

Specifying and Reasoning with Institutional Agents

Filipe Santos

Department of Information Science, ISCTE

Av. das Forças Armadas,
1649-026 Lisboa Codex, Portugal
Tel: (+351) 217903099

e-mail: Filipe.Santos@iscte.pt

Olga Pacheco

Department of Informatics

University of Minho
4710-057 Braga, Portugal
Tel: (+351) 253 604445

e-mail: omp@di.uminho.pt

ABSTRACT

This paper proposes a logic-oriented framework for institutional agents specification and analysis. Within this framework institutional agents are seen as artificial agents that aggregate a set of (real) agents, being capable to act (by means of real agents' actions) and to whom deontic qualifications may be associated. In order to analyse how institutional agents interact with the external world, the following aspects are characterised: how the obligations flow from the institutional agent to the real agents that support him, and how the actions of the latter count as actions of the former. The fundamental concept supporting these aspects is the concept of role.

Institutional agents are specified and analysed by means of a first-order role based deontic/action modal logic. The analysis is automated by means of a tableaux theorem proving method extended with additional rules to deal with the proposed classical action and deontic modalities.

Categories and Subject Descriptors

I.2.1 [Artificial Intelligence]: Applications and Expert Systems – law, organization automation.

I.2.3 [Artificial Intelligence]: Deduction and Theorem Proving – deduction, inference engines.

I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods – modal logic.

K.4.3 [Computers and Society]: Organizational Impacts – automation.

F.4.1 [Mathematical Logic and Formal Languages]: Mathematical Logic – modal logic, mechanical theorem proving, proof theory.

General Terms

Design, Legal Aspects, Experimentation, Verification.

Keywords

Tableaux, deontic/action modal logic, institutional agent, role, normative specification, representation.

1. INTRODUCTION

We are particularly interested in the formal characterisation and analysis of organisations, by using modal logic as an analytical tool. By *organisation* we mean a society of agents, comprising human or artificial agents (e.g. software agents or other organisations), whose rule-governed interactions are aimed at some specified tasks or goals. We adopt the perspective proposed in [24], seeing organizations as instances of *normative systems*, and thus expressing their activity and interaction in terms of what agents are permitted and obliged to do (allowing also the possibility that their behavior may deviate from the ideal), and in terms of other complex normative relations between them. One principal advantage of adopting this perspective is that it enables us to capitalise on a lot of the work that has been done on the application of formal-logical techniques to the analysis of law, legal systems and social systems (see, e.g., [18] [19] [22] [29] [30] [31] [32])).

There had been a considerable amount of research in order to develop appropriate modeling techniques to the automation of organisations (see, e.g., [1] [7] [20] [21] [28] [41]). Although it is recognised that the success of the automation of organisations depends on the adoption of explicit organisational models, there is a lack of approaches that provide a rigorous description of the meaning of organisational concepts (e.g., *role*, *right*, *permission*, *obligation*, *authority*, *authorisation*, *responsibility* and *delegation*). Consequently the interpretation of organisational concepts mainly depends on the intuitions of the users of such techniques, with obvious impact to the automation of organisations. Such ambiguity obstructs any rigorous specification and systematic analysis of organisations.

We believe that organisations can fruitfully be analysed using a set of action, deontic and other relevant modalities as basic building blocks to be used in the characterization of organisational notions. Our main research focuses on the logical characterisation of relevant concepts for organisation specification and analysis, and on the automation of the resulting multi-modal logic systems [3] [4] [5] [25] [26] [27] [36] [37] [38] [39] [40].

From their digression on Law, Pacheco and Carmo proposed a role based model for the normative specification of organised collective agency and agents' interaction. They stressed the need for the concept of *institutional agent* in order to deal with normative specification of collective agents. Institutional agents are seen as artificial agents that aggregate a set of (real) agents, being capable to act as a whole (by means of real agents' actions) and to whom deontic qualifications may be associated. This concept of *institutional agent* together with the concepts of *role* and *acting in a role* (playing a role) are used as basic ingredients for the

specification of collective agents, such as organisations. The characterisation of these concepts is made through a first-order role based deontic/action modal logic.

We here proceed by proposing a theorem proving method for such logic. The method is based on a first-order semantic tableaux extended with additional rules to deal with the proposed classical action and deontic modalities. This is our first step towards the automation of a systematic analysis of organisations.

An overview of the rest of the paper follows. In Section 2, we discuss our logical characterisation of the main concepts to be used on the specification of institutional agents. In Section 3, we discuss how we specify institutional agents and we illustrate how it can be used to support some interesting aspects of the organization's activity and analysis, through some simple examples. In Section 4, we describe the automated theorem proving method for the proposed logic. Conclusions and directions for future work appear in Section 5.

2. ROLE-BASED ACTION AND DEONTIC LOGIC

Our logic follows the tradition initiated by Kanger, Pörn and Lindahl of combining deontic and action logics as basic building blocks to describe social interaction and complex normative concepts (see, e.g., [18] [19] [22] [29] [30] [31] [32])). Their logics have sufficient expressive power to be able to articulate several distinctions at an appropriate abstract level, mainly in virtue of the modal logic of action they employ. They introduced a relativised modal operator, here designated by E_x , where expressions of the form E_xA are read “agent x brings it about that A ” or “agent x sees to it that A is the case”. An important feature of these logics is that actions are taken to be relationships between agents and states of affairs that they bring about, setting aside temporal aspects. (For a brief overview of the action logics proposed by these author, see [38].)

Although the formal properties assigned to the action operator vary among these authors, their logical systems all have in common the following two axiom schemas:

$$(T) \quad E_xA \rightarrow A$$

$$(C) \quad (E_xA \wedge E_xB) \rightarrow E_x(A \wedge B)$$

and the rule of inference:

$$(RE) \quad \text{If } \vdash A \leftrightarrow B \text{ then } \vdash E_xA \leftrightarrow E_xB^1$$

The (T) schema captures the intuition that if agent x brings it about that A , then A is indeed the case. The (C) schema represents the idea that an agent brings it about all he brings it about separately. (RE) is just closure under logical equivalence.

This approach to the logic of action offers a good expressive power; for instance, it facilitates the expression of the different atomic positions an agent might be in with respect to a particular state of affairs A : E_xA , $E_x\neg A$ and $\neg E_xA \wedge \neg E_x\neg A$ (the last of

these meaning “ x remains passive with respect to A ”); moreover, the importance of this operator to the characterisation of norms is brought out by the distinctions it affords when combined with a deontic operator, as has been shown in, for instance, the work of Kanger [17] and Lindahl [22]. As an example of the kind of analysis this combination of operators provides (assuming the adoption of Standard Deontic Logic - SDL), we list the seven *normative one-agent act-positions* (where the operator O represents “it is obligatory that” and its dual P represents “it is permitted that”). For any agent x , and for any state of affairs A , precisely one of the following positions obtains (cf. [16]):

$$(T1) \quad PE_xA \wedge PE_x\neg A \wedge P(\neg E_xA \wedge \neg E_x\neg A)$$

$$(T2) \quad PE_xA \wedge O\neg E_x\neg A \wedge P(\neg E_xA \wedge \neg E_x\neg A)$$

$$(T3) \quad PE_xA \wedge PE_x\neg A \wedge O(E_xA \vee E_x\neg A)$$

$$(T4) \quad O\neg E_xA \wedge PE_x\neg A \wedge P(\neg E_xA \wedge \neg E_x\neg A)$$

$$(T5) \quad OE_xA$$

$$(T6) \quad O(\neg E_xA \wedge \neg E_x\neg A)$$

$$(T7) \quad OE_x\neg A$$

In spite of the expressiveness of this action operator, it is not able to distinguish situations where an agent brings about a state of affairs in different roles for three main reasons: 1) to precisely know the effects of the actions (e.g., its juridical consequences); 2) to precisely evaluate the deontic qualifications of agents with respect to the actions they bring about; and 3) for authentication issues (cf. [3]).

Lets consider the following situations that illustrate these aspects:

(a) Agent x brings it about that B in the quality of administrator of the firm y .

(b) Agent x brings it about that B on its own behalf.

The attempt to represent both situations by E_xB is not adequate since they are different in several aspects:

1) *The effects of the acts of agent x .* In case (a) x 's action also affects institutional agent y , because when x acts as administrator of y he also acts *on behalf* of y . The role of *administrator of y* played by x is a representative role². However, in case (b) only x is affected by its own acts.

2) *The deontic qualifications of agents' acts are different.* Agent x may be authorized or obliged to bring about B as *administrator* without being authorized or obliged to bring about B in another role. So, to evaluate agent's action we have to take in consideration the *role* in which he performed that action.

3) *Authentication issues.* An agent must be authorized to act in a particular role, i.e., that role must have been formally attributed to him. In most situations, it is not enough that an agent holds a particular role. He may have to prove that he holds it. An action concept that explicitly indicates what is the role an agent plays when he acts should be helpful for authentication issues.

¹ Considering propositional logic as the underlying non-modal logic, these principles form a classical system of type ECT. (according to the classification of Chellas [6]).

² See below a brief discussion of the notion of representation between agents.

The concept of role has many different meanings, depending on the context where it is used. We will try to clarify, briefly, what do we mean by role.

Roles are qualities that agents might have that are relevant when we consider agents acting and interacting with other agents. Consider, for instance, an agent playing the role of a teacher, or an agent playing the role of an administrator of a firm.

We assume that an agent always acts playing some role (and if no specific quality is being exercised we say that he acts in the role of itself). Associated to a role there are obligations, permissions (or other deontic concepts), describing the expected behavior of the holders of that role, that is the abstract state of affairs that they are obliged or permitted to bring about. To act in a role an agent must be qualified to play that role (e.g. that role has been formally attributed to him by a contract). An agent may hold several roles but it can play only a role at a time.

Some roles represent relationships that agents have with other agents. For instance, when we say that “x acts as administrator of a company y”, we mean that agent x acts in the role of “administrator of y”, and we will represent that role by *administrator(y)* expressing the fact that being administrator of y is a kind of relationship with y

Some of these roles are *representative roles* in the sense that agents playing those roles (representative agents) are permitted to act on behalf of the represented agents. As an example, if “administrator(y)” is a representative role, when “x acts in the role of administrator of y” his actions count as actions of y.

Those aspects suggest the need for the new notion of *acting in a role*. We propose the use of a monadic operator $E_{x:r}$ in order to capture this agency concept, where expressions of the form $E_{x:r} A$ are read “agent x, playing the role r, brings it about that A”. Using this operator we may distinguish the previous situations by

- (a) $E_{x:\text{administrator}(y)} B$
- (b) $E_{x:\text{itself}} B$

where *itself* represents the role of an agent acting on its own behalf.

It is now possible to characterize the representative role played by x (when he brings about B) as administrator of y:

$$E_{x:\text{administrator}(y)} B \rightarrow E_{y:\text{itself}} B,$$

which captures the idea that the action of x as representative of y (when playing the role of administrator), counts as an action of y. Thus we may infer $E_{y:\text{itself}} B$ in situation (a). We may also properly characterize that an agent x is obliged to bring about B as *administrator* without being obliged to bring about B on its own behalf (for a suitable deontic operator O):

$$OE_{x:\text{administrator}(y)} B \wedge \neg OE_{x:\text{itself}} B.$$

Of course we are assuming that whenever agent x brings it about that B in the quality of administrator of y, x is formally qualified to play that role, i.e.,

$$E_{x:\text{administrator}(y)} B \rightarrow \text{is-administrator}(x,y).$$

The logical characterisation of these concepts is based on a first-order many-sorted logic [8] extended with action and deontic modalities. See [5] for technical details. Herein it is enough to

mention that three special sorts are used in the definition of this logic: Ag (the agent sort), R (the role sort) and AgR (the agent playing a role sort).

Starting from the logical properties of the action operator $E_{x:r}$, we adopt the following axiom schemas:

$$(T_E) \quad E_{x:r} A \rightarrow A$$

$$(C_E) \quad (E_{x:r} A \wedge E_{x:r} B) \rightarrow E_{x:r} (A \wedge B)$$

and the rule of inference:

$$(RE_E) \quad \text{If } \vdash A \leftrightarrow B \text{ then } \vdash E_{x:r} A \leftrightarrow E_{x:r} B$$

In order to deal with the mentioned authentication issues we also assume that an agent that brings it about some state of affairs, acting in a particular role, must be qualified to play that role, and thus we further adopt the following axiom schema:

$$(Qual) \quad E_{x:r} A \rightarrow \text{is-r}(x)$$

for is-r a predicate associated with each role r, where is-r(x) expresses that “agent x is qualified to play the role r”.³ The only role that an agent is always qualified to play is the role *itself*. Thus we have:

$$(Itself) \quad (\forall x^{Ag}) \text{is-itself}(x)$$

With respect to the deontic operator O we simply adopt the following rule of inference:

$$(RE_O) \quad \text{If } \vdash A \leftrightarrow B \text{ then } \vdash OA \leftrightarrow OB$$

This basic property is enough for the purposes of this paper. See, e.g., [2] [15] [23] for a detailed discussion of this concept.

In the next section we will show how to use the previous notions and their logical principles in the characterisation and analysis of institutional agents.

3. INSTITUTIONAL AGENTS SPECIFICATION AND ANALYSIS

The concept of *institutional agent*, proposed by Pacheco and Carmo, intends to model organised collective entities, like organisations, and is based on the legal concept of *artificial person*. Artificial persons are collective entities that aggregate several natural persons (human beings) allowing them to collectively pursue some interests. They have *juridical personality*, which means that they may be the subject of obligations and rights and they also have *legal qualification*, which means that they can exercise their rights and be responsible for the unfulfilment of their obligations.

³ Sometimes roles are represented with parameters in order to capture relevant relationships, for instance *administrator(y)*. In this case *is-administrator(x,y)* expresses “agent x is qualified to play the role *administrator(y)*”. We omit here the technical details of the relationship between the role r and the predicate is-r (see [5] for more details).

An artificial person has a stable structure formed by a set of roles and a set of rules regulating their behavior⁴. This structure is supported by persons, the holders of the roles. The artificial person acts through them. The holders of some roles have power to act on behalf of the artificial person: those roles are called *representative* roles, because the acts of its holders (when playing that role) count as acts of the artificial person. Since the fulfillment of obligations presuppose the performance of actions and since an artificial person, as abstract entity, cannot act directly, acting through the holders of its roles, there exists mechanisms stressing how obligations flow from the artificial person to those holders and how the actions performed by the holders count as actions performed by the artificial person. Law provides general figures of interaction between persons that regulate the relationship between a person that is holder of a role in an artificial person and that artificial person: the *mandate* and *representation* legal relationships. A detailed presentation of artificial persons and related legal concepts can be found in [27].

Based on the legal concepts just presented, the concept of institutional agent was proposed. Herein we just want to summarise the characterisation of institutional agents based on those legal concepts.

First of all, in spite of being collective entities, institutional agents are agents. They interact in a society like any other agent: they can establish normative (or other kind of) relationships with other agents (e.g. contracts), they can hold roles, they may be the subject of obligations or other normative concepts, and may be responsible for the unfulfilment of obligations or any other non ideal situation. An institutional agent has a stable structure formed by a set of roles and a set of norms defining the deontic characterisation of each role. This structure is supported by other agents: the holders of its roles. The agents that hold a particular role inherit the deontic characterization of that role, in the sense that, when acting in that role, their behavior will be evaluated according to that deontic characterization.

There must be defined a mechanism of transmission of the obligations of an institutional agent to the roles of its structure (and indirectly to the holders of those roles), stating who is responsible for the fulfilment of those obligations (c.f. point (2) below). There also must be defined what are the representative roles of the institutional agent and its respective scope of representation, stating who is authorised to act on behalf of the institutional agent and to what extent (c.f. point (3) below). The institutional agent acts through the holders of the roles of its structure. The holders of the roles of the institutional agent may change without affecting the identity of the former. The holders of the roles of an institutional agent are not necessarily human agents. They may be software agents or other institutional agents.

We will now extend the logical language with some abbreviations that will support the specification of institutional agents.

(1) The deontic characterization of roles in institutional agents is independent from the agents that hold those roles in a particular moment. So, although deontic notions are meaningful only when

applied to agents, it is useful to introduce deontic operators indexed by roles, seen as applied to the holders of those roles:

$$(Or) \quad O_r A \equiv (\forall x)(is-r(x) \rightarrow OE_{x:r}A)$$

$$(Pr) \quad P_r A \equiv (\forall x)(is-r(x) \rightarrow PE_{x:r}A)$$

$$(Fr) \quad F_r A \equiv (\forall x)(is-r(x) \rightarrow FE_{x:r}A)$$

(2) Using the above abbreviations we can express the transmission of obligations from the institutional agent to specific roles of its structure, and thus, indirectly, to the agents that support that role:

$$OE_{i:itself}A \rightarrow Org(i)A,$$

where $rg(i)$ is a role of the structure of the institutional agent i , meaning that whenever i is under the obligation of bringing about A , that obligation is attributed to the role $rg(i)$ (i.e. to the holders of that role).

(3) We have mentioned before that some roles may be representative roles of other agents. This means that the holders of that kind of roles, when acting in those roles, act on behalf of the represented agents within the scope of representation defined for those roles. In order to express representative roles, the following notation is introduced: $r:REP(a,B)$, that is read as follows: r is a representative role of agent a with the scope of representation B . This expression is seen as the following abbreviation:

$$(REP) \quad r:REP(a,B) \equiv (\forall x)(E_{x:r}B \rightarrow E_{a:itself}B)$$

Note that when an agent acts as representative of another agent he does not act on his own behalf, which is expressed by the following property:

$$(E_{x:r}B \wedge r:REP(a,B)) \rightarrow \neg E_{x:itself}B$$

Representative roles are crucial for institutional agents, because an institutional agent cannot act directly, needing other agents to act on its behalf. Those agents are the holders of the representative roles of the institutional agent structure. So, it's necessary to identify in an institutional agent model which are the representative roles and their respective scope of representation.⁵

We can now use the abbreviations we have just introduced, in the formal specification of institutional agents. The specification of an institutional agent involves a name, i , and a structure:

$$ST_i = \langle R_i, DCR_i, TO_i, RER_i, TIT_i \rangle,$$

where:

R_i : *a set of roles* – the structural roles of the institutional agent. It is constituted by a finite set of atomic formulas of the form $is\text{-}role\text{-}str(rg(i))$, stating that $rg(i)$ is a role of the structure of i ;

DCR_i : *deontic characterisation of structural roles*. It is constituted by a finite set of formulas of the form $Org(i)A$, $Prg(i)A$ or $Frg(i)A$, where $rg(i)$ is a role of the structure of i ;

TO_i : *transmission of obligations from the institutional agent to specific roles of its structure*. It is constituted by a set of formulas

⁴ To be precise, the rules regulate the behavior of the holders of those roles. Roles don't act so we cannot associate behavior to a role.

⁵ Representative roles are not necessarily roles of the structure of an institutional agent. They may result from contracts or other normative relations that agents may establish between each other.

of the form $OE_{i:\text{itself}A} \rightarrow O_{\text{rg}(i)}A$, for $\text{rg}(i)$ a role of the structure of i ;

RER_i : *representative roles of i and its respective scope of representation*. It is constituted by a set of formulas of the form $\text{rg}(i):\text{REP}(i, B)$;

TIT_i : *titularity of structural roles of i* . It is constituted by a set of formulas of the form $\text{is-rg}(x, i)$, stating that agent x holds (is titular of) the structural role $\text{rg}(i)$.

We present below an example of specification of an institutional agent ax . It specifies an association named ax , whose structure is formed by the roles:

$PAdm(ax)$ – president of administration of ax ;

$MAdm(ax)$ – member of administration of ax ;

$PWC(ax)$ – president of the watch committee of ax ;

$MWC(ax)$ – member of the watch committee of ax ;

$Ass(ax)$ – associate of ax .

$ST_{ax} = \langle R_{ax}, DCR_{ax}, TO_{ax}, RER_{ax}, TIT_{ax} \rangle$

$R_{ax} = \{ \text{is-role-str}(PAdm(ax)), \text{is-role-str}(MAdm(ax)), \\ \text{is-role-str}(PWC(ax)), \text{is-role-str}(MWC(ax)), \\ \text{is-role-str}(Ass(ax)) \}$

$DCR_{ax} = \{ O_{MAdm(ax)P1}, P_{MAdm(ax)P2}, O_{PAdm(ax)P2}, \\ O_{PAdm(ax)P3}, P_{PAdm(ax)P4}, O_{PWC(ax)P5}, \\ P_{PWC(ax)P6}, O_{MWC(ax)P7}, O_{Ass(ax)P8}, P_{Ass(ax)P9} \}$

$TO_{ax} = \{ OE_{ax:\text{itself}P1} \rightarrow O_{MAdm(ax)P1}, \\ OE_{ax:\text{itself}P2} \rightarrow O_{PAdm(ax)P2}, \\ OE_{ax:\text{itself}P4} \rightarrow O_{PAdm(ax)P4} \}$

$RER_{ax} = \{ MAdm(ax): \text{REP}(ax, p1), PAdm(ax): \text{REP}(ax, p2), \\ PAdm(ax): \text{REP}(ax, p4) \}$

$TIT_{ax} = \{ \text{is-Ass}(a, ax), \text{is-Ass}(b, ax), \text{is-Ass}(c, ax), \text{is-Ass}(d, ax), \\ \text{is-PAdm}(e, ax), \text{is-MAdm}(d, ax), \text{is-PWC}(f, ax), \\ \text{is-MWC}(g, ax) \}$

This institutional agent is part of a society of agents and interacts with other members of the society. For simplicity reasons and because it is not central to this paper, we do not specify the external society, considering only the agents that hold structural roles of the institutional agent (c.f. TIT_{ax}).

To conclude this section, let us now illustrate the kind of analysis supported by this specification. A specification of an institutional agent ST_{ax} defines a particular logical language and a set of formulas of such language. Let us call $T(ST_{ax})$, the logic obtained by adding this set of formulas to the underlying deontic/action logic here proposed, as new axioms.

Consider now two simple situations:

Case 1: Two representative agents of agent ax , d and e , act on behalf of ax , within their scope of representation. It is possible to infer in $T(ST_{ax})$ that their actions count as actions of ax :

$\{ Ed:MAdm(ax)P1, Ee:PAdm(ax)P2 \} \vdash_{T(ST_{ax})} E_{ax:\text{itself}(P1 \wedge P2)}$

Sketch of the proof:

1. $Ed:MAdm(ax)P1$ hypothesis
2. $MAdm(ax): \text{REP}(ax, p1)$ RER_{ax}
3. $E_{ax:\text{itself}P1}$ from (1), (2) and (REP)
4. $Ee:PAdm(ax)P2$ hypothesis
5. $PAdm(ax): \text{REP}(ax, p2)$ RER_{ax}
6. $E_{ax:\text{itself}P2}$ from (4), (5) and (REP)
7. $E_{ax:\text{itself}(P1 \wedge P2)}$ from (3), (6) and (C_E)

Case 2: As a result of some interaction with the external society (e.g. some established contract), ax is under the obligation to bring about $p4$. It is possible to infer in $T(ST_{ax})$ that this obligation is transmitted to agent e .

$\{ OE_{ax:\text{itself}P4} \} \vdash_{T(ST_{ax})} OE_e:PAdm(ax)P4$

Sketch of the proof:

1. $OE_{ax:\text{itself}P4}$ hypothesis
2. $OE_{ax:\text{itself}P4} \rightarrow O_{PAdm(ax)P4}$ TO_{ax}
3. $O_{PAdm(ax)P4}$ from (1) and (2)
4. $\text{is-PAdm}(e, ax)$ TIT_{ax}
5. $OE_e:PAdm(ax)P4$ from (3), (4) and (Or)

The first example illustrates how the actions of agents that hold roles of the structure of an institutional agent may affect the institutional agent (how to pass from the micro-level of individuals to the macro-level of the organization). The second example shows how an obligation of an institutional agent may be transmitted to agents that hold roles in its structure (how to pass from the macro-level to the micro-level).

4. AUTOMATED THEOREM PROVING

There exist several proposals for the automation of *normal modal* logics (i.e., modal logics with the rule of inference (RK_{\square}): If $\vdash(A_1 \wedge \dots \wedge A_n) \rightarrow A$ then $\vdash(\square A_1 \wedge \dots \wedge \square A_n) \rightarrow \square A, n \geq 0$): methods based on *semantic tableaux* (see, e.g., [10] [11] [12]); methods based on *resolution* (see, e.g., [9]); methods based on *sequent calculus* (see, e.g., [14]). As far as we know, there is no method for *classical modal* logics, i.e., modal logics with the rule of inference (RE_{\square}): if

$\vdash A \leftrightarrow B$ then $\vdash \square A \leftrightarrow \square B$.

The analysis presented in the previous section is actually supported by an automated theorem prover based on free variable first-order semantic tableaux (see, e.g. [13]), extended with additional rules to deal with the classical modal logics used to characterize the deontic

and action operators and their relationships.

Our tableaux method uses the usual expansion rules for propositional logic, i.e., the rules:

$$(R\neg) \frac{\neg\neg A}{A} \quad (R\text{ True}) \frac{\neg\text{True}}{\text{False}} \quad (R\text{ False}) \frac{\neg\text{False}}{\text{True}}$$

$$(R\alpha) \frac{\alpha}{\alpha_1} \quad (R\beta) \frac{\beta}{\beta_1 \mid \beta_2}$$

$$\alpha_2$$

together with the following free variable quantifier expansion rules for first-order logic:

$$(R\gamma) \frac{\gamma}{\gamma(x)}, \text{ for an unbound variable } x \text{ in the tableaux.}$$

$$(R\delta) \frac{\delta}{\delta(f(x_1, \dots, x_n))}$$

for f a new Skolem function symbol and x_1, \dots, x_n all the used free variables occurring on the branch.

where α and β denote respectively conjunctive and disjunctive formulas of the form $\alpha_1 \wedge \alpha_2$ and $\beta_1 \vee \beta_2$, and γ and δ denote respectively universal and existential formulas of the form $(\forall x)\gamma(x)$ and $(\exists x)\delta(x)$, according to the following tables:

Table 1: α and β - formulas and components

conjunctive formulas			disjunctive formulas		
α	α_1	α_2	β	β_1	β_2
$(A \wedge B)$	A	B	$\neg(A \wedge B)$	$\neg A$	$\neg B$
$\neg(A \vee B)$	$\neg A$	$\neg B$	$(A \vee B)$	A	B
$\neg(A \rightarrow B)$	A	$\neg B$	$(A \rightarrow B)$	$\neg A$	B
$(A \leftrightarrow B)$	$(A \rightarrow B)$	$(B \rightarrow A)$	$\neg(A \leftrightarrow B)$	$\neg(A \rightarrow B)$	$\neg(B \rightarrow A)$

Table 2: γ and δ - formulas and instances

universal formulas		existential formulas	
γ	$\gamma(t)$	δ	$\delta(t)$
$(\forall x)A$	$A\{x/t\}$	$(\exists x)A$	$A\{x/t\}$
$\neg(\exists x)A$	$\neg A\{x/t\}$	$\neg(\forall x)A$	$\neg A\{x/t\}$

The method is further extended with the following rules that correspond to the schemes and inference rules proposed in section 2 (where expression inside square brackets refers the conditions for the application of the rule):

$$(RT-E) \frac{E_{x:r}A}{A}$$

$$(RC-E) \frac{\neg E_{x:r}\alpha}{\neg E_{x:r}\alpha_1 \mid \neg E_{x:r}\alpha_2}$$

$$(RRE-E) \frac{\neg E_{x:r}A, E_{x:r}B}{\text{False}} [\vdash A \leftrightarrow B]$$

$$(RRE-O) \frac{\neg OA, OB}{\text{False}} [\vdash A \leftrightarrow B]$$

$$(RQual) \frac{E_{x:r}A}{\text{is-r}(x)}$$

$$(RItself) \frac{}{\text{is-itself}(x)}, \text{ for an unbound variable } x \text{ in the tableaux.}$$

Note that (RC-E) is just applied to α formulas. In order to apply (RC-E) during tableau construction, each time a formula of the form $\neg E_{x:r}C$ is detected, C is converted to the clause form $(C_1 \wedge \dots \wedge C_n)$. If $n > 1$ then (RC-E) is applied considering $\alpha_1 = (C_{i_1} \wedge \dots \wedge C_{i_k})$ and $\alpha_2 = (C_{i_{k+1}} \wedge \dots \wedge C_{i_n})$, where $1 \leq k < n$ and $\{i_1, \dots, i_n\}$ a permutation of $\{1, \dots, n\}$.

Moreover, to deal with the abbreviation proposed in section 3, we also use the following expansion rules:

$$(ROr) \frac{O_r A}{(\forall x)(\text{is-r}(x) \rightarrow OE_{x:r}A)}$$

$$(RPr) \frac{P_r A}{(\forall x)(\text{is-r}(x) \rightarrow PE_{x:r}A)}$$

$$(RFr) \frac{F_r A}{(\forall x)(\text{is-r}(x) \rightarrow FE_{x:r}A)}$$

$$(RREP) \frac{r:\text{REP}(a,A)}{(\forall x)(E_{x:r}A \rightarrow E_a:\text{itself}A)}$$

Finally, the next expansion rule is used to deal with the abbreviations proposed for the modal operators P and F:

$$(R\sigma) \frac{\sigma}{\text{abv}(\sigma)}$$

where σ denotes a *role free* deontic formula and $\text{abv}(\sigma)$ its equivalent *role free* obligation abbreviation, according to table 3:

Table 3: σ - formulas and abbreviations

σ	$\text{abv}(\sigma)$
FA	$O\neg A$
$\neg FA$	$\neg O\neg A$
PA	$\neg O\neg A$
$\neg PA$	$O\neg A$

Using the previous rules, for concluding that $\Delta \vdash B$, we must find a closed Δ -tableau for $\{\neg B\}$, i.e. a tableau allowing also the Δ -introduction rule:

$$(R\Delta) \frac{}{A} [A \in \Delta]$$

Since we are using free variable quantifier rules, when we consider branch closure, we must also apply the following *tableau substitution rule*: if T is a tableau for the set of formulas Ω , and the substitution σ is free for every formula in T, then $T\sigma$ (the result of replacing each formula A in T by $X\sigma$) is also a tableau for Ω .

Let's consider the following example of a tableau proof for $\{r1:\text{REP}(b,A), r2:\text{REP}(b,B), (\exists x)E_{x:r1}A, (\exists x)E_{x:r2}B\} \vdash E_{b:\text{itself}}(A \wedge B)$ pictured in figure 1 (with numbers added for

reference).

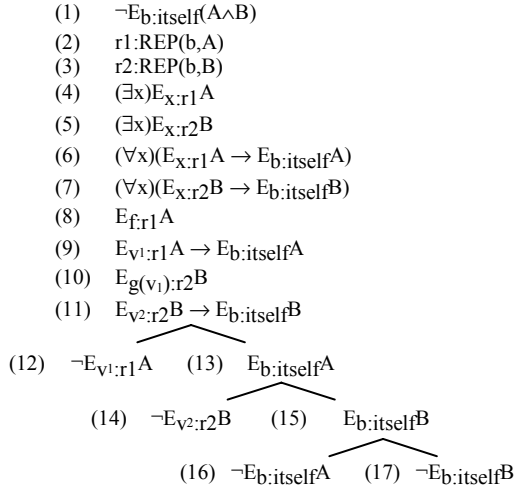


Figure 1: a closed $\{r1:\text{REP}(b,A), r2:\text{REP}(b,B), (\exists x)E_{x:r1}A, (\exists x)E_{x:r2}B\}$ -tableau for $\neg E_{b:\text{itself}}(A \wedge B)$

(2), (3) (4) and (5) are introduced by $R\{r1:\text{REP}(b,A), r2:\text{REP}(b,B), (\exists x)E_{x:r1}A, (\exists x)E_{x:r2}B\}$ rule. (6) is from (2) and (7) is from (3) by RREP rule. (8) is from (4) by $R\delta$ rule; here f is a 0-place Skolem function symbol. (9) is from (6) by $R\gamma$ rule, introducing the new free variable v_1 . (10) is from (5) by $R\delta$ rule; here g is a 1-place Skolem function symbol. (11) is from (7) by $R\gamma$ rule, introducing the new free variable v_2 . (12) and (13) are from (9) by $R\beta$ rule. (14) and (15) are from (11) by $R\beta$ rule. Finally, (16) and (17) are from (1) by RC-E rule. Applying now the *tableau substitution rule*, using the substitution $\sigma = \{v_1/f, v_2/g(f)\}$, we produce a closed tableau due to the following conflicting formulas in each branch: (8) and (12); (11) and (14); (13) and (16); (15) and (17).

Let's now consider the following example of a tableau proof for $\{OE_{b:\text{itself}}B, OE_{b:\text{itself}}B \rightarrow OrB, \text{is}_r(a)\} \vdash OE_{a:r}B$ pictured in figure 2.

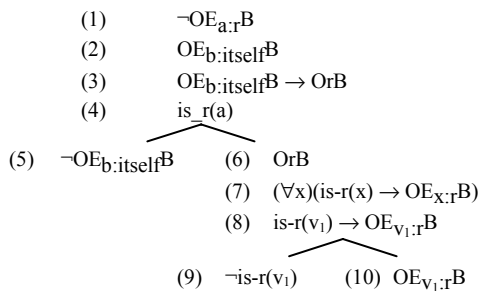


Figure 2: a closed $\{OE_{b:\text{itself}}B, OE_{b:\text{itself}}B \rightarrow OrB, \text{is}_r(a)\}$ -tableau for $\neg OE_{a:r}B$

(2), (3) and (4) are introduced by $R\{OE_{b:\text{itself}}B, OE_{b:\text{itself}}B \rightarrow OrB, \text{is}_r(a)\}$ rule. (5) and (6) are is from (3) by $R\beta$ rule. (7) is from (6) by ROr rule. (8) is from (7) by $R\gamma$ rule, introducing the new free variable v_1 . (9) and (10) are from (8) by $R\beta$ rule. Applying the *tableau substitution rule* with the substitution $\sigma = \{v_1/a\}$ we produce a closed tableau due to the following

conflicting formulas in each branch: (2) and (5); (4) and (9); (1) and (10).

It is easy to see that the previous rules yield a sound proof procedure, i.e., if B has a tableau proof then $\vdash B$. However, The completeness of the previous tableaux method is not yet established.

5. CONCLUDING REMARKS

We have proposed a logic-oriented framework for specification and analysis of organizations, based on the concepts of institutional agent, role and role-based agency. Institutional agents where described by a set of roles, a set of representative roles, their deontic characterization and by transmission of obligations principles (from institutional agents to roles). Although simple, we have shown that the proposed framework supports some interesting aspects of analysis of organizations.

We have also discussed the tableaux theorem proving method used in the automation of the first-order role based deontic/action modal logic herein proposed. This is our first step towards the automation of a systematic analysis of organisations.

We are actually working on a workbench for supporting the specification, analysis and design of institutional agents. One interesting aspect of analysis concerns the integrity evaluation of the deontic characterization of roles. Since an agent may play many roles, it is useful to analyse if there are conflicting roles in a given institutional agent, e.g. roles r_1 and r_2 s.t. $Or_1A \wedge Or_2B \wedge (Or_1A \rightarrow \neg Or_2B)$. Discovering such conflicting roles can be useful during organizational design. We intend to support this kind of analysis adapting Reiter's diagnosis theory (see [33] [34] [35]).

Our long-term research is to explore more elaborated organisational structures in order to extend the analysis to other interesting organisational subjects. We foresee two ways of attaining this objective: 1. to extend our organisational structure with other primitive concepts important to enrich the description of an organisation; and 2. to extend our organisational structure with other non-primitive relevant concepts for allowing a quick description of organisations. With respect to the first one, we intend to study and characterize other relevant normative concepts (e.g. responsibility, power, delegation, authority, authorisation). With respect to the second one, we intend to develop languages for organisation specification supporting notions usually used by real organisations.

6. ACKNOWLEDGMENTS

This work was partially supported by FCT under the Portuguese research project POSI/SRI/3951/2001 – MAGO2 (Modelling Agents for Organisations). The authors wish to thank the anonymous referees of ICAIL2003 for their suggestions and criticisms.

7. REFERENCES

- [1] Carley, K. and Les Gasser: 2000, "Computational Organization Theory". In Weiss, G., ed, Multiagent Systems: A modern approach to Distributed Artificial Intelligence, MIT Press, 299-330.

- [2] Carmo, J. and Jones, A.: 2001, "Deontic Logic and Contrary-to-Duties", in Gabbay, D. and Guenther, F., eds., *Handbook of Philosophical Logic – Second Edition*, 4, Kluwer, 287-363.
- [3] Carmo, J. and Pacheco, O.: 2000, "Deontic and Action Logics for Collective Agency and Roles". In Demolombe, R. and Hilpinen, R., eds., *Proceedings of the Fifth International Workshop on Deontic Logic in Computer Science (DEON'00)*, ONERA-DGA, Toulouse, 93-124.
- [4] Carmo, J. and Pacheco, O.: 2000, "Logics for Modeling Business and Agents Interaction". In *Proceedings of the International Conference on Advances in Infrastructure for Electronic Business, Science and Education on Internet (SSGRR 2000)*, l'Aquila, <http://www.ssgrr.it/ssgrr2000/proceedings.htm>.
- [5] Carmo, J. and Pacheco, O.: 2001, "Deontic and Action Logics for Organized Collective Agency, Modeled through Institutionalized Agents and Roles". *Fundamenta Informaticae* 34, IOS Press, 1-39
- [6] Chellas, B.J.: 1980, *Modal Logic - An Introduction*, Cambridge University Press.
- [7] Dobson, J. and Strens, R.: 1994, "Organizational Requirements Definition for Information Technology Systems". *Proceedings of IEEE International Conference on Requirements Engineering (ICRE94)*, IEEE Press, Colorado Springs.
- [8] Enderton, H.: 1972, *A Mathematical Introduction to Logic*, New York: Academic Press.
- [9] Enjalbert P. and Cerro, F.: 1989, "Modal Resolution in Clausal Form", *Theoretical Computer Science* 65, North-Holland, 1-33.
- [10] Fitting, M.: 1972, "Tableau Methods of Proof for Modal Logics", *Notre Dame Journal of Formal Logic* vol. XIII, no. 2.
- [11] Fitting, M.: 1983, *Proof Methods for Modal and Intuitionistic Logics*, D. Reidel, Dordrecht.
- [12] Fitting, M.: 1988, "First-Order Modal Tableaux", *Journal of Automated Reasoning*, no. 4, 191-213.
- [13] Fitting, M.: 1990, *First-Order Logic and Automated Theorem Proving*, Springer-Verlag.
- [14] Gallier, J.: 1987, *Logic for Computer Science: Foundations of Automatic Theorem Proving*, John Wiley & Sons.
- [15] Hilpinen, R., ed.: 1971, *Deontic Logic: Introductory and Systematic Readings*, D.Reidel.
- [16] Jones, A.J.I. and Sergot, M.: 1992, "Formal Specification of Security Requirements using the Theory of Normative Positions", *Proceedings of the European Symposium on Research in Computer Security (ESORICS'92)*, LNCS 648, Springer, 103-121.
- [17] Kanger, S. and Kanger, H.: 1966, "Rights and Parliamentarism", *Theoria* 6(2), 85-115.
- [18] Kanger, S.: 1957, *New Foundations for Ethical Theory*, Stockholm. (Reprinted in Hilpinen, R., ed.: 1971, *Deontic Logic: Introductory and Systematic Readings*, D.Reidel, 36-58).
- [19] Kanger, S.: 1972, "Law and Logic", *Theoria* 38.
- [20] Kirn, S.: 1996, "Organizational Intelligence and Distributed Artificial Intelligence". In O'Hare, G. and Jennings, N., eds., *Foundations of Distributed Artificial Intelligence*, John Wiley & Sons, 505-526.
- [21] Laudon, K. and Laudon, J.: 1996, *Management Information Systems: Organization and Technology*. Prentice Hall, Fourth Edition.
- [22] Lindahl, L.: 1977, *Position and Change - A Study in Law and Logic*, Synthese Library 112, D. Reidel, Dordrecht.
- [23] Meyer, J.-J. and Wieringa, R.: 1993, "Deontic Logic: A concise Overview". In Meyer, J.-J. and Wieringa, R., eds., *Deontic Logic in Computer Science: Normative System Specification*, John Wiley & Sons, 3-16.
- [24] Normatics: 1991, "Normatics: The Characterisation of Computer Systems and Complex Organisations as Normative Systems", *ESPRIT III BRA proposal*.
- [25] Pacheco, O.: 2002, *Especificação Normativa de Agentes Institucionais e da Interação entre Agentes*, PhD thesis, Engineering School of University of Minho.
- [26] Pacheco, O. and Carmo J.: 1998, "Collective Agents: From Law to AI". *Proceedings of Second French-American Conference on Law and Artificial Intelligence*, Nice.
- [27] Pacheco, O. and Carmo J. "Role Based Model for the Normative Specification of Organized Collective Agency and Agents Interaction". *Journal of Autonomous Agents and Multi-Agent Systems*, Vol. 6, Issue 2, Kluwer, March 2003, 145-184
- [28] Parunak, H. and Van, D.: 1996, "Applications of Distributed Artificial Intelligence in Industry". In O'Hare, G. and Jennings, N., eds., *Foundations of Distributed Artificial Intelligence*, John Wiley & Sons, 139-164.
- [29] Pörn, I.: 1970, *The Logic of Power*, Blackwell, Oxford.
- [30] Pörn, I.: 1971, *Elements of Social Analysis*, Philosophical Studies, 10, Philosophical Society and Department of Philosophy, University of Uppsala.
- [31] Pörn, I.: 1977, *Action Theory and Social Science: Some Formal Models*, Synthese Library 120, D. Reidel, Dordrecht.
- [32] Pörn, I.: 1989, "On the Nature of Social Order", in Fenstad, J.E. et al., eds., *Logic Methodology, and Philosophy of Science VIII*, Elsevier, 553-567.
- [33] Ramos, P. and Fiadeiro, J.: 1997, "Organisational Process Aided Design". In Callaos, N., Khoong, C. and Cohen, E., eds., *Proceedings of World Multiconference on Systemics, Cybernetics and Informatics (SCI'97)*, Caracas, 97-104.
- [34] Ramos, P. and Fiadeiro, J.: 1998, "a DEONTIC Logic for Diagnosis of Organisational Process Design". *Proceedings of the Fourth International Workshop on Deontic Logic in Computer Science (DEON'98)*, Università degli Studi di Bologna, 353-369.
- [35] Reiter, R.: 1987, "A Theory of Diagnosis from First Principles". *Artificial Intelligence* 32, 57-95.

- [36] Santos, F.: 1998: *Lógicas Modais de Acção para a Modelação da Interação entre Agentes e de Organizações*, PhD thesis, Science Faculty of University of Lisbon.
- [37] Santos, F.: 2002: “A Modal Logic Framework for Organization Analysis and Design”. Submitted to *Sixth International Workshop on Deontic Logic in Computer Science (DEON'02)*.
- [38] Santos, F. and Carmo, J.: 1996, “Indirect Action, Influence and Responsibility”, in Brown, M. and Carmo, J., eds., *Deontic Logic, Agency and Normative Systems*, Springer, Workshops in Computing Series, 194-215.
- [39] Santos, F. and Carmo, J.: 1996, “A Modal Action Logic Based Framework for Organization Specification and Analysis”, in Alferes, J.J., Pereira, L.M. and Orłowska, E., eds., *Proceedings of the Fifth European Workshop on Logics in AI (JELIA'96)*, Springer LNAI 1126, 119-133.
- [40] Santos, F., Jones, A. and Carmo, J.: 1997, “Responsibility for Action in Organisations: a Formal Model”. In Holmstrom-Hintikka, G. and Tuomela, R., eds., *Contemporary Action Theory, vol. II (Social Action)*, Synthese Library, 267, Kluwer, 333-350.
- [41] Skarneas, N.: 1995, Modeling Organizations using Roles and Agents. *Proceedings of the Fifth Hellenic Conference on Informatics*, Athens.