algorithms, Fourth Conference on Computability in Europe, CiE 2008, pages 87–96, Athens, 2008.

[8] Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro. Semantical concepts for a formal structural dynamics of situated multiagent systems. In J. Sichman, P. Noriega, J. Padget, and

S. Ossowski, editors, Coordination, Organizations, Institutions, and Norms in Agent Systems III, number 4870 in LNAI, pages 139–154. Springer, Berlin, 2008.

[9] Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro. Introducing social groups and group exchanges in the PopOrg model. In Proceedings of AAMAS 2009, volume 1, pages 1297–1298, Budapest, 2009. IFAAMAS.

[10] Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro. A minimal dynamical organization model. In V. Dignum, editor, Multi-Agent Systems: Semantics and Dynamics of Organizational Models, pages 419–445. IGI Global, Hershey, 2009.

[11] Antônio Carlos da Rocha Costa and Graçaliz Pereira Dimuro. On the interactional account of the social functions of agent societies. In Proceedings of the BWSS 2010 - 2nd Brazilian Workshop on Social Simulation, São Bernando do Campo, 2010. SBC - Sociedade Brasileira de Computação.

[12] Virginia Dignum, editor. Multi-Agent Systems - Semantics and Dynamics of Organizational Models. IGI Global, Hershey, 2009.

[13] Jomi F. Hübner, Jaime S. Sichman, and Olivier Boissier. Developing organised multi-agent systems using the MOISE+ model: Programming issues at the system and agent levels. International Journal of Agent-Oriented Software Engineering, 1(3-4):370–395, 2007.

[14] Carl Hewitt. Viewing control structures as patterns of passing messages. Artificial Intelligence, 8(3), June 1977.

[15] John Laird, Paul Rosenbloom, and Allen Newell. Soar: An architecture for general intelligence. Artificial Intelligence, 33(1):1–64, September 1987.

[16] Bronislaw Malinowski. A Scientific Theory of Culture, and other essays. The University of North Carolina Press, 1944.

[17] Pablo Noriega, Carles Sierra, and Juan Rodríguez. The Fishmarket Project. Reflections on agent-mediated institutions for trustworthy e-commerce. In Proc. Work. Agent-mediated Electronic Commerce, Intl. Conf. Electronic Commerce, Seoul, 1998.

[18] Carl Adam Petri. Introduction to Net Theory. In W. Brauer, editor, Net Theory and Applications, number 84 in LNCS, pages 1–19. Berlin, Springer-Verlag, 1980.

[19] Anand S. Rao and Michael P. Georgeff. Bdi agents: from theory to practice. In Proceedings of ICMAS 95, pages 312–319, San Francisco, June 1995.

[20] W. Richard Scott. Institutions and Organizations - Ideas and Interests. Sage Publications, Los Angeles, 2008.

[21] John R. Searle. The Construction of Social Reality. The Free Press, New York, 1995.

[22] Yoav Shoham. Agent oriented programming. Artificial Intelligence, 60(1):51–92, 1993.

[23] Munindar P. Singh. An ontology for commitments in multiagent systems: Toward a unification of normative concepts. Artificial Intelligence and Law, 7(1):97–113, 1999.

[24] Luca Tummolini and Cristiano Castelfranchi. The cognitive and behavioral mediation of institutions: Towards an account of institutional actions. Cognitive Systems Research, 7:307–326, 2006.

[25] Javier Vazquez-Salceda, Huib Aldewereld, and Frank Dignum. Norms in multiagent systems: some implementation guidelines. In Second European Workshop on Multi-Agent Systems, pages 737–748, Barcelona, 2004.

[26] Eric Werner. Toward a theory of communication and cooperation for multiagent planning. In Theoretical Aspects of Reasoning about Knowledge, pages 129–143. Morgan-Kaufman, 1988.