Grounding power on actions and mental attitudes

Emiliano Lorini, Nicolas Troquard, Andreas Herzig, Jan Broersen

▶ **To cite this version:**

Emiliano Lorini, Nicolas Troquard, Andreas Herzig, Jan Broersen. Grounding power on actions and mental attitudes. Logic Journal of the IGPL, 2013, Special issue of best papers of FAMAS 2007, 21 (3), pp.311-331. 10.1093/jigpal/jzr039 . hal-01153715

HAL Id: hal-01153715
https://hal.science/hal-01153715

Submitted on 20 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 12311

To link to this article : DOI:10.1093/jigpal/jzr039
URL : http://dx.doi.org/10.1093/jigpal/jzr039

To cite this version : Lorini, Emiliano and Troquard, Nicolas and Herzig, Andreas and Broersen, Jan *Grounding power on actions and mental attitudes.* (2013) Logic Journal of the IGPL, Vol. 21 (n° 3). pp. 311-331. ISSN 1367-0751

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Grounding power on actions and mental attitudes

EMILIANO LORINI[1], NICOLAS TROQUARD[2], ANDREAS HERZIG[1], AND JAN BROERSEN[3],

[1] *Université de Toulouse, IRIT-CNRS, France,* [2] *Department of Computer Science, University of Liverpool, UK and* [3] *Department of Information and Computing Sciences, Universiteit Utrecht, The Netherlands*

Abstract

The main objective of this work is to develop a logic called $\mathcal{IAL}$ (*Intentional Agency Logic*) in which we can reason about mental states of agents, action occurrences, and agentive and group powers. $\mathcal{IAL}$ will be exploited for a formal analysis of different forms of power such as an agent i's *power of* achieving a certain result and an agent i's *power over* another agent j (alias *social power*).

Keywords: Modal logic, BDI, agency, social powers.

1 Introduction

Power is one fundamental concept in a situation of agent interaction as studied in social theory and multi-agent systems. Our aim is to design a general logical framework in which various and important forms of power can be specified and their intrinsic and relational properties can be investigated.

A comprehensive formal model of power should clarify many subtle aspects of this individual and social phenomenon. First of all it should characterize the most basic form of agentive power called *power of* achieving something. When looking at an agent's *power of* achieving a certain result, we discover that this is based on the interrelation between objective level and subjective level. In fact, i's *power of* achieving a certain result φ seems to involve not only i's objective opportunity of achieving φ, but also i's awareness over such an opportunity.[1] For example, for a thief to have the power of opening a safe, he must know its combination.

Moreover, a comprehensive formal model of power should be able to characterize social forms of power which are commonly called *social powers* (or *powers over*). There are several types of social powers which need to be distinguished and which have interesting relationships among them: influencing power, persuasive power, dependence-based social power.

Influencing power is an agent's power to influence other agents to do or to refrain from doing certain actions.[2] In other words, an agent i's *power of influencing* another agent j consists in i's capacity to shape j's preferences in such a way that j will intend or will not intend to do a certain

[1] In this article, the terms belief and awareness are taken to be synonymous, even though in the logic literature they often have different meanings (see e.g. [22]).

[2] In this article, actions should be understood as action *types*, for that they are actions that can occur more than once.

action. For example, for a politician to have the power of influencing the electorate with regard to the action of voting for him, he must have the power of inducing the electorate to vote for him.

Persuasive power is an agent's power to induce other agents to believe or to abstain from believing certain things. The relationship between influencing power and persuasive power is very tight. Indeed, certain beliefs provide sufficient *reasons* for intending to do a certain action *a*: if agent *i* has the power to induce these beliefs in agent *j* then, indirectly, *i* has the power to influence *j* to do the action *a*.[3] For example, again suppose *i* is a politician before an election and *j* is a potential voter. Agent *j* wants to pay less taxes in the next year. If *i* has the power to persuade *j* that voting for him and ensuring that he will win the election is the only way to pay less taxes in the next year then, indirectly, *i* has the power to influence *j*'s decision in such a way that *j* will intend to vote for *i*.

Dependence-based social power of an agent *i* over another agent *j* is *i*'s power over *j* which is based on *j*'s dependence on *i* for the achievement of his goals. In particular, *i* has a dependence-based power over *j* if and only if, *j* has a certain goal and *j* will not achieve his goal without *i*'s intervention. Dependence-based social power is tightly related with influencing power and persuasive power. In fact, under certain conditions, dependence-based social power and persuasive power enable influencing power. For instance, if *i* has a dependence-based power over *j* and he is in a position to make credible threats to *j* (persuasive power), then *i* has the power to affect *j*'s behaviour thereby having a power of influencing *j*. By way of example, consider the situation involving two hypothetical countries called A and B in a conflict situation. In the initial situation, A has two actions it may take: either puts an embargo on against B or it does nothing. B can respond in two different ways to A's action: either by moving a military attack against A (thereby starting a war against A) or by doing nothing. Thus, after A's initial move B has a dependence-based power over A. In fact, after A's initial move, A wants to avoid a war against B and the possibility of avoiding the war depends on what B decides to do. Now, suppose B has the power of persuading A that if A makes an embargo against B then B will move a military attack against A. Hence, because of its dependence-based power and persuasive power over A, B has the power of inducing A to refrain from putting an embargo on it, thereby having an influencing power over A.

To sum it up, a comprehensive formal model and ontology of power should allow:

- to reason about knowledge of agents in order to study the discretionary aspect of their *powers of*;
- to clarify the nature of different kinds of *social power* and in particular:
  - an agent *i*'s *power of influencing* another agent *j* as *i*'s capacity to affect *j*'s intentions in such a way that *j* will do or will refrain from doing a certain action;
  - an agent *i*'s *power of persuading* another agent *j* as *i*'s capacity to induce *j* to believe or to abstain from believing certain things;
  - an agent *i*'s *dependence-based power over* another agent *j* as *j*'s dependence on a certain action of *i* for the achievement of his goals.

A number of logics have been devised to represent societies of agents in the both the philosophy and computer science literature. In particular, there exist variants of Alternating-time Temporal Logic (ATL) [1, 2, 39, 48, 49] and as 'Seeing To It That' (STIT) theories [5, 11, 28, 35, 45]. Albeit a good starting point, we claim that they miss the mark for a comprehensive ontology of powers, as their models lack one or more of the concepts that we consider essential to a theory of powers: what we need is a logic whose language allows to express the intricate relationship between the concepts that are involved in this ontology: action, agency, knowledge and goals.

[3] The view of *beliefs as reasons for intending* has been extensively debated in the philosophical literature (see, e.g. [17]).

(1) *Agency*: there are at least two candidates for a logic of agency. ATL logics [2, 39] are designed to express what coalitions can achieve by cooperating. ATL has coalition modalities $\langle\langle G\rangle\rangle$ where G is an arbitrary group of agents G. The ATL formula $\langle\langle G\rangle\rangle X\varphi$ means that coalition G has a collective strategy to ensure that, no matter what the other agents do, φ will be true in the next state. In deliberative STIT theories [5, 28] modal operators of the form $[G\ cstit:]$, are meant to capture a notion of choice *being made* by agents in G. A modal operator $\Box$ of historic necessity, whose dual operator of historic possibility is $\Diamond-$, should also be part of the language to talk about what could have happened otherwise. In formulas, $[G\ cstit:\varphi]$ and $\Diamond[G\ cstit:\varphi]$, respectively, mean that G sees to it that φ and G can see to it that φ.
(2) *Knowledge*: in [26], we have already argued for the relevance of the refined language of STIT theories when it comes to mixing the logic of agency with epistemic notions. In ATL this leads to difficulties, as acknowledged in [29], and various complexifications of the original semantics have been put forward [30, 31].
(3) *Goals*: one of the main differences between an agent and a mere process in a distributed system is a goal-oriented decision process that guides which action the agent takes. Within a society of agents, some individuals might have the incentive to collude or simply take benefit of other agents' actions to achieve their goals.
(4) *Action terms*: it is usually considered an advantage of logics of agency that they abstract away from the actions proper; however, this is also a drawback. Actually modern philosophy of action led in particular by Davidson [18] generally does not even consider a treatment of action without explicit reference to the action terms. In our case, we have seen that action terms are difficult to avoid when talking about influencing power and dependence-based power. For instance, an agent i's power of influencing another agent j consists in i's power to influence j to do or to refrain from doing a certain action in his *action repertoire*. As the elements of an agent's action repertoire are described by action terms (see [34] for a discussion on this issue), action terms become necessary in order to define influencing power.

Yet, none of the approaches cited above were intended to support the types of powers in social interactions that we described. In particular, what is still missing in the logical literature is an integration of (i) the expressiveness of logics of action and agency with (ii) the expressiveness of a logic of mental attitudes (so-called BDI logic[4]) and (iii) dynamic logic [24] where actions of agents are explicit.[5]

In this article, we try to fill this gap by developing a logic allowing to reason about mental states of agents, action occurrences, and agentive and group powers. This will enable us to capture some important properties of *power of* and *social power*.[6]

The article is organized as follows. In Section 2, we present the syntax and the semantics of a logic of powers and mental states called $\mathcal{IAL}$ (*Intentional Agency Logic*). This logic is based on a combination of the logic of group actions, powers and capabilities proposed by [25, 33, 36], and a simple BDI logic. In Section 3, we axiomatize $\mathcal{IAL}$ and study some of its properties. In Section 4, we present some interesting theorems of $\mathcal{IAL}$, and in Section 5 we briefly compare our system with some existing logics of cooperation and multi-agent interaction. In the last and main part of

[4] See e.g. [50, 54] for a survey on BDI (belief, desire, intention) logics.

[5] For a similar attempt to introduce mental attitudes in a logic of strategic interaction, see [37].

[6] Note that while ATL deals with strategies—or sequences of action—it will not be the case in the present contribution. This is certainly an important limitation. However, a strategic logic of actual agency is still an open problem, and the interesting bits of social powers are already expressible with atomic strategies.

this article (Section 6), we exploit $\mathcal{IAL}$ to formalize and study the properties of different kinds of agentive power.

2 A logic of powers and mental states: syntax and semantics

The logic $\mathcal{IAL}$ (*Intentional Agency Logic*) combines the expressiveness of a logic of actions and mental states with the expressiveness of a logic of social interaction. $\mathcal{IAL}$ is based on a combination of the logic of group actions, powers and capabilities proposed in [25, 33] and a simple BDI logic. The logic of [25, 33] is a variant of dynamic logic, which embeds STIT logic and Coalition Logic (CL). It allows to represent group actions both in terms of final outcomes ('the agents in group C ensure that φ is true by acting together') and in terms of concrete actions ('the agents in group C do the (joint) action δ_C together'). Symmetrically, it allows to represent group capabilities both in terms of final outcomes ('the agents in group C can act together to ensure that φ') and in terms of concrete actions ('the agents in group C can do the (joint) action δ_C together').

On top of that expressive logic of action and agency, we introduce modal operators for beliefs and goals of agents. We here consider intentional actions only, i.e. an action is taken by an agent only if it is the agent's goal to do so.

We denote by $AGT = \{1, 2, \ldots, n\}$ the finite and non-empty set of agents and we denote by $ACT = \{a, b, \ldots\}$ the finite and non-empty set of atomic actions. To every agent i, we associate a set of ordered pairs $Act(i) = \{i{:}a \mid a \in ACT\}$. We note 2^{AGT*} the set of non-empty subsets of AGT. We call *coalitions* the elements in 2^{AGT*}. We note Δ the set of all possible combinations of actions by the agents in AGT, i.e. $\Delta = \prod_{i \in AGT} Act(i)$. Every element $\delta = (\delta_1, \delta_2, \ldots, \delta_n)$ in Δ is then a vector of individual actions, one for every agent in AGT, commonly called an *action profile*. We denote by δ_C the *coalitional action* of the members of coalition C in action profile δ. For convenience, in the case of singleton coalitions we often write δ_i instead of $\delta_{\{i\}}$.

The logic $\mathcal{IAL}$ is a propositional modal logic. It extends propositional logic over a set of propositional variables $\Pi = \{p, q, \ldots\}$. The language $\mathcal{L}_{\mathcal{IAL}}$ is given by the following BNF:

$$\varphi \quad ::= \quad p \quad | \quad \neg\varphi \quad | \quad \varphi \vee \varphi \quad | \quad \Box\varphi \quad | \quad [i{:}a]\varphi \quad | \quad \mathbf{Does}_C\varphi \quad | \quad \mathbf{Bel}_i\varphi \quad | \quad \mathbf{Goal}_i\varphi$$

where p ranges over the set of propositional variables Π, i ranges over the set of agents AGT, a ranges over the set of actions ACT and C ranges over the set of coalitions 2^{AGT*}.

Several abbreviations are used in this article. The classical Boolean connectives $\wedge$, $\rightarrow$, $\leftrightarrow$, $\top$ (tautology) and $\bot$ (contradiction) are introduced in the usual way. $\langle i{:}a\rangle\varphi$ abbreviates $\neg[i{:}a]\neg\varphi$ and $\Diamond\varphi$ abbreviates $\neg\Box\neg\varphi$. Moreover, $\langle\delta_C\rangle\varphi$ abbreviates $\bigwedge_{i \in C}\langle\delta_i\rangle\varphi$ and $[\delta_C]\varphi$ abbreviates $\neg\langle\delta_C\rangle\neg\varphi$.

The formula $\mathbf{Bel}_i\varphi$ reads 'agent i believes that φ'. The operator $\mathbf{Goal}_i$ is intended to represent agent i's chosen goals and is similar to the operator introduced in [14]. The formula $\mathbf{Goal}_i\varphi$ then reads 'agent i has decided to pursue φ' or, for short, 'agent i wants that φ'.

The reading of $[i{:}a]\varphi$ is 'φ holds after every occurrence of action a performed by agent i'. Hence, $[i{:}a]\bot$ formalizes that the agent i does not do action a. The formula $\mathbf{Does}_C\varphi$ reads 'coalition C brings about that φ whatever the agents in $AGT \setminus C$ do' or simply 'coalition C brings about that φ'. For notational convenience, we write $\mathbf{Does}_i\varphi$ instead of $\mathbf{Does}_{\{i\}}\varphi$ for every $i \in AGT$. Finally, the operator $\Box$ is used to quantify over action profiles. Thus, $\Box\varphi$ has to be read 'φ holds whatever the agents do', or simply 'φ is necessarily true'.

As we show in Section 3, given the semantical constraints that we impose on our logic, the more natural reading of $\langle i{:}a\rangle\varphi$ is 'agent i does a and φ will be true afterwards', and the more natural reading of $[i{:}a]\varphi$ is 'if agent i does a then φ will be true afterwards'.

We can offer some more intuitions about the language and have a first glimpse of the expressions of powers. The formula $\Diamond\varphi$ is read 'there exists a world corresponding to a choice for every agent in which φ is true' or simply 'φ can/may be true'. The operators $\Diamond$ and $\mathbf{Does}_C$ (respectively $\mathbf{Does}_i$) can be exploited for expressing what a coalition C (respectively a single agent i) is able to bring about. Hence, the formula $\Diamond\mathbf{Does}_C\varphi$ reads 'there exists a world corresponding to a choice for every agent in which coalition C brings about that φ' or simply ' coalition C can bring about that φ'.

2.1 Model definition

Our semantics is in terms of possible world models. Such models are based on the following very abstract definition of Kripke frames: a set of states and a collection of binary relations over that set.

Definition 2.1 (Kripke frame)
A *Kripke frame* is a tuple

$$F=(W,H,\{R_{i:a}\mid i\in AGT,a\in ACT\},\{D_C\mid C\in 2^{AGT*}\},\{B_i\mid i\in AGT\},\{G_i\mid i\in AGT\}).$$

Each binary relation over W interprets one of the modalities of the language of $\mathcal{IAL}$.

In the logic $\mathcal{IAL}$, there is a one-to-one correspondence between worlds and actions profiles: every world $w\in W$ corresponds to a unique strategy profile that is played at w. Therefore, $H(w)=\{v\mid(w,v)\in H\}$ gives the set of action profiles which are alternative to the action profile played at w. However, H can also be viewed as the relation that models historic possibility in the sense of STIT theory, i.e. $H(w)$ gives the set of historic alternatives to the state w. For every agent i and action $a\in ACT$, $R_{i:a}(w)=\{v\mid(w,v)\in R_{i:a}\}$ models the set of the possible consequences of the execution of action a by agent i in world w. For every coalition C, $D_C(w)=\{v\mid(w,v)\in D_C\}$ is the set of worlds that the coalition C brings about at w. Finally, for every agent i, $B_i(w)=\{v\mid(w,v)\in B_i\}$ (respectively $G_i(w)=\{v\mid(w,v)\in G_i\}$) is the worlds that i considers plausible (respectively that i intends) at w.

In order to reflect the intuitions, we have to impose a few constraints on Kripke frames. All the free variables in the below constraint formulations are assumed to be universally quantified. w, w', u and v are some states in W; i and j are agent identifiers; C and C' are coalitions.

We first give the most elementary properties, that do not involve interactions between different relations.

(C.i) *H is an equivalence relation*
For every $w\in W$, $H(w)$ is a cluster of historic alternatives as traditionally modelled in the domain of STIT theory [5, 28].

(C.ii) *D_C is serial*
A coalition C always brings about something.

(C.iii) *B_i is transitive, euclidean and serial*
An agent i is assumed to have positive and negative introspection and its beliefs are consistent.

(C.iv) *G_i is serial*
An agent i always has a goal.

(C.v) *if $u\in D_{AGT}(w)$ and $v\in D_{AGT}(w)$ then $u=v$*
The grand coalition AGT brings about exactly one outcome.

The next constraints concern the way the relations interact with each other. Arguably the trickiest constraints are those involving the actions and will be introduced later. On the other hand, the

following two conditions on the relations between mental attitudes formalize assumptions that are standard in BDI logics (e.g. [14, 41]).

(C.vi) $B_i(w) \cap G_i(w) \neq \emptyset$
This is a condition of weak realism, according to which the set of i's belief-accessible worlds and the set of i's goal-accessible worlds are never disjoint.

(C.vii) *if* $w' \in B_i(w)$ *then* $G_i(w') = G_i(w)$
Worlds that are compatible with i's goals are also compatible with i's goals from those worlds which are compatible with i's beliefs.

(C.viii) *if* $w' \in B_i(w)$ *and* $v \in H(w)$ *then there is a* u *such that* $u \in H(w') \cap B_i(v)$
It is a semantic condition of confluence between the relations B_i and H.

The following constraints concern the accessibility relation corresponding to the actions. We define $R_{\delta_C} = \bigcap_{i \in C} R_{\delta_i}$ in order to make the presentation of some of them run more smoothly.

(C.ix) $\bigcup_{a \in ACT} R_{i:a}(w) \neq \emptyset$
An agent i always has at least one action to be performed. In other words, agents are never passive.

(C.x) *for* $B \cap C = \emptyset$, *if there is a* $u \in H(w)$ *such that* $R_{\delta_B}(u) \neq \emptyset$ *and there is a* $v \in H(w)$ *such that* $R_{\delta'_C}(v) \neq \emptyset$ *then there is a* $w' \in H(w)$ *such that* $R_{\delta_B}(w') \neq \emptyset$ *and* $R_{\delta'_C}(w') \neq \emptyset$
Given two disjoint coalitions B and C, if the agents in B can do together an action δ_B in some historic alternative and the agents in C can do together an action δ'_C in some other historic alternative, then there is a historic alternative where the agents in $B \cup C$ do together the collective action (δ_B, δ'_C). This constraint corresponds to the independence of agents in STIT and to the superadditivity property in social choice theory [39].

(C.xi) *if* $u \in R_{i:a}(w)$ *and* $v \in R_{j:b}(w)$ *then* $u = v$
All actions executed in one world lead to the same unique world. This semantic constraint justifies the reading of $\langle i{:}a \rangle \varphi$ and $[i{:}a]\varphi$ as 'agent i does a and φ will be true after the occurrence of a performed by i' and 'if agent i does a then φ will be true after the occurrence of a performed by i'.

(C.xii) *if* $i{:}a \neq i{:}b$ *then* $R_{i:a}(w) = \emptyset$ *or* $R_{i:b}(w) = \emptyset$
An agent cannot execute more than one action at a time.

(C.xiii) *if* $R_{\delta_C}(w) \neq \emptyset$ *then if* $v \in H(w)$ *and* $u \in R_{\delta_C}(v)$ *then* $u \in D_C(w)$
If coalition C performs the joint action δ_C at w and there is a historic alternative v to w where δ_C is also executed and leads to state u, then u is among the states brought about by the coalition C at w.

(C.xiv) *if* $R_{\delta_C}(w) \neq \emptyset$ *and* $u \in D_C(w)$ *then there exists* $v \in H(w)$ *such that* $u \in R_{\delta_C}(v)$
If the coalitional action δ_C is executed at w and u is among the states brought about by C at w, then there is a historic alternative to w where δ_C is executed leading to u.

(C.xv) *if* $R_{i:a}(w) \neq \emptyset$ *then for all* $u \in G_i(w)$ *we have* $R_{i:a}(u) \neq \emptyset$
If a is performed by i then a is performed by i at each of i's goal-accessible worlds. This means that actions are intentional: i performs a only if a is performed in all worlds that i intends.

(C.xvi) *if* $u \in R_{i:a}(w)$ *and* $u' \in B_i(v)$ *then there is a* $w' \in B_i(w)$ *such that* $u' \in R_{i:a}(w')$
What an agent i believes at a world u after performing an action a only depends on what the agent believed at the previous world w before performing a.

We are now ready to give the definition of the models of $\mathcal{IAL}$.

Definition 2.2 (model of $\mathcal{IAL}$)
A model of $\mathcal{IAL}$ is a tuple $M = (W, H, \{R_{i:a}\}, \{D_C\}, \{B_i\}, \{G_i\}, \pi)$ where:

- $(W, H, \{R_{i:a}\}, \{D_C\}, \{B_i\}, \{G_i\})$ is a Kripke frame satisfying constraints (C.i) to (C.xvi);

- $\pi : \Pi \longrightarrow 2^W$ is a valuation function.

2.2 Truth conditions

Given a model M, a world w and a formula φ, we write $M, w \models \varphi$ to mean that φ is true at world w in M, under the basic semantics. The rules defining the truth conditions of formulas of our logic are inductively defined as follows.

$M, w \models p$	iff	$w \in \pi(p)$
$M, w \models \neg\varphi$	iff	not $M, w \models \varphi$
$M, w \models \varphi \vee \psi$	iff	$M, w \models \varphi$ or $M, w \models \psi$
$M, w \models \Box\varphi$	iff	for all $w' \in W$ if $w' \in H(w)$ then $M, w' \models \varphi$
$M, w \models [i{:}a]\varphi$	iff	for all $w' \in W$ if $w' \in R_{i:a}(w)$ then $M, w' \models \varphi$
$M, w \models \mathbf{Bel}_i\varphi$	iff	for all $w' \in W$ if $w' \in B_i(w)$ then $M, w' \models \varphi$
$M, w \models \mathbf{Goal}_i\varphi$	iff	for all $w' \in W$ if $w' \in G_i(w)$ then $M, w' \models \varphi$
$M, w \models \mathbf{Does}_C\varphi$	iff	for all $w' \in W$ if $w' \in D_C(w')$ then $M, w' \models \varphi$

We write $\models \varphi$ if formula φ is *valid* in all $\mathcal{IAL}$ models, i.e. $M, w \models \varphi$ for every $\mathcal{IAL}$ model M and world w in M.

3 Axiomatization

Let $\blacksquare$ be an arbitrary modality of necessity and $\blacklozenge$ its dual modality of possibility. We are going to use some standard axioms of modal logic. We are assuming some familiarity with normal modal logic and refer to [13] for details. Axiom **K** is $\blacksquare(p \to q) \to (\blacksquare p \to \blacksquare q)$. Axiom **D** ($\blacksquare p \to \blacklozenge p$) is canonical for serial frames.[7] Axiom **T** ($\blacksquare p \to p$) is canonical for reflexive frames. Axiom **B** ($p \to \blacksquare\blacklozenge p$) is canonical for symmetric frames.[8] Axiom **4** ($\blacksquare \to \blacksquare\blacksquare p$) is canonical for transitive frames. Axiom **5** ($\blacklozenge p \to \blacksquare\blacklozenge p$) is canonical for Euclidian frames.

We call $\mathcal{IAL}$ the logic axiomatized by the principles given in Figure 1 together with the rule of Modus Ponens (from $\vdash_{\mathcal{IAL}} \varphi \to \psi$ and $\vdash_{\mathcal{IAL}} \varphi$ infer $\vdash_{\mathcal{IAL}} \psi$) and the rule of necessitation for every modality (from $\vdash_{\mathcal{IAL}} \varphi$ infer $\vdash_{\mathcal{IAL}} \blacksquare\varphi$). We write $\vdash_{\mathcal{IAL}} \varphi$ if φ is derivable in the proof system of $\mathcal{IAL}$.

We now discuss the axioms and refine our intuitions.

KT5$_\Box$ makes $\Box$ a modality of historic necessity [44]. A historic alternative is among its own alternatives; an alternative is an alternative to all alternatives. The formula $\Box\varphi$ should then be true when φ is true no matter how the future unfolds. Note that since $\Box$ obeys **T** and **5** axioms, it is an **S5** modality. $\mathbf{KD}_{\mathbf{Does}}$ ensures that a coalition cannot bring about inconsistent states of affairs. $\mathbf{KD45}_{\mathbf{Bel}}$ corresponds to the standard axiomatization of doxastic logic [27]: agents have positive and negation introspection over their beliefs and cannot have inconsistent beliefs. $\mathbf{KD}_{\mathbf{Goal}}$ ensures that an agent cannot have inconsistent goals. According to Axiom $\mathbf{Alt}_{\mathbf{Does}}$ the grand coalition AGT always produces deterministic effects. Axiom $\mathbf{D}_{\mathbf{Bel,Goal}}$ is a weak realism axiom that relates an agent's beliefs with his goals, whereas **PosIntr** and **NegIntr** are principles of positive and negative introspection for goals [20]. Axiom $\mathbf{Confl}_{\mathbf{Bel},\Box}$ says that if it is historically possible that i believes that φ is true

[7] We say that a formula of a logic is *canonical* for a class of frames when it forces the models of the logic to satisfy some property. This information will be useful in the proof of Theorem 3.1.

[8] Axiom **B** will not be explicitly used in the axiomatization.

(**ProTau**)	All tautologies of propositional calculus
(**KT5**$_{\Box}$)	All KT5-theorems for $\Box$
(**KD**$_{Does}$)	All KD-theorems for every **Does**$_C$
(**KD45**$_{Bel}$)	All KD45-theorems for every **Bel**$_i$
(**KD**$_{Goal}$)	All KD-theorems for every **Goal**$_i$
(**K**$_{Act}$)	All K-theorems for every $[i{:}a]$
(**Alt**$_{Does}$)	$\neg\mathbf{Does}_{AGT}\neg\varphi \rightarrow \mathbf{Does}_{AGT}\varphi$
(**D**$_{Bel,Goal}$)	$\mathbf{Goal}_i\varphi \rightarrow \neg\mathbf{Bel}_i\neg\varphi$
(**PosIntr**)	$\mathbf{Goal}_i\varphi \rightarrow \mathbf{Bel}_i\mathbf{Goal}_i\varphi$
(**NegIntr**)	$\neg\mathbf{Goal}_i\varphi \rightarrow \mathbf{Bel}_i\neg\mathbf{Goal}_i\varphi$
(**Confl**$_{Bel,\Box}$)	$\Diamond\mathbf{Bel}_i\varphi \rightarrow \mathbf{Bel}_i\Diamond\varphi$
(**Active**)	$\bigvee_{a\in ACT}\langle i{:}a\rangle\top$
(**Single**)	$\langle i{:}a\rangle\top \rightarrow [i{:}b]\bot$ if $a \neq b$
(**Alt**$_{Act}$)	$\langle i{:}a\rangle\varphi \rightarrow [j{:}b]\varphi$
(**Indep**)	$(\Diamond\langle\delta_C\rangle\top \wedge \Diamond\langle\delta'_B\rangle\top) \rightarrow \Diamond(\langle\delta_C\rangle\top \wedge \langle\delta'_B\rangle\top)$ if $B \cap C = \emptyset$
(**DoesDef**)	$\langle\delta_C\rangle\top \rightarrow (\mathbf{Does}_C\varphi \leftrightarrow \Box(\langle\delta_C\rangle\top \rightarrow [\delta_C]\varphi))$
(**IntAct**)	$\langle i{:}a\rangle\top \rightarrow \mathbf{Goal}_i\langle i{:}a\rangle\top$
(**NF**)	$\mathbf{Bel}_i[i{:}a]\varphi \rightarrow [i{:}a]\mathbf{Bel}_i\varphi$

FIG. 1. Axioms of $\mathcal{IAL}$.

then i believes that it is historically possible that φ is true. Axiom **Active** says that an agent always performs at least one action. Axiom **Single** says that if an action $a \in ACT$ is performed by agent i, no other action in ACT is performed by i. **Active** and **Single** together capture that every agent always executes one and only one action from his repertoire. $\mathbf{Alt}_{\mathbf{Act}}$ forces the fact that if the action a performed by agent i leads to φ then every executed action leads to φ, may it be an action of i or any other agent. (Although, an action b in ACT different from a that is also executed will be impossible here due to **Single**.) Axiom **Indep** says that for disjoint coalitions B and C, if agents in C can do together a certain combination of actions δ_C and agents in B can do together a certain combination of actions δ'_B then agents in $B \cup C$ can do together a combination of actions (δ_C, δ'_B). This axiom is the 'actional' counterpart of the axiom of *independence of agents* (called AIA_k) in the deliberative STIT theories [5, Chapter 17]. Axiom **DoesDef** is a mere local definition of the modality $\mathbf{Does}_{\mathrm{C}}$: if the agents in the coalition C execute the joint action δ_C then C brings about that φ if and only if, at every historic alternative, if the agents C execute δ_C then φ will be true thereafter. According to Axiom **IntAct**, an agent does action a only if he intends to do a. Thus, in our formal model the actions performed by an agent are intentional actions (see [34] for a discussion of this assumption). According to Axiom **NF** (*no forgetting* Axiom), if an agent believes that φ will be true after he performs a then, after he performs action a, he will believe that φ. Similar principles for the interaction between belief and action or between knowledge and action (sometimes also called *perfect recall*) have been studied in [21, 47].

We can prove that $\mathcal{IAL}$ is *sound* and *complete* with respect to the class of $\mathcal{IAL}$ models.

THEOREM 3.1
$\mathcal{IAL}$ is determined by the class of models of $\mathcal{IAL}$.

PROOF. It is a routine to prove soundness. It is also routine to check that all axioms of the logic $\mathcal{IAL}$ are in the Sahlqvist class [6]. This means that the axioms are all expressible as first-order conditions on models and are complete with respect to the defined classes of models. (This can be established e.g. by applying the SQEMA algorithm [15] and performing some predicate logic formula rewriting.)

In the following table, we sum up the correspondence between the frame constraints and the axioms.

SEMANTIC	SYNTACTIC	SEMANTIC	SYNTACTIC
(C.i)	$\mathbf{KT5}_{\Box}$	(C.ii)	$\mathbf{KD}_{\mathbf{Does}}$
(C.iii)	$\mathbf{KD45}_{\mathbf{Bel}}$	(C.iv)	$\mathbf{KD}_{\mathbf{Goal}}$
(C.v)	$\mathbf{Alt}_{\mathbf{Does}}$	(C.vi)	$\mathbf{D}_{\mathbf{Bel,Goal}}$
(C.vii)	**PosIntr** and **NegIntr**	(C.viii)	$\mathbf{Confl}_{\mathbf{Bel},\Box}$
(C.ix)	**Active**	(C.x)	**Indep**
(C.xi)	$\mathbf{Alt}_{\mathbf{Act}}$	(C.xii)	**Single**
(C.xiii) and (C.xiv)	**DoesDef**	(C.xv)	**IntAct**
(C.xvi)	**NF**		

4 Some properties of $\mathcal{IAL}$

Observe that due to determinism of $\mathbf{Does}_{\mathrm{AGT}}$ (constraint (C.v)), we may conceive the unique state w' such that $\{w'\} = D_{\mathrm{AGT}}(w)$ as the temporal successor of w. Therefore, we can read the formula $\mathbf{Does}_{\mathrm{AGT}}\varphi$ as 'φ will be true in the next state'. Thus, $\mathbf{Does}_{\mathrm{AGT}}$ can be interpreted as a standard operator **X** (*next*) of temporal logic.

DEFINITION 4.1
The **X** operator is defined as follows:

$$\mathbf{X}\varphi \stackrel{\text{def}}{=} \mathbf{Does}_{\mathrm{AGT}}\varphi$$

The next proposition highlights some interesting properties of actions and goals in $\mathcal{IAL}$.

PROPOSITION 4.2
For every $i \in AGT$, $a \in ACT$, $\delta \in \Delta$ and $B, C \in 2^{AGT*}$:

$$\vdash_{\mathcal{IAL}} (\langle\delta_C\rangle\top \wedge \mathbf{Does}_{\mathrm{C}}\varphi) \rightarrow \Box(\langle\delta_C\rangle\top \rightarrow \mathbf{Does}_{\mathrm{C}}\varphi) \tag{4.1a}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Does}_{\mathrm{B}}\varphi \rightarrow \mathbf{Does}_{\mathrm{B\cup C}}\varphi \tag{4.1b}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Goal}_i\langle i{:}a\rangle\top \vee \mathbf{Goal}_i[i{:}a]\bot \tag{4.1c}$$

PROOF. We prove $\mathcal{IAL}$-theorem 4.1b as an example.

$\mathbf{Does}_{\mathrm{B}}\varphi$ is equivalent to $\bigvee_{\delta_B}(\langle\delta_B\rangle\top \wedge \Box(\langle\delta_B\rangle\top \rightarrow [\delta_B]\varphi))$ (by Axiom **DoesDef** and Axiom **Active**). The latter implies $\bigvee_{\delta_B,\delta_C}(\langle\delta_B\rangle\top \wedge \langle\delta_C\rangle\top \wedge \Box(\langle\delta_B\rangle\top \rightarrow [\delta_B]\varphi))$ (by Axiom **Active**), which in turn implies $\bigvee_{\delta_{B\cup C}}(\langle\delta_{B\cup C}\rangle\top \wedge \Box(\langle\delta_{B\cup C}\rangle\top \rightarrow [\delta_{B\cup C}]\varphi))$. The latter is equivalent to $\mathbf{Does}_{\mathrm{B\cup C}}\varphi$. ∎

According to the $\mathcal{IAL}$-theorem 4.1a in Proposition 4.2, if coalition C brings about that φ by doing the joint action δ_C then, necessarily, if coalition C does the joint action δ_C then it will bring about that φ. $\mathcal{IAL}$-theorem 4.1b expresses monotonicity for coalitions: if coalition B brings about that φ then coalition $B \cup C$ brings about that φ. According to the $\mathcal{IAL}$-theorem 4.1c, at each moment an agent either decides (intends) to do an action or decides (intends) not to do it.

The following proposition highlights some other noteworthy properties of $\mathcal{IAL}$.

PROPOSITION 4.3
For every $i \in AGT$, $a \in ACT$, $\delta \in \Delta$ and $B, C \in 2^{AGT*}$ such that $B \cap C = \emptyset$:

$$\vdash_{\mathcal{IAL}} (\Diamond \mathbf{Does}_{\mathrm{B}} \varphi \wedge \Diamond \mathbf{Does}_{\mathrm{C}} \psi) \rightarrow \Diamond \mathbf{Does}_{\mathrm{B \cup C}} (\varphi \wedge \psi) \tag{4.2a}$$

$$\vdash_{\mathcal{IAL}} (\Diamond \mathbf{Does}_{\mathrm{B}} \varphi \wedge \Diamond \mathbf{Does}_{\mathrm{C}} \neg \varphi) \rightarrow \bot \tag{4.2b}$$

$$\vdash_{\mathcal{IAL}} \langle \delta_C \rangle \top \rightarrow ([\delta_C] \varphi \leftrightarrow \mathbf{X} \varphi) \tag{4.2c}$$

$$\vdash_{\mathcal{IAL}} \mathbf{X} \varphi \leftrightarrow \neg \mathbf{X} \neg \varphi \tag{4.2d}$$

$$\vdash_{\mathcal{IAL}} \Box \mathbf{X} \varphi \rightarrow \Box \mathbf{Does}_{\mathrm{C}} \varphi \tag{4.2e}$$

$$\vdash_{\mathcal{IAL}} \langle i{:}a \rangle \top \leftrightarrow \mathbf{Goal}_i \langle i{:}a \rangle \top \tag{4.2f}$$

$$\vdash_{\mathcal{IAL}} \langle i{:}a \rangle \top \leftrightarrow \mathbf{Bel}_i \langle i{:}a \rangle \top \tag{4.2g}$$

$$\vdash_{\mathcal{IAL}} [i{:}a] \bot \leftrightarrow \mathbf{Goal}_i [i{:}a] \bot \tag{4.2h}$$

$$\vdash_{\mathcal{IAL}} [i{:}a] \bot \leftrightarrow \mathbf{Bel}_i [i{:}a] \bot \tag{4.2i}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Bel}_i \Box \varphi \rightarrow \Box \mathbf{Bel}_i \varphi \tag{4.2j}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Bel}_i \mathbf{X} \varphi \rightarrow \mathbf{X} \mathbf{Bel}_i \varphi \tag{4.2k}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Bel}_i \Box \mathbf{X} \varphi \rightarrow \Box \mathbf{X} \mathbf{Bel}_i \varphi \tag{4.2l}$$

$$\vdash_{\mathcal{IAL}} \mathbf{Bel}_i \mathbf{X} \Box \varphi \rightarrow \mathbf{X} \Box \mathbf{Bel}_i \varphi \tag{4.2m}$$

$$\vdash_{\mathcal{IAL}} (\mathbf{Goal}_i \varphi \wedge \mathbf{Bel}_i (\varphi \rightarrow \langle i{:}a \rangle \top)) \rightarrow \mathbf{Goal}_i \langle i{:}a \rangle \top \tag{4.2n}$$

PROOF. We prove the $\mathcal{IAL}$-theorems 4.2a, 4.2e and 4.2j as an example.

We first prove 4.2a. Suppose $B \cap C = \emptyset$. $\Diamond \mathbf{Does}_{\mathrm{B}} \varphi \wedge \Diamond \mathbf{Does}_{\mathrm{C}} \psi$ is equivalent to $\bigvee_{\delta_B, \delta_C} (\Diamond (\langle \delta_B \rangle \top \wedge \Box (\langle \delta_B \rangle \top \rightarrow [\delta_B] \varphi)) \wedge \Diamond (\langle \delta_C \rangle \top \wedge \Box (\langle \delta_C \rangle \top \rightarrow [\delta_C] \psi)))$ (by Axiom **DoesDef** and Axiom **Active**). As $B \cap C = \emptyset$, the latter implies $\bigvee_{\delta_{B \cup C}} \Diamond (\langle \delta_{B \cup C} \rangle \top \wedge \Box (\langle \delta_B \rangle \top \rightarrow [\delta_B] \varphi) \wedge \Box (\langle \delta_C \rangle \top \rightarrow [\delta_C] \psi))$ (by Axiom **Indep**) which in turn implies $\bigvee_{\delta_{B \cup C}} \Diamond (\langle \delta_{B \cup C} \rangle \top \wedge \Box (\langle \delta_{B \cup C} \rangle \top \rightarrow ([\delta_B] \varphi \wedge [\delta_C] \psi)))$. By the valid equivalence $([\delta_B] \varphi \wedge [\delta_C] \psi) \leftrightarrow [\delta_{B \cup C}] (\varphi \wedge \psi)$ the latter is equivalent to $\bigvee_{\delta_{B \cup C}} \Diamond (\langle \delta_{B \cup C} \rangle \top \wedge \Box (\langle \delta_{B \cup C} \rangle \top \rightarrow [\delta_{B \cup C}] (\varphi \wedge \psi)))$, which in turn is equivalent to $\Diamond \mathbf{Does}_{\mathrm{B \cup C}} (\varphi \wedge \psi)$ (by Axiom **DoesDef** and Axiom **Active**).

We now prove 4.2e. $\Box \mathbf{X} \varphi$ is equivalent to $\Box \bigvee_{\delta} (\langle \delta \rangle \top \wedge \Box (\langle \delta \rangle \top \rightarrow [\delta] \varphi))$ (by Axiom **DoesDef** and Axiom **Active**). The latter implies $\Box \bigvee_{\delta} (\langle \delta \rangle \top \wedge [\delta] \varphi)$ (by Axiom T for $\Box$) which in turn implies $\Box \bigvee_{\delta_C} (\langle \delta_C \rangle \top \wedge [\delta_C] \varphi)$. The latter implies $\Box \bigvee_{\delta_C} \langle \delta_C \rangle \varphi$ which in turn implies $\Box \bigwedge_{\delta_C} [\delta_C] \varphi$ (by Axiom $\mathbf{Alt}_{\mathbf{Act}}$). From the latter it follows that $\bigwedge_{\delta_C} \Box [\delta_C] \varphi$ which in turn implies $\bigwedge_{\delta_C} \Box (\langle \delta_C \rangle \top \rightarrow [\delta_C] \varphi)$. From the latter, by Axiom 4 for $\Box$, it follows that $\Box \bigwedge_{\delta_C} \Box (\langle \delta_C \rangle \top \rightarrow [\delta_C] \varphi)$ which in turn implies $\Box \bigvee_{\delta_C} (\langle \delta_C \rangle \top \wedge \Box (\langle \delta_C \rangle \top \rightarrow [\delta_C] \varphi))$ (by Axiom **Active**). The latter is equivalent to $\Box \mathbf{Does}_{\mathrm{C}} \varphi$ (by Axiom **DoesDef**).

We now prove 4.2j. $\mathbf{Bel}_i \Box \varphi$ implies $\Box \Diamond \mathbf{Bel}_i \Box \varphi$ (by Axiom K, Axiom T and necessitation rule for $\Box$) which in turn implies $\Box \mathbf{Bel}_i \Diamond \Box \varphi$ (by Axiom K and necessitation rule for $\Box$ and Axiom $\mathbf{Confl}_{\mathbf{Bel}, \Box}$). The latter implies $\Box \mathbf{Bel}_i \Box \varphi$ (by the equivalence $\Diamond \Box \varphi \leftrightarrow \Box \varphi$ which is derivable by Axiom T and Axiom 5 for $\Box$) which in turn implies $\Box \mathbf{Bel}_i \varphi$ (by Axiom K, Axiom T and necessitation rule for $\Box$, Axiom K and necessitation rule for $\mathbf{Bel}_i$). ∎

$\mathcal{IAL}$-theorem 4.2a in Proposition 4.3 says that two disjoint coalitions can combine their efforts to ensure a conjunction of outcomes. This corresponds to the *superadditivity* axiom of Coalition Logic [39] of the form $[B] \varphi \wedge [C] \psi \rightarrow [B \cup C] (\varphi \wedge \psi)$ (when $B \cap C = \emptyset$). $\mathcal{IAL}$-theorem 4.2b, which is a direct consequence of $\mathcal{IAL}$-theorem 4.2a, says that two disjoint coalitions can never bring about conflicting effects. According to the $\mathcal{IAL}$-theorem 4.2c, if the agents in coalition C execute the joint action δ_C then, φ will be true in the next state if and only if φ will be true after the execution of δ_C. $\mathcal{IAL}$-theorem 4.2d shows the tight correspondence between our definition of *next* and the

standard operator *next* of linear temporal logic (LTL). According to the $\mathcal{IAL}$-theorem 4.2e, if φ will be necessarily true in the next state then, necessarily, coalition C will ensure φ in the next state. $\mathcal{IAL}$-theorems 4.2f-4.2i are about the relations between intention, belief and action occurrences: an agent i executes (respectively does not execute) an action a if and and only if i has the intention to do (respectively not to do) a, an agent i executes (respectively does not execute) an action a if and and only if i believes that he does (respectively does not do) a. $\mathcal{IAL}$-theorems 4.2j-4.2m are about the relations between belief, historic necessity and the temporal modality *next*.

$\mathcal{IAL}$-theorem 4.2n expresses a sort of generative principle for intentions according to which, if i wants φ to be true and believes that φ will be true only if he does a then i comes to intend to do a. The belief $\mathbf{Bel}_i(\varphi \rightarrow \langle i{:}a\rangle \top)$ should be called agent i's *reason for intending* to do action a [3, 17, 53] (see also Section 6.2.1).

It has to be noted that neither $\mathbf{Does}_i\varphi \rightarrow \mathbf{Bel}_i\mathbf{Does}_i\varphi$ nor $\neg\mathbf{Does}_i\varphi \rightarrow \mathbf{Bel}_i\neg\mathbf{Does}_i\varphi$ are $\mathcal{IAL}$ valid. This highlights that an agent is not necessarily aware of what he will bring about.

5 Comparison with STIT and Coalition Logic

5.1 Main differences between STIT and $\mathcal{IAL}$

There are some substantial differences between $\mathcal{IAL}$ and STIT theories. Formulas in STIT logic are built by means of the Boolean connectives together with the modal operator $\Box$ of historic necessity and the so-called *Chellas* STIT operator $[i\ cstit:\ ]$. The modal construction $\Box\varphi$ is read 'φ is true in all possible histories', whereas $[i\ cstit:\varphi]$ is read 'agent i sees to it that φ'. Thus, $\Diamond[i\ cstit:\varphi]$ and $\Box[i\ cstit:\varphi]$ can be read 'agent i can see to it that φ' and 'agent i necessarily sees to it that φ'.

STIT formulas such as $[i\ cstit:\varphi]$ being interpreted by means of reflexive relations, $[i\ cstit:\varphi]\rightarrow\varphi$ is valid in STIT. This implies that in STIT theory, actions are supposed to represent *ex post acto* action sentences, or finished actions. In $\mathcal{IAL}$, the construct $\mathbf{Does}_C\varphi$ indicates that a coalition of agents is about to take an action that brings about φ, and that action has its results at the next moment. For these reasons, for every $C \subseteq AGT$ the relation D_C is simply a serial relation and for every $C \subseteq AGT$ the formula $\mathbf{Does}_C\varphi \wedge \neg\varphi$ is satisfiable.[9]

5.2 Relationship between Coalition Logic and $\mathcal{IAL}$

Pauly's Coalition Logic (CL) [39] is a popular logic for multi-agent systems that stems from social choice theory. CL has been introduced to reason about what single agents and groups of agents are able to achieve. CL has coalition modalities of the form $[C]$ where C is an arbitrary coalition of agents $C \subseteq AGT$ (where AGT is the set of all agents). The CL formula $[C]\varphi$ is read 'the coalition C can bring about (can enforce an outcome state satisfying) φ'. Space restrictions prevent presenting in the detail the mathematical relationship between CL and $\mathcal{IAL}$. However, we have explored them in [33], in which it is proved that a slightly different variant of $\mathcal{IAL}$ without modal operators for beliefs and goals embeds CL. In particular, the CL formula $[C]\varphi$ can be translated into $\mathcal{IAL}$ by the formula $\Diamond\mathbf{Does}_C\varphi$.

One can observe that logics for agency and multi-agent systems have three dimensions: historic necessity/possibility, agent's choice and time. In CL and ATL, these three components are *fused* and make up a single non-normal modal operator. We have seen that in STIT logic, these three

[9] Note that in STIT theory $\Box\varphi \rightarrow [i\ cstit:\varphi]$ is also valid. This is not the case in $\mathcal{IAL}$ where for every $C \subseteq AGT$ $\Box\varphi \wedge \neg\mathbf{Does}_C\varphi$ is satisfiable.

ingredients are separated, and each has its own modal operator. In $\mathcal{IAL}$, we explore the middle ground: we fuse the choice and the temporal *next* operator. A similar construction is used in [8].

6 Varieties of power

This last part of the article provides a comprehensive formal ontology of power in the logic $\mathcal{IAL}$. We start with an analysis of the general concept of *power of* (Section 6.1). Then, in Section 6.2, we study social power by distinguishing the three general concepts of *influencing power*, *persuasive power* and *dependence-based social power*.

6.1 Power of

The aim of this section is to provide a formal characterization of the concept of *power of* by exploiting the expressiveness of $\mathcal{IAL}$. As argued in [4, 12], for an agent i to have the power of achieving φ, i must have the objective capability to achieve φ and must be aware of this.[10] In fact, without i's discretion over his objective capability i would be unable to exploit it in order to ensure φ. A first rough pre-formal definition of i's *power of* achieving φ is given by the following two conditions:

(1) i can bring about that φ (*objective capability*); and
(2) i believes that he can bring about that φ (*discretion over the capability*).

In $\mathcal{IAL}$, the former condition is expressed by the formula $\Diamond\mathbf{Does}_i\varphi$, while the latter condition is expressed by $\mathbf{Bel}_i\Diamond\mathbf{Does}_i\varphi$.

Let us we denote by $\mathbf{K}_i\varphi$ agent i's correct belief that φ holds.

Definition 6.1
For every $i \in AGT$:

$$\mathbf{K}_i\varphi \stackrel{\text{def}}{=} \mathbf{Bel}_i\varphi \wedge \varphi$$

The modal operator $\mathbf{K}_i$ is normal, and obeys the principles of the logic **S4**. That is, if an agent i has a correct belief that φ then φ is true ($\vdash_{\mathcal{IAL}} \mathbf{K}_i\varphi \rightarrow \varphi$), and an agent i has positive introspection over his correct beliefs ($\vdash_{\mathcal{IAL}} \mathbf{K}_i\varphi \rightarrow \mathbf{K}_i\mathbf{K}_i\varphi$). Moreover, $\mathbf{K}_i$ satisfies Axiom K and necessitation. Thus, one might try to formalize the concept i's *power of* achieving φ by the formula $\mathbf{K}_i\Diamond\mathbf{Does}_i\varphi$. But this is not sufficient to formalize a genuine concept of power. In fact, the formula $\mathbf{K}_i\Diamond\mathbf{Does}_i\varphi$ simply says 'i correctly believes that there exists some action whose execution can ensure φ'. It does not say 'there is some action such that if agent i chooses it, i correctly believes that he will ensure φ by doing that action'.[11] To see why $\mathbf{K}_i\Diamond\mathbf{Does}_i\varphi$ is insufficient to capture the concept of power, consider the scenario in Figure 2. Agent i is at world w_1 and is in front of two doors A and B. Behind door A 'there is a treasure' (proposition t), behind door B there is nothing. Besides, i believes that behind one of the two doors there is a treasure, whereas behind the other there is nothing, but he is not sure whether the treasure is behind door A or B. The agent can either open door A (action a) or open door B (action b). In the world w_1 and in each world which is compatible with i's beliefs at w_1 (worlds w_3

A similar argument is given in [52] where the notion of *practical possibility* is distinguished from the notion of *power*.

[11] The necessity to distinguish *de dicto* sentences of the form 'i knows that there exists some action by doing which he can ensure φ' from *de re* sentences of the form 'there is some action such that if agent i chooses it, i correctly believes that he will ensure φ by doing that action' has also been stressed in [11, 31, 46].

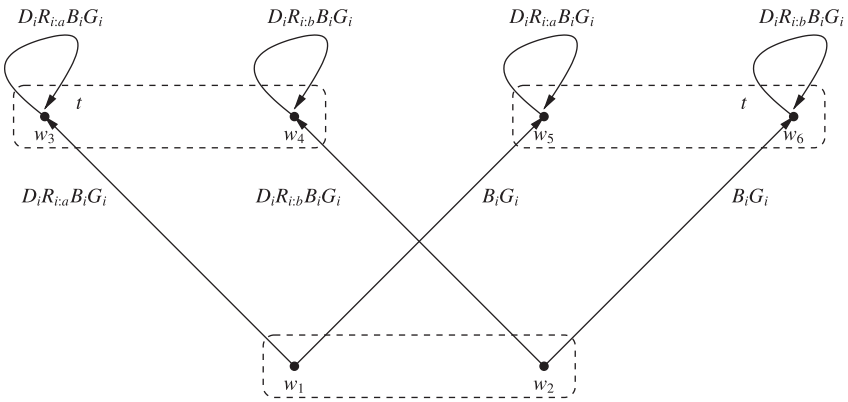

FIG. 2. Example of $\mathcal{IAL}$-model. The equivalence relation H is depicted by the dashed lines and form clusters of historic alternatives.

and w_5), it is the case that he can get the treasure. From this, we conclude that at w_1 agent i correctly believes that he can get the treasure: $\mathbf{K}_i\Diamond\mathbf{Does}_i t$ holds at w_1. Unfortunately, there is no action such that if agent i chooses it, he correctly believes that he will get the treasure. So, it is reasonable to say that in the example, i does not have the power of getting the treasure. At world w_1 agent i cannot correctly believe that he will get the treasure by opening door A, nor can he correctly believe that he will get the treasure by opening door B: $\neg\Diamond\mathbf{K}_i(\langle i{:}a\rangle\top\wedge\mathbf{Does}_i t)$ and $\neg\Diamond\mathbf{K}_i(\langle i{:}b\rangle\top\wedge\mathbf{Does}_i t)$ hold at w_1. More generally, at w_1 there is no action such that if agent i chooses it, he correctly believes that he will get the treasure: $\neg\Diamond\mathbf{K}_i\mathbf{Does}_i t$ holds at w_1.

From the previous example, we have to conclude that an agent i does not have the *power of* achieving φ unless:

(*) there is some action such that if agent i chooses it, he correctly believes that he will ensure the property φ.

Statement (*) corresponds to a notion of *power of* that can be formalized in $\mathcal{IAL}$.

DEFINITION 6.2
For every $i\in AGT$:

$$\mathsf{PowerOf}(i,\varphi)\stackrel{\text{def}}{=}\Diamond\mathbf{K}_i\mathbf{Does}_i\varphi$$

By Axiom $\mathbf{Confl}_{Bel,\Box}$, we can show that $\mathsf{PowerOf}(i,\varphi)$ implies $\mathbf{K}_i\Diamond\mathbf{Does}_i\varphi$ (but not *vice versa*) which, as discussed above, characterizes a situation of uncertainty in which i cannot determine what action must be taken to ensure φ. The following $\mathcal{IAL}$-theorem highlights another noteworthy property of the previous notion of power of. For every $i\in AGT$:

$$\vdash_{\mathcal{IAL}}\mathsf{PowerOf}(i,\varphi)\leftrightarrow\mathbf{K}_i\mathsf{PowerOf}(i,\varphi)\tag{6.1}$$

PROOF. We only prove left-to-right direction of $\mathcal{IAL}$-theorem 6.1. The other direction is trivially satisfied by definition of $\mathbf{K}_i\varphi$. By definition of $\mathsf{PowerOf}(i,\delta_i,\varphi)$ and $\mathbf{K}_i\varphi$, we have that $\mathsf{PowerOf}(i,\delta_i,\varphi)$ implies $\Diamond\mathbf{K}_i\mathbf{Does}_i\varphi$. Moreover, $\Diamond\mathbf{K}_i\varphi\rightarrow\mathbf{K}_i\Diamond\varphi$ and $\mathbf{K}_i\varphi\rightarrow\mathbf{K}_i\mathbf{K}_i\varphi$ are theorems of $\mathcal{IAL}$ (by definition of $\mathbf{K}_i\varphi$, Axiom 4 for $\mathbf{Bel}_i$ and Axiom $\mathbf{Confl}_{Bel,\Box}$). Therefore, $\Diamond\mathbf{K}_i\mathbf{Does}_i\varphi$ implies $\Diamond\mathbf{K}_i\mathbf{K}_i\mathbf{Does}_i\varphi$

which in turn implies $\mathbf{K}_i \Diamond \mathbf{K}_i \mathbf{Does}_i \varphi$. From this and the definition of $\mathsf{PowerOf}(i, \varphi)$, we can infer $\mathbf{K}_i \mathsf{PowerOf}(i, \varphi)$. ∎

According to the $\mathcal{IAL}$-theorem 6.1, an agent has the power of achieving φ if and only if he correctly believes this.

6.2 Social power

An interesting form of power on which many authors in sociology have focused is the intrinsically social power called *power over*. However, there is no consensus on the meaning of the expression 'an agent has power over another agent with respect a given issue, fact, etc'. Several kinds of social power have been investigated and defined.

While power, in its most general sense, can be seen as an agent's capability of producing causal effects and the agent's awareness of this capability (Section 6.1), social power is an agent's causal power to affect the conduct of other agents. Therefore, the most important aspect of social power is that it is a bipartite relation between two agents, one of whom is the principal agent, and the other the subordinate agent.

A point of disagreement is whether *i*'s *power over j* should be based on *i*'s ability to affect the behaviour of *j* by inducing *j* to intend to do a certain action or to refrain from doing a certain action (*power of influencing*) or whether it should be based on *j*'s dependence on *i* for the achievement of his goals (*dependence-based social power*). The former concept of power is studied for instance in [16], whereas the latter is investigated in [12, 23]. The authors in [12, 23] emphasize that the two notions of power are closely interdependent. In fact, if *i* has a dependence-based power over *j* and he knows this, then he is in a position to make threats or offers to *j* in order to affect his behaviour thereby having a power of influencing *j*.

The aim of this section is to formalize in our logic $\mathcal{IAL}$ the most important kinds of social power and to provide an analysis of their logical relationships.

6.2.1 Influencing power and persuasive power

The first kind of social power we consider is *influencing power*. We say that agent *i* has the power of influencing agent *j* when *i* is in a position to induce *j* to intend to do certain action or in a position to induce *j* to refrain from doing a certain action. In the former case, *i* has the power of shaping *j*'s preferences in such a way that *j* will intend to do a certain action, in the latter case *i* has the power of shaping *j*'s preferences in such a way that *j* will intend not to do a certain action. More succinctly, we say that *i* has the *power of influencing j* to do (respectively not to do) action *a*, denoted by $\mathsf{InflPower}(i,j,a)$ (respectively $\mathsf{InflPower}(i,j,\sim a)$), if and only if *i* has the power of ensuring that *j* will intend to do (respectively will intend not to do) action *a*. Formally:

Definition 6.3
For every $i,j \in AGT$, $a \in ACT$:

$$\mathsf{InflPower}(i,j,a) \stackrel{\text{def}}{=} \mathsf{PowerOf}(i, \mathbf{Goal}_j \langle j{:}a \rangle \top)$$

Definition 6.4
For every $i,j \in AGT$, $a \in ACT$:

$$\mathsf{InflPower}(i,j,\sim a) \stackrel{\text{def}}{=} \mathsf{PowerOf}(i, \mathbf{Goal}_j [j{:}a] \bot)$$

We here distinguish influencing power from *persuasive power* (see [42] for an account of this concept in sociological theory). We say that i has the *power of persuading* j to believe φ, noted $\mathsf{PersPower}(i,j,\varphi)$, if and only if i has the power of ensuring that j will believe φ.

DEFINITION 6.5
For every $i,j \in AGT$:

$$\mathsf{PersPower}(i,j,\varphi) \stackrel{\text{def}}{=} \mathsf{PowerOf}(i, \mathbf{Bel}_j\varphi)$$

As emphasized in Section 1 and 4, certain beliefs can provide reasons for intending to perform a certain action. For instance, suppose an agent has a certain goal φ and believes that he will not achieve φ unless he performs action a. Then, since the agent follows the general principle of instrumental reasoning expressed by the $\mathcal{IAL}$-theorem 4.2n in Proposition 4.3 (Section 4), he will intend to perform action a in order to achieve his goal that φ.[12] The belief in the premises of this piece of practical reasoning provides a *reason for intending* to do action a. Thus, as the following $\mathcal{IAL}$-theorem highlights, if agent i correctly believes that necessarily agent j will have the goal φ and agent i has the persuasive power of giving to j the reason for intending to do action a in order to achieve φ then, indirectly, i has the power of influencing j to do a. In particular, if i has the power of persuading j that he will not achieve φ unless he does a and i correctly believes that, necessarily, in the next state j will have the goal that φ then, i has the power of influencing j to do action a. For every $i,j \in AGT$, $a \in ACT$:

$$\vdash_{\mathcal{IAL}} (\mathbf{K}_i\Box\mathbf{XGoal}_j\varphi \wedge \mathsf{PersPower}(i,j,\varphi \rightarrow \langle j{:}a\rangle\top)) \rightarrow \mathsf{InflPower}(i,j,a) \tag{6.2}$$

PROOF. $\mathsf{PersPower}(i,j,\varphi \rightarrow \langle j{:}a\rangle\top)$ is equivalent to $\Diamond\mathbf{K}_i\mathbf{Does}_{\mathrm{i}}\mathbf{Bel}_j(\varphi \rightarrow \langle j{:}a\rangle\top)$.
$\mathbf{K}_i\Box\mathbf{XGoal}_j\varphi$ implies $\mathbf{K}_i\Box\Box\mathbf{XGoal}_j\varphi$ (by Axiom 4 for $\Box$, Axiom K and necessitation for $\mathbf{K}_i$) which in turn implies $\Box\mathbf{K}_i\Box\mathbf{XGoal}_j\varphi$ (by $\mathcal{IAL}$-theorem 4.2j). Thus, $\mathsf{PersPower}(i,j,\varphi \rightarrow \langle j{:}a\rangle\top)$ and $\mathbf{K}_i\Box\mathbf{XGoal}_j\varphi$ together imply $\Diamond\mathbf{K}_i\mathbf{Does}_{\mathrm{i}}\mathbf{Bel}_j(\varphi \rightarrow \langle j{:}a\rangle\top) \wedge \Box\mathbf{K}_i\Box\mathbf{XGoal}_j\varphi$. The latter implies $\Diamond\mathbf{K}_i(\mathbf{Does}_{\mathrm{i}}\mathbf{Bel}_j(\varphi \rightarrow \langle j{:}a\rangle\top) \wedge \Box\mathbf{XGoal}_j\varphi)$ (by the $\mathcal{IAL}$-theorem $(\Box\varphi \wedge \Diamond\psi) \rightarrow \Diamond(\varphi \wedge \psi)$) which in turn implies $\Diamond\mathbf{K}_i(\mathbf{Does}_{\mathrm{i}}\mathbf{Bel}_j(\varphi \rightarrow \langle j{:}a\rangle\top) \wedge \mathbf{Does}_{\mathrm{i}}\mathbf{Goal}_j\varphi)$ (by $\mathcal{IAL}$-theorem 4.2e, Axiom T for $\Box$, Axiom K and necessitation for $\mathbf{K}_i$, Axiom K and necessitation for $\Box$). The latter implies $\Diamond\mathbf{K}_i\mathbf{Does}_{\mathrm{i}}(\mathbf{Bel}_j(\varphi \rightarrow \langle j{:}a\rangle\top) \wedge \mathbf{Goal}_j\varphi)$ which in turn implies $\Diamond\mathbf{K}_i\mathbf{Does}_{\mathrm{i}}\mathbf{Goal}_j\langle j{:}a\rangle\top$ (by $\mathcal{IAL}$-theorem 4.2n, Axiom K and necessitation for $\mathbf{K}_i$, Axiom K and necessitation for $\Box$, Axiom K and necessitation for $\mathbf{Does}_{\mathrm{i}}$). The latter is equivalent to $\mathsf{InflPower}(i,j,a)$. ∎

As in [40], we also distinguish influencing power from *indirect power*. In our view, agent i has the *indirect power* of achieving φ by means of agent j, noted $\mathsf{IndPower}(i,j,\varphi)$, if and only if i has the power of ensuring that j will bring about that φ.

DEFINITION 6.6
For every $i,j \in AGT$, $a \in ACT$:

$$\mathsf{IndPower}(i,j,\varphi) \stackrel{\text{def}}{=} \mathsf{PowerOf}(i, \mathbf{Does}_{\mathrm{j}}\varphi)$$

[12] Instrumental reasoning (generally opposed to theoretical reasoning), is the kind of reasoning that concludes in an action or in an intention [3].

6.2.2 Dependence-based social power

Social dependence have been extensively studied in the multi-agent system (MAS) domain as a fundamental concept for understanding social relations and their dynamics (see, e.g. [43]). As emphasized in [12], social power is often based on social dependence: agent i has a power over agent j because j can achieve his goals only by the aid of i (in this sense j depends on i). Our aim here is to formalize the concept of social dependence and to study its logical relationships with influencing power and persuasive power defined above.

We say that an agent j depends on agent i with respect to the achievement of φ (or i has a dependence-based power over j with respect to φ), noted $\mathsf{Dep}(j,i,\varphi)$, if and only if agent j wants φ to be true and, necessarily, the coalition $AGT \setminus \{i\}$ cannot bring about that φ no matter what agent i does. In other words, j depends on i with respect to φ if and only if j wants φ to be true and the intervention of i is necessary to ensure that φ will be true.

Definition 6.7
For every $i,j \in AGT$:

$$\mathsf{Dep}(j,i,\varphi) \stackrel{\text{def}}{=} \mathbf{Goal}_j \mathbf{X}\varphi \wedge \Box\neg \mathbf{Does}_{AGT\setminus\{i\}}\varphi$$

Remark 6.8
Note that the clause $\Box\neg\mathbf{Does}_{AGT\setminus\{i\}}\varphi$ in the definition of social dependence corresponds to the game-theoretic concept of β-ability or $\forall\exists$-capability (see, e.g. [38, 51]). Intuitively, a coalition C is said to have β-ability for φ if and only if, for every joint action (or collective choice) $\delta_{AGT\setminus C}$ of the agents in $AGT \setminus C$, there exists a possible joint action (or collective choice) δ'_C of the agents in C such that, necessarily if C does δ'_C and $AGT \setminus C$ does $\delta_{AGT\setminus C}$, then φ will be true. Thus, the formula $\Box\neg\mathbf{Does}_{AGT\setminus\{i\}}\varphi$ just expresses that agent i has the β-ability for $\neg\varphi$.

A 4-arguments definition of dependence can also be given in which the action of agent i on which agent j depends is specified. We say that an agent j depends on the execution of action a by agent i with respect to the achievement of φ, noted $\mathsf{Dep}(j,i,a,\varphi)$, if and only if agent j wants φ to be true and, necessarily, φ will be true in the next state only if i performs action a. In other words, j depends on the execution of action a by agent i with respect to φ if and only if j wants φ to be true and the occurrence of action a performed by i is necessary to ensure φ.

Definition 6.9
For every $i,j \in AGT$, $a \in ACT$: $\mathsf{Dep}(j,i,a,\varphi) \stackrel{\text{def}}{=} \mathbf{Goal}_j \mathbf{X}\varphi \wedge \Box(\mathbf{X}\varphi \rightarrow \langle i{:}a\rangle \top)$

As the following $\mathcal{IAL}$-theorem highlights, the 4-argument definition of social dependence is stronger than the 3-argument definition: if agent j depends on agent i's action a for the achievement of φ then j depends on i for the achievement of φ. For every $i,j \in AGT$, $a \in ACT$:

$$\vdash_{\mathcal{IAL}} \mathsf{Dep}(j,i,a,\varphi) \rightarrow \mathsf{Dep}(j,i,\varphi) \tag{6.3}$$

Social dependence on an agent's action has a symmetric concept of social dependence on an agent's inaction. We say that an agent j depends on agent i's inexecution of action a with respect to φ, noted $\mathsf{Dep}(j,i,\sim a,\varphi)$, if and only if agent j wants φ to be true and, necessarily, φ will be true in the next state only if i does not perform action a.

Definition 6.10
For every $i,j \in AGT$, $a \in ACT$:

$$\mathsf{Dep}(j,i,\sim a,\varphi) \stackrel{\text{def}}{=} \mathbf{Goal}_j \mathbf{X}\varphi \wedge \Box(\mathbf{X}\varphi \rightarrow [i{:}a]\bot)$$

As for social dependence on action, the 4-argument definition of social dependence on inaction is stronger than the 3-argument definition of social dependence: if agent j depends on agent i's inexecution of action a for the achievement of φ then j depends on i for the achievement of φ. For every $i,j \in AGT$, $a \in ACT$:

$$\vdash_{\mathcal{IAL}} \mathsf{Dep}(j,i,\sim a,\varphi) \rightarrow \mathsf{Dep}(j,i,\varphi) \tag{6.4}$$

We conclude with two $\mathcal{IAL}$-theorems about the logical relationship between dependence-based social power, influencing power and persuasive power. For every $i,j \in AGT$ and $a,b \in ACT$:

$$\vdash_{\mathcal{IAL}} (\mathsf{PersPower}(i,j,[j{:}b]\bot \rightarrow \mathbf{X}[i{:}a]\bot) \wedge \mathbf{K}_i\mathbf{Bel}_j \Box \mathbf{XX}\ \mathsf{Dep}(j,i,a,\varphi)) \rightarrow \mathsf{InflPower}(i,j,b) \tag{6.5}$$

$$\vdash_{\mathcal{IAL}} (\mathsf{PersPower}(i,j,\langle j{:}b\rangle \top \rightarrow \mathbf{X}\langle i{:}a\rangle \top) \wedge \mathbf{K}_i\mathbf{Bel}_j \Box \mathbf{XX}\ \mathsf{Dep}(j,i,\sim a,\varphi)) \rightarrow \mathsf{InflPower}(i,j,\sim b) \tag{6.6}$$

According to the $\mathcal{IAL}$-theorem 6.5, if i has the power of persuading j that if j does not do b then i will not do a and, i correctly believes that j believes that necessarily in two steps from now he will depend on i's action a for the achievement of φ then, i has the power of influencing j to do b. $\mathcal{IAL}$-theorem 6.6 is the corresponding version for dependence on an agent's inaction. Consider the example we gave in Section 1 of this article in which i and j are two countries in a conflict situation. Suppose i has the power of persuading j that if j makes an embargo against i then i will move a military attack against j: $\mathsf{PersPower}(i,j,\langle j{:}embargo\rangle \top \rightarrow \mathbf{X}\langle i{:}attack\rangle \top)$. Moreover, i correctly believes that j believes that, necessarily, the achievement of his goal of avoiding a war against i depends on the fact that i will not attack: $\mathbf{K}_i\mathbf{Bel}_j \Box \mathbf{XX}\mathsf{Dep}(j,i,\sim attack,\neg war)$. By the $\mathcal{IAL}$-theorem 6.6, we conclude that i has the power of influencing j not to make an embargo against him: $\mathsf{InflPower}(i,j,\sim embargo)$.

7 Conclusion

In this article, we have developed a logical framework that allows to formalize different forms of power and to clarify their relationships with the concept of action and with intentional concepts like the concepts of belief and goal. There are important forms of power that we did not consider here and that are crucial for a theory of organization. For instance, we did not consider the concept of *institutional power* that an agent has *qua* player of a certain *role* within the context of an organization and which is specified by means of rules of the form 'an act a performed by an agent i playing a certain role r *counts as* i's act of ensuring φ' (e.g. a priest's act of performing certain gestures during a wedding ceremony counts as the priest's act of marrying a couple). See, e.g. [32] for a logical account of institutional power. In other words, in this work we have been mainly interested in the logical analysis of the cognitive constituents of power, without relating this concept to the theory of organization. Thus, our logical approach is somehow complementary to the logical approach proposed by Dignum & Dignum [19] in which the cognitive aspect of agency and of social interaction is not considered, but which is interested in clarifying the relationships between the concept of action and the organizational concepts of role (e.g. role dependency, role hierarchy), responsibility and delegation.

There are several ways in which the work presented in this article can be extended. An interesting direction of application of the logic $\mathcal{IAL}$ is the theory of collective powers [12]. $\mathcal{IAL}$'s constructions for groups of agents of the form $\mathbf{Does}_C\varphi$ can be useful for understanding how powers of coalitions interact with powers and mental attitudes of individuals. We have argued that, for an agent i to have the power of achieving φ, i must have both the objective capability to ensure φ and must be aware of his capability. The same argument applies to collective powers. Indeed, it seems reasonable to suppose that, for a group of agents C to have the power of achieving φ, the agents in C must be able to perform a joint action that will ensure φ and must be collectively aware of this, where being collectively aware of something seems to require some group attitude notions such as common belief.

We also think that $\mathcal{IAL}$ extended with modal operators for common belief of the form $\mathbf{CBel}_C$ (see, e.g. [21] for an analysis of these modalities) is a suitable framework for formalizing the concept of joint intention as defined in [7]. In Bratman's analysis, there are three basic conditions in the definition of joint intention. We can approximately say that a group of agents C has the joint intention that φ if and only if every agent in C intends that the group C ensures φ (*joint goal* condition),[13] every agent in C intends that the group C performs the joint action δ_C in order to ensure φ (*joint plan* condition), and these two facts are common belief between the agents in C (*common ground* condition). Thus, Bratman's concept of joint intention can be translated into $\mathcal{IAL}$ extended with common belief by the following formula: $\bigwedge_{i\in C}\mathbf{Goal}_i(\langle\delta_C\rangle\top\wedge\mathbf{Does}_C\varphi)\wedge\mathbf{CBel}_C(\bigwedge_{i\in C}\mathbf{Goal}_i(\langle\delta_C\rangle\top\wedge\mathbf{Does}_C\varphi))$.

Another interesting avenue for future research is to enrich the ontology of action and time of $\mathcal{IAL}$. Indeed, at the current stage $\mathcal{IAL}$ only allows to reason about next states and single-step actions, and is therefore too weak to account for strategies in the sense of ATL [2] and strategic STIT [9, 10, 28]. A way to overcome this limitation is to enrich the dynamic logic fragment of our two logics by introducing additional PDL constructs such as action composition (;), choice ($\cup$) and iteration (*).

Acknowledgements

We first would like to thank the participants of the workshop FAMAS'007 in Durham, as well as the two reviewers for this long version.

Funding

E. Lorini and A. Herzig are partially supported by French ANR project ForTrust. N. Troquard is partially supported by the EPSRC grant (EP/E061397/1) *Logic for Automated Mechanism Design and Analysis*.

References

[1] T. Ågotnes. Action and knowledge in alternating-time temporal logic. *Synthese*, **149**, 377–409, 2006.

[2] R. Alur, T. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, **49**, 672–713, 2002.

[3] R. Audi. A theory of practical reasoning. *American Philosophical Quarterly*, **19**, 25–39, 1982.

[4] B. Barnes. *The Nature of Power*. Polity Press, 1988.

[13] This condition, according to Bratman, defines the concept of *we*-intention in the sense that 'I (and you) intend that *we* do something together'.

[5] N. Belnap, M. Perloff, and M. Xu. *Facing the Future: Agents and Choices in our Indeterminist World*. Oxford University Press, 2001.
[6] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.
[7] M. Bratman. Shared cooperative activity. *The Philosophical Review*, **101**, 327–341, 1992.
[8] J. M. Broersen. A logical analysis of the interaction between 'obligation-to-do' and 'knowingly doing'. In *Proceedings of the Ninth International Conference on Deontic Logic in Computer Science (DEON'08)*, Vol. 5076 of *Lecture Notes in Computer Science*, pp. 140–154. Springer, 2008.
[9] J. M. Broersen. A stit-logic for extensive form group strategies. In *WI-IAT '09: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, J. Lang, S. Mitra, and S. Parsons, eds, pp. 484–487. IEEE Computer Society, 2009.
[10] J. M. Broersen, A. Herzig, and N. Troquard. A STIT-extension of ATL. In *Proceedings Tenth European Conference on Logics in Artificial Intelligence (JELIA'06)*, M. Fisher, ed., Vol. 4160 of *Lecture Notes in Artificial Intelligence*, pp. 69–81. Springer, 2006.
[11] J. M. Broersen, A. Herzig, and N. Troquard. Normal coalition logic and its conformant extension. In *Proceedings of Eleventh Conference on Theoretical Aspects of Rationality and Knowledge (TARK XI)*, ACM, New York, 2007.
[12] C. Castelfranchi. The micro-macro constitution of power. *Protosociology*, 18–19, 2003.
[13] B. F. Chellas. *Modal Logic: an Introduction*. Cambridge University Press, 1980.
[14] P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, **42**, 213–261, 1990.
[15] W. Conradie, V. Goranko, and D. Vakarelov. Algorithmic correspondence and completeness in modal logic: the core algorithm SQEMA. *Logical methods in computer science*, **2**, 1–26, 2006.
[16] R. Dahl. The concept of power. *Behavioral Science*, **2**, 201–215, 1957.
[17] D. Davidson. Actions, reasons and causes. *Journal of Philosophy*, **60**, 685–700, 1963.
[18] D. Davidson. The logical form of action sentences. In *The Logic of Decision and Action*. N. Rescher, ed., University of Pittsburgh Press, 1967.
[19] V. Dignum and F. Dignum. A Logic for Agent Organizations. In *Proceedings of the Workshop on Formal Approaches to Multi-Agent Systems Formal Approaches to Multi-Agent Systems (FAMAS 2007)*, Durham University Press, 2007.
[20] B. Dunin-Keplicz and R. Verbrugge. Collective intentions. *Fundamenta Informaticae*, **51**, 271–295, 2002.
[21] R. Fagin, J. Halpern, Y. Moses, and M. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.
[22] R. Fagin and J. Y. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, **34**, 39–76, 1987.
[23] A. I. Goldman. Toward a theory of social power. *Philosophical studies*, **23**, 221–268, 1972.
[24] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. MIT Press, 2000.
[25] A. Herzig and E. Lorini. A dynamic logic of agency I: STIT, capabilities and powers. *Journal of Logic, Language and Information*, **19**, 89–121, 2009.
[26] A. Herzig and N. Troquard. Knowing how to play: Uniform choices in logics of agency. In *Proceedings of AAMAS'06*, H. Nakashima, M. P. Wellman, G. Weiss, and P. Stone, eds, pp. 209–216. ACM Press, 2006.
[27] J. Hintikka. *Knowledge and Belief*. Cornell University Press, 1962.
[28] J. F. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.

[29] W. Jamroga. Some remarks on alternating temporal epistemic logic. In *Proceedings of the International Workshop on Formal Approaches to Multi-Agent Systems (FAMAS'03)*, 2003.
[30] W. Jamroga and T. Ågotnes. Constructive knowledge: what agents can achieve under imperfect information. *Journal of Applied Non-Classical Logics*, **17**, 423–475, 2007.
[31] W. J. Jamroga and W. van der Hoek. Agents that know how to play. *Fundamenta Informaticae*, **63**, 185–219, 2004.
[32] A. Jones and M. J. Sergot. A formal characterization institutionalised power. *Journal of the IGPL*, **4**, 429–445, 1996.
[33] E. Lorini. A dynamic logic of agency II: deterministic DLA, coalition logic, and game theory. *Journal of Logic, Language and Information*, **19**, 327–351, 2010.
[34] E. Lorini and A. Herzig. A logic of intention and attempt. *Synthese*, **163**, 45–77, 2008.
[35] E. Lorini and F. Schwarzentruber. A logic for reasoning about counterfactual emotions. In *Proceedings of the twenty-first International Joint Conference on Artificial Intelligence (IJCAIÂ'09)*, pp. 867–872. AAAI Press, 2009.
[36] E. Lorini, F. Schwarzentruber, and A. Herzig. Epistemic Games in Modal Logic: Joint Actions, Knowledge and Preferences all together. In *LORI-II Workshop on Logic, Rationality and Interaction*, X. He, J. F. Horty, and E. Pacuit, eds, pp. 212–226. Springer, 2009.
[37] E. Lorini, N. Troquard, A. Herzig, and C. Castelfranchi. Delegation and mental states. In *Proceedings of Sixth International Joint Conference on Autonomous Agents in Multi-Agent Systems (AAMAS'07)*. ACM Press, 2007.
[38] M. Pauly. *Logic for Social Software*. PhD Thesis, University of Amsterdam, 2001.
[39] M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, **12**, 149–166, 2002.
[40] I. Pörn. *Action Theory and Social Science: Some Formal Models*. Synthese Library 120, D. Reidel, 1977.
[41] A. S. Rao and M. Georgeff. BDI agents: from theory to practice. In *Proceedings of the first international conference on Multi-Agent Systems (ICMAS-95)*, pp. 312–319. AAAI Press, 1995.
[42] J. Scott. Modes of power and the re-conceptualization of elites. *Sociological Review*, **56**, 25–43, 2008.
[43] J. S. Sichman and R. Conte. Multi-agent dependence by dependence graphs. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2002)*, pp. 483–490. ACM Press, 2002.
[44] R. H. Thomason. Combinations of tense and modality. In *Handbook of Philosophical Logic: Extensions of Classical Logic*, D. Gabbay and F. Guenthner, eds, pp. 135–165. Reidel, 1984.
[45] N. Troquard and L. Vieu. Towards a logic of agency and actions with duration . In *European Conference on Artificial Intelligence 2006 (ECAI'06) , Riva del Garda, Italy*, pp. 775–776. IOS Press, 2006.
[46] J. van Benthem. Games in dynamic-epistemic logic. *Bulletin of Economic Research*, **53**, 219–248, 2001.
[47] J. van Benthem and E. Pacuit. The tree of knowledge in action: Towards a common perspective. In *Proceedings of Advances in Modal Logic Volume 6 (AiML 2006)*, G. Governatori, I. Hodkinson, and Y. Venema, eds, pp. 87–106. College Publications, 2006.
[48] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proceedings of Fourth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'05)*, ACM Press, 2005.

[49] W. van der Hoek and M. Wooldridge. Cooperation, knowledge, and time: Alternating-time temporal epistemic logic and its applications. *Studia Logica*, **75**, 125–157, 2003.
[50] W. van der Hoek and M. Wooldridge. Towards a logic of rational agency. *Logic Journal of the IGPL*, **11**, 133–157, 2003.
[51] W. van der Hoek and M. Wooldridge. On the logic of cooperation and propositional control. *Artificial Intelligence*, **64**, 81–119, 2005.
[52] B. van Linder, W. van der Hoek, and J.-J. Ch. W., Meyer. Formalising abilities and opportunities. *Fundamenta Informaticae*, **34**, 53–101, 1998.
[53] G. H. Von Wright. On so-called practical inference. *The Philosophical Review*, **15**, 39–53, 1972.
[54] M. Wooldridge. *Reasoning about rational agents*. MIT Press, 2000.