A logical account of institutions: from acceptances to norms via legislators

Emiliano Lorini and **Dominique Longin**
Institut de Recherche en Informatique de Toulouse (IRIT)
118 route de Narbonne
F-31062 Toulouse Cedex 04 (France)

Abstract

The aim of this paper is to provide a logical framework which enables reasoning about institutions and their dynamics. In our approach an institution is grounded on the *acceptances* of its members. We devote special emphasis to the role of *legislator*. We characterize the legislator as the role whose function is the creation and the modification of legal facts (e.g. permissions, obligations, *etc.*): the acceptance of the legislators that a certain norm is valid ensures that the norm is valid. The second part of the paper is devoted to the logical characterization of two important notions in the domain of legal and social theory: the notion of *constitutive rule* and the notion of *norm of competence*. A constitutive rule is a rule which is responsible for the creation of new kinds of (institutional) facts. A norm of competence is a rule which assigns powers to the agents playing certain roles within the institution. We show that norms of competence provide the criteria for institutional change.

Introduction

The problem of devising artificial institutions and modeling their dynamics is a fundamental problem in the multi-agent system domain (Dignum and Dignum 2001). Following (North 1990, p. 3), artificial institutions can be conceived as human-like: "the rules of the game in a society or the humanly devised constraints that structure agents' interaction". Starting from this concept of institution, many researchers working in the field of normative multi-agent systems have been interested in developing models which describe the different kinds of rules and norms that agents have to deal with. In some models of artificial institutions norms are conceived as means to achieve coordination among agents and agents are supposed to comply with them and to obey the authorities of the system as an end (Esteva, Padget, and Sierra 2001). More sophisticated models of institutions leave to the agents' autonomy the decision whether to comply or not with the specified rules and norms of the institution (Ågotnes et al. 2007; Lopez y Lopez, Luck, and d'Inverno 2004). However, all previous models abstract way from the legislative source of the norms of an institution, and from how institutions are created, maintained and changed by their members and not imposed from the outside by an external designer.

The aim of this work is to advance the state of the art on artificial institutions by proposing a logical model in which the existence and the dynamics of an institution (norms, rules, institutional facts, *etc.*) depend on the individual and collective attitudes of the agents which identify themselves as members of the institution. In particular, we propose a model in which an institution is grounded on the (individual and collective) *acceptances* of its followers and members, and its dynamics depend on the dynamics of these acceptances. On this aspect we agree with (Mantzavinos, North, and Shariq 2004), when the authors say that (p. 77):

> "only because institutions are anchored in peoples minds do they ever become behaviorally relevant. The *elucidation of the internal aspect is the crucial step* in adequately explaining the emergence, evolution, and effects of institutions." [Emphasis added].

In our model the agents are supposed to play certain social roles in one or more institutions and to accept things while playing these roles in the institutions. We devote special emphasis to the social role of *legislators* and show that the acceptances of the legislators directly affect the dynamics of the rules and the norms of the institution: the acceptances of the legislators are responsible for creating and modifying the obligations and the permissions of the institution.

It is worth noting that other authors in the MAS field have emphasized the need for a model which explains the origin and the evolution of institutions in terms of the agents' attitudes (Conte, Castelfranchi, and Dignum 1998; Conte and Dignum 2001; Boella and Van der Torre 2007). For instance, in agreement with (Hart 1992), Conte et al. (Conte, Castelfranchi, and Dignum 1998; Conte and Dignum 2001) have stressed that the existence of a norm in an institution (but also in a group, organization, *etc.*) depends on the recognition and acceptance of the norm by the members of the institution. In Conte et al.'s perspective, the agents contribute to the enforcement and the propagation of the norm. Furthermore, it has to be noted that, although in our approach institutions are anchored in agents' attitudes, we do not claim that institutions can be conceived as agents. Thus, our approach is different from (Boella and van der Torre 2004), in which the metaphor of normative systems as agents is used and institutions are described in terms of mental attitudes such as beliefs and goals.

The paper is organized as follows. The second section of the article will be devoted to discussing the notion of acceptance and to distinguishing it from the classical notion of belief. In the third section we will introduce a modal logic which enables reasoning about obligations and individual and collective acceptances of agents while playing a certain *social role* within the institution (*i.e.* acceptance *qua* players of a certain role within the institution). This logic is an extension of the logic of acceptance we have presented in (Gaudou et al. 2008), in which social roles and obligations were not considered. We will devote special emphasis to the logical characterization of the acceptances of the agents playing the social role of *legislators* within the institution. We will formally characterize the legislators' power to create and modify the legal level of an institution. In the last sections of the article we will extend our analysis to the distinction between regulative and non-regulative components of an institution (Searle 1995). First, we will formalize the concept of *constitutive rule*, that is, the kind of rules accepted by the legislators which are responsible for the creation of new kinds of (institutional) facts. Since (Searle 1995; 1969) and (Jones and Sergot 1996), these rules have been expressed in terms of assertions of the form "X counts as Y in the context of institution x" (*e.g.* in the institutional context of US, a piece of paper with a certain shape, color, *etc.* counts as a five-dollar bill). We will conclude with a logical analysis of a particular form of constitutive rule, the so-called *norm of competence*. Norms of competence are rules which assign powers to the agents playing certain roles within the institution. We will show that norms of competence provide the criteria for institutional change.

An overview of the notion of *acceptance*

Before presenting our logical framework, we provide a brief overview of the concept of acceptance.

Whereas beliefs have been studied for decades, acceptances have only been examined since (Stalnaker 1984) and (Cohen 1992) while studying the nature of argument premises or reformulating Moore's paradox (Cohen 1992). If a belief that p is an attitude constitutively aimed at the truth of p, an acceptance is the output of "a decision to treat p as true in one's utterances and actions" (Hakli 2006; Bratman 1992) without being necessarily connected to the actual truth of the proposition. In order to better distinguish these two notions, it has been suggested (Hakli 2006) that while beliefs are not subject to the agent's will, acceptances are voluntary; while beliefs aim at truth, acceptances are sensitive to pragmatic considerations; while beliefs are shaped by evidence, acceptances need not be; while beliefs come in degrees, acceptances are qualitative; finally, while beliefs are context-independent, acceptance depends on context.

For the aims of this article we are particularly interested in the last feature, namely the fact that acceptances are context-dependent. In fact, one can decide (say for prudential reasons) to reason and act by "accepting" the truth of a proposition in a specific context, and possibly rejecting the very same proposition in another context. This aspect of acceptance has been studied both with respect to cooperative contexts (Gilbert 1989) (*e.g.* the context of a team) and with respect to institutional contexts (Tuomela 2007). We here continue the work initiated in (Gaudou et al. 2008) by exploring the role of acceptance with respect to institutional contexts. Institutional contexts are either rule-governed social practices (informal institutions) (*e.g.* language, games) or legal institutions, in which agents play certain social roles and on the background of which they reason. Consider a legal institution such as a trading company. The institutional context is the set of rules and norms which the agents conform to when they play the role of employees in the company. On the background of such contexts, we are interested in the individual and collective acceptances that can be formally captured. In the context of the trading company, for instance, the agents accept that something is true *qua* employees of the company. The state of acceptance *qua* player of a certain role in a certain institution is the kind of acceptance one is committed to when one is functioning as a player of a certain role in the institution (Tuomela 2007).

A logic of acceptance and obligation

Syntax

The syntactic primitives of our logic $\mathcal{L}$ of acceptance and obligation are the following: a finite set of $n > 0$ agents $AGT = \{i, j, \ldots\}$; a nonempty finite set of *atomic actions* $ACT = \{\alpha, \beta, \ldots\}$; a finite set of atomic formulas $ATM = \{p, q, \ldots\}$; a finite set of labels denoting institutional contexts $INST = \{x, y, \ldots\}$; a finite set of labels denoting social roles $ROLE = \{a, b, \ldots\}$. We suppose that $ROLE$ contains a (single) special role leg corresponding to the role of legislator of a certain institution. Moreover, we note $2^{AGT^\star} = 2^{AGT} \setminus \{\emptyset\}$ the set of all nonempty subsets of agents, $\Delta = 2^{AGT^\star} \times ROLE \times INST$ the set of all triples of non empty subsets of agents, social roles, and institutional contexts. We note $C{:}a{:}x$ the elements of Δ.

The language $\mathcal{LANG}$ of the logic $\mathcal{L}$ is defined by the following BNF:

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \vee \varphi \mid After_{i:\alpha}\varphi \mid [C{:}a{:}x]\,\varphi \mid O_x\varphi$$

where i ranges over AGT, α ranges over ACT, C ranges over $2^{AGT^\star}$, a ranges over $ROLE$ and x ranges over $INST$. The classical boolean connectives $\wedge$, $\rightarrow$, $\leftrightarrow$, $\top$ and $\bot$ are defined from $\vee$ and $\neg$ in the usual manner. For notational convenience we write $[i{:}a{:}x]$ instead of $[\{i\}{:}a{:}x]$, for any $i \in AGT$.

Formula $[C{:}a{:}x]\,\varphi$ reads "the agents in group C accept that φ while playing role a together in the institution x". Operators of the form $[C{:}a{:}x]$ are extensions of the operators of acceptance $[C{:}x]$ we have introduced in (Gaudou et al. 2008) where we completely ignored social roles.

EXAMPLE 1. $[C{:}activist{:}Greenpeace]\,protectEarth$ *is read "the agents in C accept that their mission is to protect the Earth while playing together the role of activists in Greenpeace".*

The formula $[C{:}a{:}x]\,\bot$ has to be read "agents in C do not play role a together in the institution x" because we assume that playing a role together in a certain institution is, at least in this minimal sense, a rational activity; conversely, $\neg [C{:}a{:}x] \bot$ has to be read "agents in C play role a together in the institution x"; $\neg [C{:}a{:}x] \bot \wedge [C{:}a{:}x] \varphi$ stands for "agents in C play role a together in institution x and they accept that φ while playing role a together in institution x" or simply "agents in C accept that φ *qua* players of role a in the institution x" (*i.e.* **collective acceptance**).

For the individual case, formula $\neg [i{:}a{:}x] \bot \wedge [i{:}a{:}x] \varphi$ has to be read "agent i accepts that φ *qua* player of role a in the institution x" (*i.e.* **individual acceptance**).

O_x are operators of obligation of standard deontic logic (SDL) indexed by institutional contexts and are used to express those facts which are legal with respect to a certain institution. Formula $O_x\varphi$ has to be read "φ is obligatory in the institution x". The dynamic operators of the form $After_{i:\alpha}$ are similar to the standard operators of dynamic logic (Harel, Kozen, and Tiuryn 2000) where both the action and its author are specified. Formula $After_{i:\alpha}\varphi$ has to be read "after agent i does action α, it is the case that φ".

We introduce four concepts by means of abbreviations. Their meanings will become clearer later in the analysis where the axioms and some theorems of the logic $\mathcal{L}$ will be discussed. For any $i \in AGT$, $\alpha \in ACT$, $C \in 2^{AGT\star}$ and $x \in INST$:

$$Happens_{i:\alpha}\varphi \stackrel{def}{=} \neg After_{i:\alpha}\neg\varphi$$

$$Leg(C,x) \stackrel{def}{=} \neg [C{:}leg{:}x] \bot \wedge \bigwedge_{C \subset B} [B{:}leg{:}x] \bot$$

$$Leg_x\varphi \stackrel{def}{=} \bigvee_{C \in 2^{AGT\star}} (Leg(C,x) \wedge [C{:}leg{:}x] \varphi)$$

$$Leg_{Univ}\varphi \stackrel{def}{=} \bigwedge_{x \in INST} Leg_x\varphi$$

Formula $Happens_{i:\alpha}\varphi$ has to be read "agent i performs action α and φ is true afterward". Formula $Leg(C,x)$ stands for "C is *the* group of legislators of institution x". Indeed, we suppose that the group of legislators of a certain institution x is the group C whose agents play together the role of legislators in x and there is no super-group B of C whose agents play together the role of legislators in x. We will show below that, for every institution x, there is only one group of legislators of x (this is the reason why we do not read $Leg(C,x)$ as "C is *a* group of legislators of institution x".) Formula $Leg_x\varphi$ stands for "there exists a group of legislators of x which accept φ". This can be shortened to "the legislators of x accept that φ" (due to the fact that in our logic every institution has only one group of legislators) or more simply "within the institutional context x, it is the case that φ". Finally, formula $Leg_{Univ}\varphi$ has to be read "the legislators of all institutions accept that φ" or simply "φ is universally accepted as true".

Semantics

We use a possible worlds semantics. A model of the logic $\mathcal{L}$ is a tuple $\mathcal{M} = \langle W, \mathscr{A}, \mathscr{R}, \mathscr{O}, \mathscr{V} \rangle$ where:

- W is a set of possible worlds;
- $\mathscr{A} : \Delta \longrightarrow (W \longrightarrow 2^W)$ associates each $C{:}a{:}x \in \Delta$ and world w with the set $\mathscr{A}_{C:a:x}(w)$ of worlds accepted by the group C at w, where the agents in C are playing role a in the institution x;
- $\mathscr{O} : INST \longrightarrow (W \longrightarrow 2^W)$ associates each $x \in INST$ and possible world w with the set $\mathscr{O}_x(w)$ of worlds which are ideal with regard to the institution x;
- $\mathscr{R} : AGT \times ACT \longrightarrow (W \longrightarrow 2^W)$ associates each agent $i \in AGT$, action $\alpha \in ACT$ and world w with the set $\mathscr{R}_{i:\alpha}(w)$ of worlds that are reachable from w through the occurrence of action α performed by i;
- $\mathscr{V} : W \longrightarrow 2^{ATM}$ is a truth assignment which associates each world w with the set $\mathscr{V}(w)$ of atomic propositions true in w.

The truth conditions of formulas are recursively defined as follows:

- $\mathcal{M}, w \models p$ iff $p \in \mathscr{V}(w)$;
- $\mathcal{M}, w \models \neg\varphi$ iff not $\mathcal{M}, w \models \varphi$;
- $\mathcal{M}, w \models \varphi \vee \psi$ iff $\mathcal{M}, w \models \varphi$ or $\mathcal{M}, w \models \psi$;
- $\mathcal{M}, w \models [C{:}a{:}x]\varphi$ iff for all $w' \in W$, if $w' \in \mathscr{A}_{C:a:x}(w)$ then $\mathcal{M}, w' \models \varphi$;
- $\mathcal{M}, w \models O_x\varphi$ iff for all $w' \in W$, if $w' \in \mathscr{O}_x(w)$ then $\mathcal{M}, w' \models \varphi$;
- $\mathcal{M}, w \models After_{i:\alpha}\varphi$ iff for all $w' \in W$, if $w' \in \mathscr{R}_{i:\alpha}(w)$ then $\mathcal{M}, w' \models \varphi$.

Axiomatization

Every operator of type $[C{:}a{:}x]$, O_x and $After_{i:\alpha}$ is supposed to be a normal modal operator satisfying standard axioms and rules of inference of the basic modal logic K. The rest of the section contains other axioms of the logic $\mathcal{L}$ and corresponding semantic constraints over $\mathcal{L}$ models.

Action. We suppose the following constraint over $\mathcal{L}$ models. For every $w \in W$, $i, j \in AGT$ and $\alpha, \beta \in ACT$:

$$\text{if } w' \in \mathscr{R}_{i:\alpha}(w) \text{ and } w'' \in \mathscr{R}_{j:\beta}(w) \text{ then } w' = w'' \qquad \textbf{S1}$$

The property **S1** says that all actions occurring in a world w lead to the same world. Thus, all actions occur in parallel and they do not have non-deterministic effects. This explains why we have phrased $Happens_{i:\alpha}\varphi$ "i does α and φ holds afterward" rather than "*it is possible that* i does α and φ holds afterward". Constraint **S1** corresponds to the following axiom of our logic. For every $i, j \in AGT$ and $\alpha, \beta \in ACT$:

$$Happens_{i:\alpha}\varphi \rightarrow After_{j:\beta}\varphi \qquad \textbf{Determ}$$

Acceptance and role playing. We suppose that: if agents in C accept that φ while playing role a together in the institution x then, for every subset B of C, while playing role a together in the institution x, the agents in B accept that the agents in C accept that φ, while playing role a together in the institution x. This means that given a group of agents C, all subgroups of C have access to all the facts that are accepted by the agents in C while playing together a certain role in an institution. Such property is expressed by the following axiom. For every $B, C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$, if $B \subseteq C$ then:

$$[C{:}a{:}x]\,\varphi \rightarrow [B{:}a{:}x]\,[C{:}a{:}x]\,\varphi \qquad \mathbf{4}_{\boldsymbol{Accept}}$$

Axiom $\mathbf{4}_{\boldsymbol{Accept}}$ corresponds to the following semantic constraint over $\mathcal{L}$ models. For every $w \in W$, $B, C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$, if $B \subseteq C$ then:

$$\text{if } w' \in \mathscr{A}_{B:a:x}(w) \text{ then } \mathscr{A}_{C:a:x}(w') \subseteq \mathscr{A}_{C:a:x}(w) \qquad \mathbf{S2}$$

Moreover, we assume that if the agents in C accept that φ *qua* players of role a in the institution x then, for every subset B of C, it holds that the agents in B accept φ *qua* players of role a in the institution x. Thus, for every $B, C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$, if $B \subseteq C$ then:

$$(\neg\,[C{:}a{:}x]\bot \wedge [C{:}a{:}x]\,\varphi) \rightarrow (\neg\,[B{:}a{:}x]\,\bot \wedge [B{:}a{:}x]\,\varphi) \qquad \mathbf{Inc}_{\boldsymbol{Accept}}$$

EXAMPLE 2. *Imagine three agents* i, j, k *that,* qua *Clue players, accept that someone called Mrs. Red, has been killed:* $\neg\,[\{i,j,k\}\,{:}player{:}Clue]\,\bot \wedge [\{i,j,k\}\,{:}player{:}Clue]\, killedMrsRed$. *This implies that also the two agents* i, j qua *Clue players accept that someone called Mrs. Red has been killed:* $\neg\,[\{i,j\}\,{:}player{:}Clue]\,\bot \wedge [\{i,j\}\,{:}player{:}Clue]\, killedMrsRed$.

Axiom $\mathbf{Inc}_{\boldsymbol{Accept}}$ corresponds to the following semantic constraint over $\mathcal{L}$ models. For every $w \in W$, $B, C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$, if $B \subseteq C$ then:

$$\text{if } \mathscr{A}_{C:a:x}(w) \neq \emptyset \text{ then } \mathscr{A}_{B:a:x}(w) \neq \emptyset \text{ and } \mathscr{A}_{B:a:x}(w) \subseteq \mathscr{A}_{C:a:x}(w) \qquad \mathbf{S3}$$

The last axiom concerning acceptance and role playing says that: the agents in a group C, while playing together role a in the institution x, accept that they play together role a in the institution x. Formally for every $C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$:

$$[C{:}a{:}x]\,\neg\,[C{:}a{:}x]\,\bot \qquad \mathbf{RolePlay}$$

Intuitively, Axiom $\mathbf{RolePlay}$ means that if the agents in a group C play together a certain role within a certain institution then, this fact is public for the group C.

EXAMPLE 3. *Suppose that, during a concert, the agents in* C *play together the role of musicians in the context of the Philharmonic Orchestra. Then, this is public for the agents in* C. *That is, while playing together the role of musicians in the Philharmonic Orchestra, the agents in* C *accept that they are playing together the role of musicians in the Philharmonic Orchestra:* $[C{:}musician{:}Orchestra]\,\neg\,[C{:}musician{:}Orchestra]\,\bot$.

Axiom $\mathbf{RolePlay}$ corresponds to the following semantic constraint over $\mathcal{L}$ models. For every $w \in W$, $C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$:

$$\forall w' \in \mathscr{A}_{C:a:x}(w),\, \mathscr{A}_{C:a:x}(w') \neq \emptyset \qquad \mathbf{S4}$$

Acceptance and action. We also suppose the axiom of no forgetting for acceptance. This axiom describes how operators of acceptance interact with dynamic operators of the form $After_{i:\alpha}$. For every $i \in AGT$, $\alpha \in ACT$, $C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$:

$$([C{:}a{:}x]\, After_{i:\alpha}\varphi \wedge \neg\,[C{:}a{:}x]\, After_{i:\alpha}\bot) \rightarrow After_{i:\alpha}\,[C{:}a{:}x]\,\varphi \qquad \mathbf{NF}$$

A lot of researchers have studied similar principles for the interaction between belief and action or between knowledge and action. Among them we should mention (Fagin et al. 1995; Gerbrandy 1999; Scherl and Levesque 2003; Herzig, Lang, and Polacsek 2000). It has to be noted that axiom $\mathbf{NF}$ relies on an assumption of complete and correct perception information. It is supposed that an agent i's action α occurs if and only if every group of agents (viz. single agent) is informed of this fact and updates its acceptances accordingly. Hence all action occurrences are supposed to be public. The axiom corresponds to the following semantic constraint over $\mathcal{L}$ models. For every $w \in W$, $i \in AGT$, $\alpha \in ACT$, $C \in 2^{AGT^\star}$, $x \in INST$ and $a \in ROLE$:[1]

$$\text{if } \mathscr{A}_{C:a:x} \circ \mathscr{R}_{i:\alpha}(w) \neq \emptyset \text{ then } \mathscr{R}_{i:\alpha} \circ \mathscr{A}_{C:a:x}(w) \subseteq \mathscr{A}_{C:a:x} \circ \mathscr{R}_{i:\alpha}(w) \qquad \mathbf{S5}$$

Legal level of institutions. As far as the legal level of institutions is concerned, we suppose the standard deontic logic principle: if φ is obligatory in the context of institution x then, $\neg\varphi$ is not obligatory in the context of the same institution. Formally, for every $x \in INST$:

$$\neg(O_x\varphi \wedge O_x\neg\varphi) \qquad \boldsymbol{D_{Obl}}$$

Axiom $\boldsymbol{D_{Obl}}$ corresponds to the following semantic constraints over $\mathcal{L}$ models. For every $w \in W$ and $x \in INST$:

$$\mathscr{O}_x(w) \neq \emptyset \qquad \mathbf{S6}$$

Legislator. We add two specific axioms for the social role *legislator*. According to the first axiom, we cannot have two groups of agents which play *separately* the role of legislators in the same institution.

More precisely, given two groups B and C, if the agents in B play together the role of legislators in the institution x and the agents in C play together the role of legislators in the institution x, then the agents in $B \cup C$ play together the role of legislators in the institution x.

Formally, for every $B, C \in 2^{AGT^\star}$ and $x \in INST$:

$$(\neg\,[B{:}leg{:}x]\,\bot \wedge \neg\,[C{:}leg{:}x]\,\bot) \rightarrow \neg\,[B \cup C{:}leg{:}x]\,\bot \qquad \mathbf{LegSum}$$

Axiom $\mathbf{LegSum}$ corresponds to the following semantic constraint over $\mathcal{L}$ models. For every $w \in W$, $B, C \in 2^{AGT^\star}$ and $x \in INST$:

$$\text{if } \mathscr{A}_{B:leg:x}(w) \neq \emptyset \text{ and } \mathscr{A}_{C:leg:x}(w) \neq \emptyset \text{ then } \mathscr{A}_{B\cup C:leg:x}(w) \neq \emptyset \qquad \mathbf{S7}$$

[1] We note $\circ$ the standard composition operator such that, given two arbitrary functions $\mathscr{T}^1$ and $\mathscr{T}^2$ over worlds in W, $\mathscr{T}^1 \circ \mathscr{T}^2(w) = \bigcup \{\mathscr{T}^2(v) \mid v \in \mathscr{T}^1(w)\}$.

The following axiom **LegPower** is intended to capture the peculiar power of the group of legislators to create legal facts. We assume that the legislators of x accept that φ is obligatory in the institution x if and only if φ is obligatory in this institution. Formally, for every $x \in INST$:

$$Leg_x O_x \varphi \leftrightarrow O_x \varphi \qquad \textbf{LegPower}$$

It is worth noting that axiom **LegPower** do not express the intermediate step between the acceptance of the legislators that a norm is valid and the instantiation of the norm in the institution. This step, which is left implicit in the present analysis, is based on the legislators' act of *proclaiming* that the norm is valid.[2] Moreover, in axiom **LegPower** the legislators' power to create obligations is just expressed by means of a material implication (the left to right direction of the axiom). A more adequate characterization of this concept would require to substitute the material implication with a conditional which better expresses the fact that the legislators are *responsible* for the creation of the obligation.

Axiom **LegPower** corresponds to the following two semantic constraints over $\mathcal{L}$ models. For every $w \in W$, $C \in 2^{AGT^\star}$ and $x \in INST$:

$$\text{if } \mathscr{A}_{C:leg:x}(w) \neq \emptyset \text{ and } \forall B \text{ such that } C \subset B, \mathscr{A}_{B:leg:x}(w) = \emptyset, \text{ then } \mathscr{O}_x(w) \subseteq \mathscr{A}_{C:leg:x} \circ \mathscr{O}_x(w) \qquad \textbf{S8}$$

For every $w \in W$ and $x \in INST$:

$$\exists C \in 2^{AGT^\star} \text{ such that } \mathscr{A}_{C:leg:x}(w) \neq \emptyset \text{ and } \forall B \text{ such that } C \subset B, \mathscr{A}_{B:leg:x}(w) = \emptyset \text{ and } \mathscr{A}_{C:leg:x} \circ \mathscr{O}_x(w) \subseteq \mathscr{O}_x(w) \qquad \textbf{S9}$$

We call $\mathcal{L}$ the logic axiomatized by the principles presented above and we write $\vdash_{\mathcal{L}} \varphi$ iff formula φ is a theorem of $\mathcal{L}$ provable from our axioms by the inference rules of modus ponens and necessitation for every modal operator. Moreover, we write $\models_{\mathcal{L}} \varphi$ iff formula φ is *valid* in all $\mathcal{L}$ models, *i.e.* $\mathcal{M}, w \models \varphi$ for every $\mathcal{L}$ model $\mathcal{M}$ and world w in $\mathcal{M}$. Finally, we say that a formula φ is *satisfiable* if $\not\models_{\mathcal{L}} \neg\varphi$ We can prove that the logic $\mathcal{L}$ is *sound* and *complete* with respect to the class of $\mathcal{L}$ models. Namely:

Theorem 1 *$\mathcal{L}$ is determined by the class of $\mathcal{L}$ models.*

Properties of acceptances

This section provides further clarifications of the concept of acceptance. In particular, we focus on the distinction between acceptance and belief.

As said, there is a large literature about the distinction between belief and acceptance. For us, belief and acceptance are clearly different concepts in several senses. (For convenience, we adopt $Bel_i\varphi$ as a notation for "the agent i believes that φ is true", and we suppose that these operators are defined as usual in a KD45 modal logic).

Individual belief and individual acceptance are both private mental attitudes but: a belief does not depend on contexts, whilst an acceptance is a context-dependent attitude which is entertained by an agent *qua* player of a certain role within a given institution. Therefore, an agent can privately disbelieve something he accepts while playing a certain role within a given institution. Formally: $Bel_i\varphi \wedge [i{:}a{:}x]\neg\varphi$ should be satisfiable. In a similar way, as emphasized in (Tuomela 1992), a collective acceptance that φ of a group of agents C (*qua* players of a certain role within a given institution) might be compatible with the fact that every agent in C does not believe that φ (or that every agent in C believes that $\neg\varphi$). The following example, inspired from (Tuomela 1992, p. 285), illustrates this.

EXAMPLE 4.*At the end of the 80s, the Communist Party of Ruritania accepted that capitalist countries will soon perish (but none of its members really believed so).*

We can formalize the example as follows: $\neg [C{:}member{:}CPR] \bot \wedge [C{:}member{:}CPR]\, ccwp \wedge \bigwedge_{i \in C} \neg Bel_i\, ccwp$. This means that the agents in C accept that capitalist countries will perish ($ccwp$) *qua* members of the Communist Party of Ruritania (CPR) but nobody in C believes this.

Properties of legislators

This section is devoted to studying the notion of legislator. The following theorem highlights some of its properties.

Theorem 2 *For every $x \in INST$ and $B, C \in 2^{AGT^\star}$ such that $B \neq C$:*

$$\vdash_{\mathcal{L}} Leg(C,x) \rightarrow \neg Leg(B,x) \qquad (2a)$$

$$\vdash_{\mathcal{L}} \bigvee_{C \in 2^{AGT^\star}} Leg(C,x) \qquad (2b)$$

$$\textit{If } \vdash_{\mathcal{L}} \varphi \textit{ then } \vdash_{\mathcal{L}} Leg_x\varphi \qquad (2c)$$

$$\vdash_{\mathcal{L}} (Leg_x(\varphi \rightarrow \psi) \wedge Leg_x\varphi) \rightarrow Leg_x\psi \qquad (2d)$$

$$\vdash_{\mathcal{L}} \neg(Leg_x\varphi \wedge Leg_x\neg\varphi) \qquad (2e)$$

$$\vdash_{\mathcal{L}} Leg_x\varphi \rightarrow Leg_x Leg_x\varphi \qquad (2f)$$

Theorem 2a ensures that, for every institution, there is only one group of legislators. Theorem 2b says that for every institution x, there is always a group of legislators of x. Theorems 2c–2f highlight the fact that operators of type Leg_x are normal modal operators satisfying the axioms and rules of inference of the system KD4 (Chellas 1980). In particular, Theorem 2f says that, if the legislators of x accept that φ, then the legislators of x accept that the legislators of x accept that φ. This latter property captures a sort of 'introspective' capacity of legislators: legislators have access to those facts that they accept *qua* legislators.

Weak permissions *vs.* strong permissions. As discussed above, the legislator of a certain institution is the social role which has the function of creating and modifying legal facts. In particular, we have assigned to the legislators the power to create obligations (Axiom **LegPower**). Now, let us consider permissions in order to establish a new formal relationship between the legal level of institutions and the legislators. We define permission in the usual way by taking the dual of the operator of obligation. We say that "φ is something permitted in the institutional context x" (noted $P_x\varphi$) if

[2] For a logical characterization of the act of *proclaiming*, see (Gelati et al. 2004).

and only if $\neg\varphi$ is not obligatory in the institutional context x. Formally:

$$P_x\varphi \stackrel{def}{=} \neg O_x\neg\varphi.$$

The following theorem can be proved.

Theorem 3 *For every $x \in INST$:*

$$\vdash_{\mathcal{L}} Leg_x P_x\varphi \rightarrow P_x\varphi$$

Thus, in our logic the legislators are also endowed with the power of creating permissions. It has to be noted that the converse of Theorem 3 is not a theorem of our logic: $P_x\varphi \wedge \neg Leg_x P_x\varphi$ is satisfiable in the logic $\mathcal{L}$.

Thus, φ might be permitted within institution x while the legislators of x do not accept φ to be permitted within the context x. (In this sense, permissions behave differently from obligations, cf. Axiom $\mathbf{LegPower}$.)

This property is justified by the distinction between weak permission and strong permission, which was emphasized by several authors in analytical philosophy (Alchourrón and Bulygin 1971; Raz 1975; Von Wright 1963) and in the domain of normative MAS (Boella and van der Torre 2003). According to Von Wright for instance "[...] An act will be said to be permitted in the weak sense if it is not forbidden; and it will be said to be permitted in the strong sense if it is not forbidden but subject to norm." (Von Wright 1963, p. 86). A weak permission corresponds to the absence in the institution of a norm prohibiting φ. This concept is captured by the formula $P_x\varphi$ of our logic. A strong permission corresponds to the existence in the institution of an explicit norm, accepted by the legislators, according to which φ is permitted, which is captured by the formula $Leg_x P_x\varphi$ of our logic. In this perspective, Theorem 3 states that a strong permission implies a weak permission. In contrast, the converse is not valid.

In the rest of the article we will investigate the fundamental concepts of *constitutive rule*, *norm of competence* and *institutionalized power* within the formal framework of $\mathcal{L}$.

From constitutive rules to norms of competence

According to many philosophers working on social theory (Rawls 1955; Alchourrón and Bulygin 1971) and researchers in the field of normative multi-agent systems (Boella and van der Torre 2004), institutions are based both on regulative as well as constitutive (*i.e.* non-regulative) components. That is, institutions are not only defined in terms of sets of permissions, obligations, and prohibitions (*i.e. norms of conduct*) but also in terms of rules which specify and create new forms of behavior and concepts. According to Searle for instance "[...] regulative rules regulate antecedently or independently existing forms of behavior [...]. But constitutive rules do not merely regulate, they create or define new forms of behavior" (Searle 1969, p. 33). In Searle's theory of institutions (Searle 1969; 1995), constitutive rules are expressed by means of "counts-as" assertions of the form "X counts as Y in context x" where the context x refers to the institution/normative system in which the rule is specified. For example, in the institutional context of US, a piece of paper with a certain shape, color, *etc.* counts as a five-dollar bill.

The distinction between regulative rules and constitutive rules can be expressed in our formal language $\mathcal{L}$. Regulative rules are characterized in $\mathcal{L}$ by the constructions $O_x\varphi$ (obligation), $P_x\varphi$ (weak permission) and $Leg_x P_x\varphi$ (strong permission) introduced above.

The following two subsections are devoted to presenting a formal characterization of the concept of constitutive rule. We will first provide a formal analysis of the general notion of constitutive rule. Then, we will investigate a particular form of constitutive rule which is commonly referred to as *norm of competence* (Bulygin 1992). A norm of competence of a certain institution x is a norm on the basis of which special (institutionalized) powers are assigned to the agents playing a certain role in the institution.

Constitutive rules

A notion of *constitutive rule* of the form "φ counts as ψ in the institutional context x" can be defined in our logic $\mathcal{L}$ by means of the operator Leg_x. We conceive a constitutive rule as a material implication of the form $\varphi \rightarrow \psi$ in the scope of an operator Leg_x. Thus, "φ counts as ψ in the institutional context x" only if the legislators of institution x accept that φ entails ψ. Furthermore, we suppose that a constitutive rule is intrinsically contextual, that is, a rule that is not universally valid while it is accepted by the legislators of a certain institution. More precisely, we exclude the situation in which $Leg_{Univ}(\varphi \rightarrow \psi)$ is true (the legislators of every institution accept that φ entails ψ). More generally, for every $x \in INST$ the following abbreviation $\varphi \stackrel{x}{\rhd} \psi$ (that stands for "φ counts as ψ in the institutional context x") is given:

$$\varphi \stackrel{x}{\rhd} \psi \stackrel{def}{=} Leg_x(\varphi \rightarrow \psi) \wedge \neg Leg_{Univ}(\varphi \rightarrow \psi)$$

EXAMPLE 5. *The formula $sixteen \stackrel{Brazil}{\rhd} votingAge$ stands for "in Brazil, the fact that a person is sixteen year old counts as the fact that he has the voting age". This means that the legislators of Brazil accept that being sixteen year old entails having the voting age,* i.e. $Leg_{Brazil}(sixteen \rightarrow votingAge)$, *and there are legislators of other countries who do not accept this,* i.e. $\neg Leg_{Univ}(sixteen \rightarrow votingAge)$. *Indeed, there are other countries such as Italy and France in which the voting age is set at eighteen years and not at sixteen.*[3]

It is interesting to note that $\varphi \stackrel{x}{\rhd} \psi$ satisfies some intuitive properties of counts-as conditionals as isolated in (Jones and Sergot 1996).

[3]Note that a more precise characterization of this example requires a quantification over the set of agents AGT. That is, the constitutive rule should be specified by the formula $\bigwedge_{i \in AGT}(sixteen(i) \stackrel{Brazil}{\rhd} votingAge(i))$ which is meant to stand for " in Brazil, for every agent i, i is sixteen year old counts as i has the voting age".

Theorem 4 *For every $x \in INST$:*

$$\textit{If } \vdash_{\mathcal{L}} (\varphi_2 \leftrightarrow \varphi_3) \textit{ then } \vdash_{\mathcal{L}} (\varphi_1 \overset{x}{\rhd} \varphi_2 \leftrightarrow \varphi_1 \overset{x}{\rhd} \varphi_3) \quad (4a)$$

$$\textit{If } \vdash_{\mathcal{L}} (\varphi_1 \leftrightarrow \varphi_3) \textit{ then } \vdash_{\mathcal{L}} (\varphi_1 \overset{x}{\rhd} \varphi_2 \leftrightarrow \varphi_3 \overset{x}{\rhd} \varphi_2) \quad (4b)$$

$$\vdash_{\mathcal{L}} (\varphi_1 \overset{x}{\rhd} \varphi_2 \wedge \varphi_1 \overset{x}{\rhd} \varphi_3) \rightarrow (\varphi_1 \overset{x}{\rhd} (\varphi_2 \wedge \varphi_3)) \quad (4c)$$

$$\vdash_{\mathcal{L}} (\varphi_1 \overset{x}{\rhd} \varphi_2 \wedge \varphi_3 \overset{x}{\rhd} \varphi_2) \rightarrow ((\varphi_1 \vee \varphi_3) \overset{x}{\rhd} \varphi_2) \quad (4d)$$

$$\vdash_{\mathcal{L}} (\varphi_1 \overset{x}{\rhd} \varphi_2 \wedge (\varphi_1 \wedge \varphi_2) \overset{x}{\rhd} \varphi_3) \rightarrow (\varphi_1 \overset{x}{\rhd} \varphi_3) \quad (4e)$$

For instance, Theorem 4e corresponds to a property of cumulative transitivity (cut). We can easily show that the operator $\overset{x}{\rhd}$ does not satisfy reflexivity, transitivity and weakening of the antecedent, that is: $\varphi \overset{x}{\rhd} \varphi$, $(\varphi_1 \overset{x}{\rhd} \varphi_2 \wedge \varphi_2 \overset{x}{\rhd} \varphi_3) \rightarrow \varphi_1 \overset{x}{\rhd} \varphi_3$, and $\varphi_1 \overset{x}{\rhd} \varphi_2 \rightarrow (\varphi_1 \wedge \varphi_3) \overset{x}{\rhd} \varphi_2$ are not valid in $\mathcal{L}$. This is due to the "local" nature of the operator $\overset{x}{\rhd}$. For instance, $\varphi_1 \overset{x}{\rhd} \varphi_2$ and $\varphi_2 \overset{x}{\rhd} \varphi_3$ might be constitutive rules of the institution x, while $\varphi_1 \overset{x}{\rhd} \varphi_3$ fails to be a constitutive rule of x since it is not intrinsically contextual (*i.e.* $Leg_{Univ}(\varphi_1 \overset{x}{\rhd} \varphi_3)$ holds).

It has to be noted that our notion of "counts as" is similar to the notion of *proper classificatory rule* defined in (Grossi, Meyer, and Dignum 2006). [4]

Norms of competence and institutionalized power

According to some legal theorists (Bulygin 1992; Searle 1969; Hart 1992), norms of competence are power-conferring rules which should not be reduced to norms of conduct such as obligations, prohibitions, commands and permissions. These kinds of rules assign special powers to the agents playing certain roles within the institution. They have a fundamental function in normative and legal systems since they provide the criteria for institutional change, that is, they provide the criteria for the creation and modification of institutional facts (*e.g.* agent i and agent j are married, this house is i's property, *etc.*) and normative facts (*e.g.* obligations and permissions).

Norms of competence can be specified in the logic $\mathcal{L}$ in the following way. For any $x \in INST$, $a \in ROLE$ and $\alpha \in ACT$:

$$Power(a,\alpha,\varphi,x) \stackrel{def}{=} \bigwedge_{i \in AGT} (\neg\, [i{:}a{:}x] \perp \overset{x}{\rhd} After_{i:\alpha}\varphi)$$

$Power(a,\alpha,\varphi,x)$ reads "in institution x there is a norm of competence which assigns to the agents playing role a in x the power to ensure φ by performing action α" or simply "in the institutional context x, the agents playing role a (in x) have the power to ensure φ by performing action α".

EXAMPLE 6. *The formula $Power(priest, gesture, married, church)$ is meant to stand for "in the institutional context of Catholic Church, the agents playing the role of priest (in the Church) have the power of marrying a couple by performing certain gestures".*

From the previous concept of *institutionalized power*, we can define a corresponding notion of *exercise of institutionalized power*. We say that in the institutional context x, an agent i playing role a (in x) *exercises* its power of ensuring φ by doing action α if and only if:

i. in context x, the agents playing role a (in x) have the power to ensure φ by performing action α;

ii. the legislators of x accept that i is playing role a in x and that agent i performs action α.

Formally, for any $x \in INST$, $a \in ROLE$, $i \in AGT$ and $\alpha \in ACT$:

$$ExPower(i,a,\alpha,\varphi,x) \stackrel{def}{=} Power(a,\alpha,\varphi,x) \wedge Leg_x(\neg\, [i{:}a{:}x] \perp \wedge Happens_{i:\alpha}\top)$$

Our aim is to show how the exercise of an institutionalized power by an agent modifies the current structure of the institution through the creation of new institutional facts. To this end, we have to introduce the following definition. For any $x \in INST$, $i \in AGT$ and $\alpha \in ACT$:

$$NoChange(x, i{:}\alpha) \stackrel{def}{=} \bigwedge_{C \in 2^{AGT\star}} (Leg(C,x) \rightarrow After_{i:\alpha} Leg(C,x))$$

$NoChange(x, i{:}\alpha)$ is meant to stand for "the group of legislators of x do not change after agent i performs action α" (*i.e.* for any $C \in 2^{AGT\star}$, if C is the group of legislators of x then, after agent i performs action α, C is still the group of legislators of x).

We are now in the position to prove two theorems which highlight the dynamic aspect of institutions based on the exercise of institutionalized power.

Theorem 5 *For every $x \in INST$, $a \in ROLE$, $i \in AGT$ and $\alpha \in ACT$:*

$$\vdash_{\mathcal{L}} (ExPower(i,a,\alpha,\varphi,x) \wedge NoChange(x, i{:}\alpha)) \rightarrow After_{i:\alpha} Leg_x \varphi$$

According to Theorem 5, if in the institution x agent i playing role a exercises its power of ensuring φ by doing action α, under the condition that the legislators of x do not change after agent i performs action α then, after i performs α, it is the case that φ is true within the institution x.

EXAMPLE 7. *Suppose that in the Church, agent i playing the role of priest exercises its power of marrying a couple by performing certain gestures, noted $ExPower(i, priest, gesture, married, church)$. Then, under the condition $NoChange(church, i{:}gesture)$ (the group of legislators of the Church do not change after i's action), after i performs the gestures, the couple will be married in the context of the Church, noted $After_{i:gesture} Leg_{church} married$.*

Theorem 6 *For every $x \in INST$, $a \in ROLE$, $i \in AGT$ and $\alpha \in ACT$:*

$$\vdash_{\mathcal{L}} (NoChange(x, i{:}\alpha) \wedge ExPower(i,a,\alpha,O_x\varphi,x)) \rightarrow After_{i:\alpha} O_x \varphi$$

[4] We refer to (Grossi, Meyer, and Dignum 2006) for interesting arguments why proper classificatory rules should not necessarily satisfy reflexivity, transitivity and weakening of the antecedent.

Theorem 6 highlights the dynamics of obligations in an institution due to the exercise of institutionalized powers by agents. It says that: if in the institution x agent i playing role a exercises its power of creating the obligation that φ by doing action α, then, after i performs α, it is the case that φ is obligatory in x. (Under the condition that the legislators of x do not change.)

Conclusion

We have presented in this article a logic of acceptance and obligation and applied it to the analysis of institutions and their dynamics. Our logic of acceptance and obligation allows to express that certain agents accept something to be true *qua* players of a role within an institution. We have devoted special emphasis to the social role *legislator* by discussing its influence on the creation and the modification of the norms of an institution. In the second part of the paper we have formalized the concept of constitutive rule, that is, a rule of the form "X counts as Y in the context of institution x" which is responsible for the creation of institutional facts. While constitutive rules are usually defined from the external perspective of a normative system or institution, we have, once again, anchored these rules in the acceptances of the legislators. We have concluded with an analysis of a particular form of constitutive rule, the so-called norm of competence. A norm of competence is a norm on the basis of which special institutionalized powers are assigned to the agents playing certain roles in the institution.

Directions for future research are manifold. Our future works will be devoted to better clarify the relationships between the legislators of an institution and the acceptances of the members of the institution. In particular, we will integrate the following two general principles into our logical framework. According to first principle, all members of an institution have to accept, *qua* members of the institution, all facts which are accepted by the legislators of the institution. This principle expresses that the members of an institution are necessarily subject to what the legislators of the institution accept. According to the second principle, the agents in a set C are the legislators of a certain institution only if the members of the institution accept that the agents in C are the legislators of the institution and recognize them as the legislators. This second principle expresses that the legitimacy of the legislator's authority is necessarily based on the consent of the members of the institution. Our logic is sufficiently expressive to capture the two principles:

- $[C{:}a{:}x]\, Leg_x\varphi \rightarrow [C{:}a{:}x]\,\varphi$
- $Leg(B{,}x) \rightarrow [C{:}a{:}x]\, Leg(B{,}x)$

The first formula expresses that, if the agents in C, while playing role a in institution x, accept that the legislators of x accept φ then, the agents in C have to accept φ while playing role a in x. The second formula expresses that, if B is the group of legislators of institution x then, for every set of agents C and role a, the agents in C, while playing role a in x, have to accept that B is the group of legislators of x.

Furthermore, in future extensions of this work, we will investigate the decision to join or not to join (and the decision to leave or to remain member of) a given institution. This decision is influenced by the inconsistency between the agent's preferences and goals and the current norms and rules of the institution. For instance, if the agent's goals conflict with the norms proclaimed by the legislators then, the agent will probably decide not to join the institution. In order to model this form of reasoning, we will extend our logical framework to modalities expressing agents' goals and preferences, such as the ones provided in (Cohen and Levesque 1990).

Acknowledgments

Emiliano Lorini is financed by the project ForTrust (Social Trust Analysis and Formalization) funded by the french Agence Nationale de la Recherche (ANR). Thanks are due to our colleague Andreas Herzig and to the anonymous reviewers of this paper for their helpful comments and suggestions.

Appendix: proofs of theorems

Proof of Theorem 1. It is a routine to prove soundness, whereas completeness is obtained by Sahlqvist completeness theorem (Blackburn, de Rijke, and Venema 2001). Indeed, all axioms of the logic $\mathcal{L}$ are in the Sahlqvist class.

Proof of Theorem 2a.

$$\vdash_{\mathcal{L}} Leg(C{,}x) \rightarrow \neg Leg(B{,}x) \text{ if } B \neq C$$

It is enough to prove the Theorem for the following two cases.

- CASE 1. $B \subset C$ or $C \subset B$
- CASE 2. $B \not\subset C$ and $C \not\subset B$

The proof for the first case is straightforward. Let us give a proof for the second case.

1. $(Leg(C{,}x) \wedge Leg(B{,}x)) \rightarrow$
   $(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C \subset D} [D{:}leg{:}x]\,\bot \wedge \neg\,[B{:}leg{:}x]\,\bot)$
   from def. $Leg(C{,}x)$ and def. $Leg(B{,}x)$
2. $(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C \subset D} [D{:}leg{:}x]\,\bot \wedge \neg\,[B{:}leg{:}x]\,\bot) \rightarrow$
   $(\neg\,[B \cup C{:}leg{:}x]\,\bot \wedge \bigwedge_{C \subset D} [D{:}leg{:}x]\,\bot)$
   From Axiom **LegSum**
3. $(\neg\,[B \cup C{:}leg{:}x]\,\bot \wedge \bigwedge_{C \subset D} [D{:}leg{:}x]\,\bot) \rightarrow \bot$
   From the facts $B \neq C$, $B \neq \emptyset$, $C \neq \emptyset$[5], $B \not\subset C$ and $C \not\subset B$
4. $(Leg(C{,}x) \wedge Leg(B{,}x)) \rightarrow \bot$
   From 1,2,3
5. $Leg(C{,}x) \rightarrow \neg Leg(B{,}x)$ From 4

Proof of Theorem 2b.

$$\vdash_{\mathcal{L}} \bigvee_{C \in 2^{AGT^\star}} Leg(C{,}x)$$

1. $O_x\top$

[5] Remember that B and C are member of the set $2^{AGT^\star}$ of non empty subsets of agents.

2. $O_x\top \rightarrow Leg_x\top$
   From Axiom $\mathbf{LegPower}$
3. $Leg_x\top$
   From 1,2
4. $Leg_x\top \rightarrow (\bigvee_{C\in 2^{AGT^\star}} Leg(C,x))$
   From def. $Leg_x\top$ and def. $Leg(C,x)$
5. $\bigvee_{C\in 2^{AGT^\star}} Leg(C,x)$
   From 3,4

Proof of Theorem 2c.

$$\text{From } \vdash_{\mathcal{L}} \varphi \text{ infer } \vdash_{\mathcal{L}} Leg_x\varphi$$

Let us suppose that φ is a theorem of $\mathcal{L}$. We prove that $Leg_x\varphi$ is a theorem of $\mathcal{L}$ as well.

1. φ
   From hypothesis
2. $\bigwedge_{C\in 2^{AGT^\star}} [C{:}leg{:}x]\,\varphi$
   From 1 and necessitation rule for $[C{:}leg{:}x]$
3. $\bigvee_{C\in 2^{AGT^\star}} Leg(C,x)$
   From Theorem 2b
4. $\bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi)$
   From 2,3
5. $Leg_x\varphi$
   From 4 and def. $Leg_x\varphi$

Proof of Theorem 2d.

$$\vdash_{\mathcal{L}} (Leg_x(\varphi \rightarrow \psi) \wedge Leg_x\varphi) \rightarrow Leg_x\psi$$

1. $Leg_x(\varphi \rightarrow \psi) \leftrightarrow \bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,(\varphi \rightarrow \psi))$
   From def. $Leg_x(\varphi \rightarrow \psi)$
2. $Leg_x\varphi \leftrightarrow \bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi)$
   From def. $Leg_x\varphi$
3. $\bigwedge_{B,C\in 2^{AGT^\star},B\neq C}(Leg(C,x) \rightarrow \neg Leg(B,x))$
   From Theorem 2a
4. $(Leg_x(\varphi \rightarrow \psi) \wedge Leg_x\varphi) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,(\varphi \rightarrow \psi) \wedge [C{:}leg{:}x]\,\varphi)$
   From 1,2,3
5. $\bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,(\varphi \rightarrow \psi) \wedge [C{:}leg{:}x]\,\varphi) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\psi)$
   From Axiom **K** for $[C{:}leg{:}x]$
6. $\bigvee_{C\in 2^{AGT^\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\psi) \rightarrow Leg_x\psi$
   From def. $Leg_x\psi$
7. $(Leg_x(\varphi \rightarrow \psi) \wedge Leg_x\varphi) \rightarrow Leg_x\psi$
   From 4,5,6

Proof of Theorem 2e.

$$\vdash_{\mathcal{L}} \neg(Leg_x\varphi \wedge Leg_x\neg\varphi)$$

1. $\bot \rightarrow O_x\bot$
   From standard principles of propositional calculus
2. $Leg_x\bot \rightarrow Leg_xO_x\bot$
   From 1, Theorems 2c and 2d
3. $Leg_xO_x\bot \rightarrow O_x\bot$
   From Axiom $\mathbf{LegPower}$
4. $O_x\bot \rightarrow (O_x\varphi \wedge O_x\neg\varphi)$
5. $(O_x\varphi \wedge O_x\neg\varphi) \rightarrow \bot$
   From Axiom $\boldsymbol{D_{Obl}}$
6. $Leg_x\bot \rightarrow \bot$
   From 2,3,4,5
7. $\neg Leg_x(\varphi \wedge \neg\varphi)$
   From 6
8. $\neg(Leg_x\varphi \wedge Leg_x\neg\varphi)$
   From 7 and standard principles of the normal modal operator Leg_x (*i.e.* $(Leg_x\varphi \wedge Leg_x\psi) \leftrightarrow Leg_x(\varphi \wedge \psi)$ is a theorem of $\mathcal{L}$)

Proof of Theorem 2f.

$$\vdash_{\mathcal{L}} Leg_x\varphi \rightarrow Leg_xLeg_x\varphi$$

1. $Leg_x\varphi \rightarrow \bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\varphi)$
   From def. $Leg_x\varphi$
2. $\bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\varphi) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[C{:}leg{:}x]\,[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,[C{:}leg{:}x]\,\varphi)$
   From Axiom $\mathbf{4}_{\boldsymbol{Accept}}$ and Axiom $\mathbf{RolePlay}$
3. $(\bigwedge_{1\leq i\leq n}[C{:}leg{:}x]\,\varphi_i) \leftrightarrow ([C{:}leg{:}x]\bigwedge_{1\leq i\leq n}\varphi_i)$
   Standard principle of the normal modal operator $[C{:}leg{:}x]$
4. $\bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[C{:}leg{:}x]\,[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,[C{:}leg{:}x]\,\varphi) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\neg\,[C{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,[C{:}leg{:}x]\,\varphi)$
   From 3
5. $\bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\neg\,[C{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,[C{:}leg{:}x]\,\varphi) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\varphi))$
   From 3
6. $\bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,\varphi)) \rightarrow \bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,Leg_x\varphi)$
   From def. $Leg_x\varphi$
7. $\bigvee_{C\in 2^{AGT^\star}}(\neg\,[C{:}leg{:}x]\,\bot \wedge \bigwedge_{C\subset B}[B{:}leg{:}x]\,\bot \wedge [C{:}leg{:}x]\,Leg_x\varphi) \rightarrow Leg_xLeg_x\varphi$
   From def. $Leg_xLeg_x\varphi$
8. $Leg_x\varphi \rightarrow Leg_xLeg_x\varphi$
   From 1-2, 4-7

Proof of Theorem 3.

$$\vdash_{\mathcal{L}} Leg_x P_x\varphi \rightarrow P_x\varphi$$

1. $Leg_x P_x\varphi \rightarrow Leg_x \neg O_x \neg\varphi$
   From def. $P_x\varphi$
2. $\neg(Leg_x \neg O_x \neg\varphi \wedge Leg_x O_x \neg\varphi)$
   From Theorem 2e
3. $Leg_x \neg O_x \neg\varphi \rightarrow \neg Leg_x O_x \neg\varphi$
   From 2
4. $Leg_x P_x\varphi \rightarrow \neg Leg_x O_x \neg\varphi$
   From 1,3
5. $\neg Leg_x O_x \neg\varphi \leftrightarrow \neg O_x \neg\varphi$
   From Axiom **LegPower**
6. $Leg_x P_x\varphi \rightarrow P_x\varphi$
   From 4,5 and def. $P_x\varphi$

Proof of Theorem 5.

$$\vdash_{\mathcal{L}} (ExPower(i,a,\alpha,\varphi,x) \wedge NoChange(x, i{:}\alpha)) \rightarrow$$

$$After_{i:\alpha} Leg_x\varphi$$

1. $(ExPower(i,a,\alpha,\varphi,x) \wedge NoChange(x, i{:}\alpha)) \rightarrow$
   $(Power(a,\alpha,\varphi,x) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge Happens_{i:\alpha}\top) \wedge$
   $NoChange(x, i{:}\alpha))$
   From def. $ExPower(i,a,\alpha,\varphi,x)$
2. $(Power(a,\alpha,\varphi,x) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge Happens_{i:\alpha}\top) \wedge$
   $NoChange(x, i{:}\alpha)) \rightarrow$
   $((\neg [i{:}a{:}x] \bot \overset{x}{\rhd} After_{i:\alpha}\varphi) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge$
   $Happens_{i:\alpha}\top) \wedge NoChange(x, i{:}\alpha))$
   From def. $Power(a,\alpha,\varphi,x)$
3. $((\neg [i{:}a{:}x] \bot \overset{x}{\rhd} After_{i:\alpha}\varphi) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge$
   $Happens_{i:\alpha}\top) \wedge NoChange(x, i{:}\alpha)) \rightarrow$
   $(Leg_x(\neg [i{:}a{:}x] \bot \rightarrow After_{i:\alpha}\varphi) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge$
   $Happens_{i:\alpha}\top) \wedge NoChange(x, i{:}\alpha))$
   From def. $\neg [i{:}a{:}x] \bot \overset{x}{\rhd} After_{i:\alpha}\varphi$
4. $(Leg_x(\neg [i{:}a{:}x] \bot \rightarrow After_{i:\alpha}\varphi) \wedge Leg_x(\neg [i{:}a{:}x] \bot \wedge$
   $Happens_{i:\alpha}\top) \wedge NoChange(x, i{:}\alpha)) \rightarrow$
   $(Leg_x(\neg [i{:}a{:}x] \bot \rightarrow After_{i:\alpha}\varphi) \wedge Leg_x \neg [i{:}a{:}x] \bot \wedge$
   $Leg_x Happens_{i:\alpha}\top \wedge NoChange(x, i{:}\alpha))$
   From standard principles of the normal modal operator Leg_x (*i.e.* $Leg_x(\varphi \wedge \psi) \leftrightarrow (Leg_x\varphi \wedge Leg_x\psi)$ is a theorem of $\mathcal{L}$)
5. $(Leg_x(\neg [i{:}a{:}x] \bot \rightarrow After_{i:\alpha}\varphi) \wedge Leg_x \neg [i{:}a{:}x] \bot \wedge$
   $Leg_x Happens_{i:\alpha}\top \wedge NoChange(x, i{:}\alpha)) \rightarrow$
   $(Leg_x After_{i:\alpha}\varphi \wedge Leg_x Happens_{i:\alpha}\top \wedge$
   $NoChange(x, i{:}\alpha))$
   From Theorem 2d
6. $(Leg_x After_{i:\alpha}\varphi \wedge Leg_x Happens_{i:\alpha}\top \wedge$
   $NoChange(x, i{:}\alpha)) \rightarrow$
   $(Leg_x(After_{i:\alpha}\varphi \wedge Happens_{i:\alpha}\top) \wedge NoChange(x, i{:}\alpha))$
   From standard principles of the normal modal operator Leg_x
7. $(Leg_x(After_{i:\alpha}\varphi \wedge Happens_{i:\alpha}\top) \wedge$
   $NoChange(x, i{:}\alpha)) \rightarrow$
   $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
   $[C{:}leg{:}x]\,(After_{i:\alpha}\varphi \wedge Happens_{i:\alpha}\top)))$
   From def. $Leg_x(After_{i:\alpha}\varphi \wedge Happens_{i:\alpha}\top)$
8. $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
   $[C{:}leg{:}x]\,(After_{i:\alpha}\varphi \wedge Happens_{i:\alpha}\top))) \rightarrow$
   $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
   $[C{:}leg{:}x]\, After_{i:\alpha}\varphi \wedge [C{:}leg{:}x]\, Happens_{i:\alpha}\top))$
   From standard principles of the normal modal operator $[C{:}leg{:}x]$
9. $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
   $[C{:}leg{:}x]\, After_{i:\alpha}\varphi \wedge [C{:}leg{:}x]\, Happens_{i:\alpha}\top)) \rightarrow$
   $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
   $\neg [C{:}leg{:}x] \bot \wedge [C{:}leg{:}x]\, After_{i:\alpha}\varphi \wedge$
   $[C{:}leg{:}x]\, Happens_{i:\alpha}\top))$
   From def. $Leg(C,x)$
10. $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
    $\neg [C{:}leg{:}x] \bot \wedge [C{:}leg{:}x]\, After_{i:\alpha}\varphi \wedge$
    $[C{:}leg{:}x]\, Happens_{i:\alpha}\top)) \rightarrow$
    $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
    $\neg [C{:}leg{:}x]\, After_{i:\alpha}\bot \wedge [C{:}leg{:}x]\, After_{i:\alpha}\varphi \wedge))$
    From def. $Happens_{i:\alpha}\top$ and standard principles of the operator $[C{:}leg{:}x]$ (*i.e.* $\neg [C{:}leg{:}x] \bot \wedge [C{:}leg{:}x]\, Happens_{i:\alpha}\top$ implies $\neg [C{:}leg{:}x]\, After_{i:\alpha}\bot$)
11. $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
    $\neg [C{:}leg{:}x]\, After_{i:\alpha}\bot \wedge [C{:}leg{:}x]\, After_{i:\alpha}\varphi)) \rightarrow$
    $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
    $After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi))$
    From Axiom **NF**
12. $(NoChange(x, i{:}\alpha) \wedge \bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge$
    $After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi)) \rightarrow$
    $(\bigwedge_{C \in 2^{AGT\star}}(Leg(C,x) \rightarrow After_{i:\alpha} Leg(C,x)) \wedge$
    $\bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi))$
    From def. $NoChange(x, i{:}\alpha)$
13. $(\bigwedge_{C \in 2^{AGT\star}}(Leg(C,x) \rightarrow After_{i:\alpha} Leg(C,x)) \wedge$
    $\bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi)) \rightarrow$
    $\bigvee_{C \in 2^{AGT\star}}(After_{i:\alpha} Leg(C,x) \wedge After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi)$
14. $\bigvee_{C \in 2^{AGT\star}}(After_{i:\alpha} Leg(C,x) \wedge$
    $After_{i:\alpha}\,[C{:}leg{:}x]\,\varphi) \rightarrow$
    $\bigvee_{C \in 2^{AGT\star}}(After_{i:\alpha}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi))$
    From standard principles of the normal modal operator $After_{i:\alpha}$
15. $\bigvee_{C \in 2^{AGT\star}}(After_{i:\alpha}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi)) \rightarrow$
    $After_{i:\alpha}(\bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi))$
    From standard principles of the normal modal operator $After_{i:\alpha}$
16. $After_{i:\alpha}(\bigvee_{C \in 2^{AGT\star}}(Leg(C,x) \wedge [C{:}leg{:}x]\,\varphi)) \rightarrow$
    $After_{i:\alpha} Leg_x\varphi$
    From def. $Leg_x\varphi$
17. $(ExPower(i,a,\alpha,\varphi,x) \wedge NoChange(x, i{:}\alpha)) \rightarrow$
    $After_{i:\alpha} Leg_x\varphi$
    From 1-16

References

Ågotnes, T.; van der Hoek, W.; Rodriguez-Aguilar, J.; Sierra, C.; and Wooldridge, M. 2007. On the logic of normative systems. In *Proceedings of the Twentieth International Joint Conference on Artificial Intelligence (IJCAI'07)*, 1181–1186. AAAI Press.

Alchourrón, C., and Bulygin, E. 1971. *Normative systems*. New York: Springer.

Blackburn, P.; de Rijke, M.; and Venema, Y. 2001. *Modal Logic*. Cambridge: Cambridge University Press.

Boella, G., and van der Torre, L. 2003. Permissions and obligations in hierarchical normative systems. In *Proceedings of the Eighth International Conference on Artificial Intelligence and Law (ICAIL'03)*, 109–118. ACM Press.

Boella, G., and van der Torre, L. 2004. Regulative and constitutive norms in normative multiagent systems. In *Proceedings of the Ninth International Conference on Principles on Principles of Knowledge Representation and Reasoning(KR 2004)*, 255–266. AAAI Press.

Boella, G., and Van der Torre, L. 2007. Norm negotiation in multiagent systems. *International Journal of Cooperative Information Systems* 16(1):97–122.

Bratman, M. E. 1992. Practical reasoning and acceptance in context. *Mind* 101(401):1–15.

Bulygin, E. 1992. On norms of competence. *Law and Philosophy* 11(3):201–216.

Chellas, B. F. 1980. *Modal logic: an introduction*. Cambridge: Cambridge University Press.

Cohen, P. R., and Levesque, H. J. 1990. Intention is choice with commitment. *Artificial Intelligence* 42:213–261.

Cohen, L. J. 1992. *An essay on belief and acceptance*. New York, USA: Oxford University Press.

Conte, R., and Dignum, F. 2001. From social monitoring to normative influence. *The Journal of Artificial Societies and Social Simulation* 4(2).

Conte, R.; Castelfranchi, C.; and Dignum, F. 1998. Autonomous norm acceptance. In *Intelligent Agents V (ATAL'98)*, volume 1555 of *LNCS*, 99 – 112. Berlin: Springer Verlag.

Dignum, V., and Dignum, F. 2001. Modelling agent societies: Coordination frameworks and institutions. In Brazdil, P., and Jorge, A., eds., *Proceedings of the Tenth Portuguese Conference in Artificial Intelligence (EPIA'01)*, volume 2258 of *LNAI*. Berlin: Springer-Verlag.

Esteva, M.; Padget, J.; and Sierra, C. 2001. Formalizing a language for institutions and norms. In *Intelligent Agents VIII (ATAL'01)*, volume 2333 of *LNAI*, 348 – 366. Berlin: Springer Verlag.

Fagin, R.; Halpern, J.; Moses, Y.; and Vardi, M. 1995. *Reasoning about Knowledge*. Cambridge: MIT Press.

Gaudou, B.; Longin, D.; Lorini, E.; and Tummolini, L. 2008. Anchoring Institutions in Agents' Attitudes: Towards a Logical Framework for Autonomous MAS. In *Proceedings of the Seventh International Joint Conference on Autonomous Agents & Multi-Agent Systems (AAMAS'08)*, 728–735. ACM Press.

Gelati, J.; Rotolo, A.; Sartor, G.; and Governatori, G. 2004. Normative autonomy and normative co-ordination: Declarative power, representation, and mandate. *Artificial Intelligence and Law* 12(1-2):53–81.

Gerbrandy, J. 1999. *Bisimulations on Planet Kripke*. The Netherlands: PhD thesis, University of Amsterdam.

Gilbert, M. 1989. *On Social Facts*. London and New York: Routledge.

Grossi, D.; Meyer, J.-J. C.; and Dignum, F. 2006. Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation* 16(5):613–643.

Hakli, P. 2006. Group beliefs and the distinction between belief and acceptance. *Cognitive Systems Research* 7:286–297.

Harel, D.; Kozen, D.; and Tiuryn, J. 2000. *Dynamic Logic*. Cambridge: MIT Press.

Hart, H. L. A. 1992. *The concept of law*. Oxford: Clarendon Press. new edition.

Herzig, A.; Lang, J.; and Polacsek, T. 2000. A modal logic for epistemic tests. In *Proceedings of the Fourteenth European Conference on Artificial Intelligence (ECAI 2000)*, 553–557. Berlin: IOS Press.

Jones, A., and Sergot, M. J. 1996. A formal characterization of institutionalised power. *Journal of the IGPL* 4:429–445.

Lopez y Lopez, F.; Luck, M.; and d'Inverno, M. 2004. Normative agent reasoning in dynamic societies. In *Proceedings of the Third International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'04)*, 732 – 739. ACM Press.

Mantzavinos, C.; North, D.; and Shariq, S. 2004. Learning, institutions, and economic performance. *Perspectives on Politics* 2:75–84.

North, D. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.

Rawls, J. 1955. Two concepts of rules. *The Philosophical Review* 64:3–32.

Raz, J. 1975. *Practical reason and norms*. London: Hutchinson.

Scherl, R. B., and Levesque, H. 2003. Knowledge, action, and the frame problem. *Artificial Intelligence* 144:1–39.

Searle, J. R. 1969. *Speech acts: An essay in the philosophy of language*. New York: Cambridge University Press.

Searle, J. R. 1995. *The Construction of Social Reality*. New York: The Free Press.

Stalnaker, R. 1984. *Inquiry*. Cambridge: MIT Press.

Tuomela, R. 1992. Group beliefs. *Synthese* 91:285–318.

Tuomela, R. 2007. *The Philosophy of Sociality*. Oxford: Oxford University Press.

Von Wright, G. H. 1963. *Norm and Action*. London: Routledge and Kegan.