A dynamic logic of institutional actions

Andreas Herzig[1], Emiliano Lorini[1], and Nicolas Troquard[2]

[1] University of Toulouse, CNRS, IRIT, France
[2] Laboratory of Applied Ontology, ISTC-CNR, Italy

Abstract. We propose a logical framework to represent and reason about some important aspects of a theory of institutional action: (1) the distinctions between physical facts and actions and institutional facts and actions; (2) the distinction between causality and 'counts-as'; (3) the notion of institutional power. Technically, our contribution consists in extending a dynamic logic of propositional assignments with constructions allowing to express that an agent plays a given role; that a physical action causes another physical action; that a physical action performed by an agent playing a given role counts as an institutional action.

1 Introduction

We present a logical framework in which we can represent and reason about some important aspects of a theory of institutional action: (1) the distinctions between physical facts and actions and institutional facts and actions; (2) the distinction between causality and 'counts-as'; (3) the notion of institutional power. Our framework is that of a dynamic logic of propositional assignments in the sense of [32, 33, 6, 2]. In preceding work [14] we have shown that this logic allows to reason about agent capabilities in the sense of coalition logic [26] and coalition logic of propositional control [16, 15]. We here refine our account by distinguishing 'brute', physical facts from institutional facts. This leads us to the distinction between brute actions (changing brute facts) from institutional actions (changing institutional facts). We moreover add constructions allowing to express that an agent plays a given role; that a physical action causes another physical action (e.g. Jack's action of shooting Joe causes Jack's action of killing Joe); that a physical action performed by an agent playing a given role counts as an institutional action (e.g. an agent's act of performing certain gestures during the wedding ceremony while playing the role of priest counts as the act of marrying the couple). This provides a full-blown account of normative systems.

The paper is organized as follows. Section 2 establishes the conceptual basis of the logical analysis developed in the rest of the paper, providing a detailed discussion of the philosophical theory of institutional action developed by Goldman, Searle and other scholars, and it explains how the logic presented in the paper takes into account its different dimensions and aspects. Section 3 presents the syntax and the semantics of the logic, while Section 4 provides a complete axiomatization as well as a complexity result for the satisfiability problem. In Section 5 the logic is exploited to formalize the concept of institutional power. In Section 6 we discuss related works in the area of logic of normative systems. We finally conclude in Section 7 by discussing some perspectives for future work.

2 Institutional actions: conceptual analysis

Some background and clarifications of the notion of institutional action are needed in order to ground the logical analysis presented in the rest of the paper on a solid conceptual basis.

Physical facts and actions vs. institutional facts and actions. According to several authors working in legal theory and in the field of normative multiagent systems (MAS) (see e.g. [1, 3]), normative systems are based both on regulative as well as constitutive (i.e., non-regulative) components. That is, normative systems 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" [29, p. 33]. In Searle's theory [29, 30], constitutive rules are expressed by means of 'counts-as' assertions of the form "X counts as Y in context C" where the context C refers to the normative system in which the rule is specified. Constitutive rules relate "brute" physical facts and actions with institutional facts and actions. For example, in the context of the US federal reserve, receiving a piece of paper with a certain shape, color, etc. (a physical action) counts as receiving a certain amount of money (an institutional action); or in the context of Catholic Church the priest's action of performing certain gestures during the wedding ceremony (which is a physical action) counts as the act of marrying the couple (which is an institutional action). Although Searle's counts-as relation is between objects in general (such as a piece of paper counting as an amount of money), in this work we only consider the counts-as' relation between actions.

As pointed out by [17], the counts-as relation may also relate two institutional actions. For example, in the context of chess, the action of checkmating the opponent (an institutional action) counts as the action of winning the game (an institutional action).

Causality vs. counts-as. In his seminal work on the philosophical theory of action, Goldman studied a fundamental relation between actions and events of the form "action α is done *by* doing a different action β" [8]. The word "by" expresses a relation between actions which Goldman calls *generation*. This means that an action can be performed by way of one or more actions. According to Goldman's theory, there are actions which have a deep recursive structure. In fact, there could be an action α done by doing an action β which in turn is done by doing a further action γ and so on. Such a decomposition of an agent's action α stops at the level of basic actions. Basic actions therefore represent the agent's only direct intervention in the process of doing α. As Davidson puts it, "the rest is up to nature" [5].[3] By way of example, consider Jack's action of killing Joe. Jack kills Joe by shooting him and Jacks shoots Joe by pulling the trigger of the gun. Jack's bodily movement of pulling the trigger (which consists in Jack's moving his forefinger in a certain way) is a basic action, as it is the only part of the action of killing Joe which is directly controlled by Jack.

[3] See [19] for a formal analysis of basic actions in dynamic logic.

Goldman's theory opposes "causal generation" to "conventional generation". The latter can be identified with Searle's counts-as relation. According to Goldman, physical actions are causally generated, that is, they just consist in an agent bringing about (i.e. causing) a certain state of affairs to be true. On the other hand, institutional actions are conventionally generated, by which he meant that actions such as signaling before making a turn, and checkmating one's opponent, exist in virtue of rules or conventions relating physical actions with institutional effects. For example, in the sentence "a player wins a chess game against his opponent by checkmating him" the word "by" expresses a relation of conventional generation, that is, the action of checkmating the opponent *counts-as* the action of winning the chess game. On the contrary, in the sentence "Jack kills Joe by shooting him" the word "by" expresses a relation of causal generation, that is, Jack's action of shooting Joe *causes* the action of killing Joe (i.e. the action of making Joe dead). To carry the example further, the action of killing Joe might conventionally generate the action of murdering Joe (which is an institutional action).

We here explore Goldman's view. We assume that the causal relation and the counts-as' relation between actions are ontologically different for at least two reasons. While the former relates a physical action to another physical action, the latter relates a physical action to an institutional action, or an institutional action to another institutional action. Moreover, while the causal relation is merely a relation between physical actions performed by an agent, counts-as' is a relation between actions performed by an agent playing a certain role in a given institutional context. For example, in the institutional context of Catholic Church, an agent's act of performing certain gestures during the wedding ceremony while playing the role of priest counts as the act of marrying the couple. As the next paragraph highlights, this aspect of counts-as is fundamental to understand the notion of institutional power.

Institutional power. Some legal and social theorists [4, 29, 12] as well as some logicians [17, 20, 22] have emphasized the tight relationship between counts-as and the notion of institutional power. According to these authors, there exists a specific kind of norms called *norms of competence* whose function in a legal system is to assign institutional powers to the agents playing certain roles within a given institution.[4] Such power-conferring norms should not be reduced to norms of conduct such as obligations, prohibitions, commands and permissions. On the contrary, they are expressed by means of counts-as assertions relating physical or institutional actions to institutional actions. 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.). In other words, according to these authors, saying that "an agent playing the role r has the institutional power to do the institutional action α *by* doing action β" just means that "an agent's performance of action β while playing the role r counts as the agent's performance of the institutional action α". For example, "an agent playing the role of priest has the institutional power to marry a couple *by* performing certain

[4] From this perspective, a given role can be identified with the set of norms of competence concerning it. This view is compatible with [27] in which a role is defined as a set (or cluster) of norms.

gestures during the wedding ceremony" just means that "an agent's act of performing certain gestures during the wedding ceremony while playing the role of priest counts as the act of marrying the couple".

The interesting aspect of this notion of institutional power is that it allows to properly understand how (human or software) agents, conceived as physical entities, can produce institutional effects by way of performing physical actions and by playing certain social roles. To summarize, the crucial point is the following. A given agent i has the ability to perform a certain institutional action α by way of performing another action β because: (1) the agent plays a certain social role r; (2) there is a norm of competence establishing that the performance of action α by an agent playing the role r counts as the performance of the institutional action β. For example, an agent i has the ability of marrying a couple by performing certain gestures during a wedding ceremony because: (1) i plays the role of priest; (2) there is a norm of competence establishing that an agent's physical act of performing certain gestures during the wedding ceremony while playing the role of *priest* counts as the institutional act of marrying the couple.

Remarks on the nature of roles and brute abilities. In the framework presented in Section 3, the world the agents populate will be a mere database listing the atomic facts that are true at this very moment. The world is dynamic because the agents can act upon their environment by changing the truth value of these atomic facts. Our proposal relies on the assumption that an agent's brute abilities can be identified with the set of propositional assignments that he can perform. As shown in [14], such a framework also supports the models of *propositional control* [16, 7].

Roles are central in our study of institutional abilities. It is by occupying roles that an agent's brute action generates an institutional event. However, we do not need at this stage to ground our proposal on a rigorous ontology of roles. We refer to [23] for a foundational ontology of roles and a detailed interdisciplinary review of the literature.

The rest of the paper. We start by splitting the set of propositional variables into two disjoint sets: atomic physical facts and atomic institutional facts. In the logic presented in Section 3, a physical action just consists in setting to 'true' or to 'false' some atomic physical fact, while an institutional action consists in setting to 'true' or to 'false' some atomic institutional fact. The latter actions can only be performed indirectly, i.e. by performing a physical action. Moreover, we distinguish two different relations between actions: "counts-as" and "causes". There are conditionals of the form $\alpha_1 \Rightarrow \alpha_2$ expressing that i's performance of the physical action α_1 causes the performance of the physical action α_2, and there are conditionals of the form $\alpha_1 \overset{r}{\rightsquigarrow} \alpha_2$ expressing that i's performance of the physical or institutional action α_1 in the social role r counts as the performance of the institutional action α_2. Finally, in Section 5, we study the notion of institutional power by introducing special modal operators describing an agent's capability of producing a given institutional effect by way of performing a physical action while playing a given social role.

For the sake of simplicity we suppose that there is only one institution.

3 Logic

This section introduces the syntax and the semantics of the logic. It is basically an extension of our logic of [14] by a causality relation.

3.1 Language

We suppose that there is a finite set of *agents* $\mathbb{A}$, a finite set of *roles* $\mathbb{R}$, and a countable set of propositional variables.

Propositional variables are meant to capture the atomic facts about the world. They are partitioned into two kinds: *facts about the physical world (or brute facts)* and *facts about the institutional world (or institutional facts)* and are collected in two sets $\mathbb{P}_{\mathsf{phys}}$ and $\mathbb{P}_{\mathsf{inst}}$. These sets form a partition, and we thus assume that $\mathbb{P} = \mathbb{P}_{\mathsf{phys}} \cup \mathbb{P}_{\mathsf{inst}}$ and $\mathbb{P}_{\mathsf{phys}} \cap \mathbb{P}_{\mathsf{inst}} = \emptyset$. Our propositional variables are therefore typed.

These sets are again composed of variables of different sub-types: we suppose that they respectively contain a countable set $\mathbb{P}^0_{\mathsf{phys}} \subseteq \mathbb{P}_{\mathsf{phys}}$ of *basic physical facts*; and a countable set $\mathbb{P}^0_{\mathsf{inst}} \subseteq \mathbb{P}_{\mathsf{inst}}$ of *basic institutional facts*. Beyond these basic variables the set $\mathbb{P}_{\mathsf{inst}}$ contains variables denoting that an agent plays a role and the set $\mathbb{P}_{\mathsf{phys}}$ contains variables denoting that an agent is able to assign a variable to true or false. We suppose that the latter covers the case of ability variables being themselves assigned; we will therefore need a recursive definition.

In the sequel we are going to analyze in more detail what kinds of propositions actually populate $\mathbb{P}$.

Institutional facts. We have assumed the existence of a finite set of roles $\mathbb{R}$. These roles are occupied by agents. We write $R_i(r)$ to formalize the fact that agent $i \in \mathbb{A}$ occupies the role $r \in \mathbb{R}$. Holding a role is a societal construct, and an atomic institutional fact. It is also a contingent fact (or anti-rigid [23]) meaning that role occupations can change. Hence, we assume that expressions of the form $R_i(r)$ are propositional variables in $\mathbb{P}_{\mathsf{inst}}$. These are the only 'special' atomic institutional facts in $\mathbb{P}_{\mathsf{inst}}$ and we therefore have $\mathbb{P}_{\mathsf{inst}} = \mathbb{P}^0_{\mathsf{inst}} \cup \{R_i(r) \ : \ r \in \mathbb{R}, i \in \mathbb{A}\}$. We can write this in a BNF:

$$\mathbb{P}_{\mathsf{inst}} : p_{\mathsf{inst}} ::= p^0_{\mathsf{inst}} \mid R_i(r)$$

where p^0_{inst} ranges over the set of basic institutional facts $\mathbb{P}^0_{\mathsf{inst}}$, i ranges over the set of agents $\mathbb{A}$, and r ranges over the set of roles $\mathbb{R}$. Then the language $\mathcal{L}_{\mathsf{inst}}$ of (complex) institutional facts is defined by the following grammar:

$$\mathcal{L}_{\mathsf{inst}} : \varphi_{\mathsf{inst}} ::= p_{\mathsf{inst}} \mid \neg\varphi_{\mathsf{inst}} \mid \varphi_{\mathsf{inst}} \vee \varphi_{\mathsf{inst}}$$

where p_{inst} ranges over the set of atomic institutional facts $\mathbb{P}_{\mathsf{inst}}$.

Assignments. Assignments are expressions of the form $p{\leftarrow}\top$ or $p{\leftarrow}\bot$, where p is a propositional variable from $\mathbb{P}$. Assignments and formulas are different entities: the former are *events* modifying the truth values of propositional variables. The event $p{\leftarrow}\top$ sets p to true, and the event $p{\leftarrow}\bot$ sets p to false. We sometimes write $p{\leftarrow}\tau$ in order to talk about $p{\leftarrow}\top$ and $p{\leftarrow}\bot$ in an economic way; τ is therefore a placeholder for either $\top$ or $\bot$.

The sets

$$\begin{aligned}\mathcal{A}_{\mathsf{phys}} &= \{p{\leftarrow}\tau \ : \ p \in \mathbb{P}_{\mathsf{phys}}, \tau \in \{\top, \bot\}\}\\ \mathcal{A}_{\mathsf{inst}} &= \{p{\leftarrow}\tau \ : \ p \in \mathbb{P}_{\mathsf{inst}}, \tau \in \{\top, \bot\}\}\end{aligned}$$

respectively collect the assignments of brute facts and the assignments of institutional facts. Observe that $\mathcal{A}_{\mathsf{phys}} \cap \mathcal{A}_{\mathsf{inst}} = \emptyset$ because $\mathbb{P}_{\mathsf{phys}} \cap \mathbb{P}_{\mathsf{inst}} = \emptyset$. The set of all assignments is $\mathcal{A} = \mathcal{A}_{\mathsf{phys}} \cup \mathcal{A}_{\mathsf{inst}}$. We write $\alpha_{\mathsf{phys}}, \alpha'_{\mathsf{phys}}, \ldots$ to denote assignments from $\mathcal{A}_{\mathsf{phys}}$ and $\alpha_{\mathsf{inst}}, \alpha'_{\mathsf{inst}}, \ldots$ to denote assignments from $\mathcal{A}_{\mathsf{inst}}$. We sometimes use $\alpha, \alpha', \ldots$ to denote generic assignments from $\mathcal{A}$.

Physical facts. We assume that an agent can act upon his environment by assigning values to some propositional variables. We also assume that these variables can only be the atomic physical facts in $\mathbb{P}_{\mathsf{phys}}$, while the values of the atomic institutional facts of $\mathbb{P}_{\mathsf{inst}}$ can only be modified indirectly.

For a physical assignment event $\alpha \in \mathcal{A}_{\mathsf{phys}}$ and an agent $i \in \mathbb{A}$, we formalize that i has the *physical ability* to perform α by writing $\mathrm{A}_i(\alpha)$. We take that the physical ability of an agent to perform an action is itself an atomic physical fact. Moreover, we assume that agent's physical abilities are contingent facts. Hence, we assume that expressions of the form $\mathrm{A}_i(\alpha)$ are propositional variables in $\mathbb{P}_{\mathsf{phys}}$.

We allow this to be recursive: for every ability variable $\mathrm{A}_i(\alpha)$ there is an ability variable $\mathrm{A}_j(\mathrm{A}_i(\alpha){\leftarrow}\top)$ that is also an atomic physical fact.

We suppose that the set of atomic physical facts $\mathbb{P}_{\mathsf{phys}}$ is made up of the basic physical facts of $\mathbb{P}^0_{\mathsf{phys}}$ plus all these ability variables, and nothing else. This set is therefore built according to the following grammar:

$$\mathbb{P}_{\mathsf{phys}} : p_{\mathsf{phys}} ::= p^0_{\mathsf{phys}} \mid \mathrm{A}_i(p_{\mathsf{phys}}{\leftarrow}\tau)$$

where p^0_{phys} ranges over the set of basic brute facts $\mathbb{P}^0_{\mathsf{phys}}$ and τ ranges over the set $\{\top, \bot\}$. Then the language $\mathcal{L}_{\mathsf{phys}}$ of (complex) physical facts is defined by the following grammar:

$$\mathcal{L}_{\mathsf{phys}} : \varphi_{\mathsf{phys}} ::= p_{\mathsf{phys}} \mid \neg\varphi_{\mathsf{phys}} \mid \varphi_{\mathsf{phys}} \vee \varphi_{\mathsf{phys}}$$

where p_{phys} ranges over the set of atomic brute facts $\mathbb{P}_{\mathsf{phys}}$.

'Causes'. We have a binary connective $\Rightarrow$ relating assignments of physical facts: if α_1 and α_2 are both assignments in $\mathcal{A}_{\mathsf{phys}}$ then $\alpha_1 \Rightarrow \alpha_2$ expresses that the performance of α_1 triggers α_2. For example, $p{\leftarrow}\top \Rightarrow q{\leftarrow}\bot$ says that the atomic physical fact q is made false by making the atomic physical fact p true.

'Counts-as'. We have a ternary connective $\rightsquigarrow$ whose arguments are a role, an atomic fact, and an atomic institutional fact, written $\alpha_1 \overset{r}{\rightsquigarrow} \alpha_2$. For instance, we write $p{\leftarrow}\top \overset{r}{\rightsquigarrow} q{\leftarrow}\bot$ to formalize the fact that setting the (physical or institutional) fact p to true while acting in role r counts as setting the institutional fact q to false.

Achieving by doing. Actions are physical events performed by an agent. The formula $\langle i{:}p_{\mathsf{phys}}{\leftarrow}\tau\rangle\varphi$ reads "i can achieve φ by performing $p_{\mathsf{phys}}{\leftarrow}\tau$". By convention, we adopt a strong reading of 'can' and assume that if i does not have the ability to perform $p_{\mathsf{phys}}{\leftarrow}\tau$, then i cannot achieve anything by performing $p_{\mathsf{phys}}{\leftarrow}\tau$.

Definition of the language. The language $\mathcal{L}$ of the logic is fully defined by the following grammar:

$$\begin{aligned}\mathcal{L} : \varphi ::= \; & p_{\mathsf{phys}} \mid p_{\mathsf{inst}} \mid \alpha_{\mathsf{phys}} \Rightarrow \alpha_{\mathsf{phys}} \mid \alpha_{\mathsf{phys}} \overset{r}{\rightsquigarrow} \alpha_{\mathsf{inst}} \mid \alpha_{\mathsf{inst}} \overset{r}{\rightsquigarrow} \alpha_{\mathsf{inst}} \mid \\ & \neg\varphi \mid \varphi \vee \varphi \mid \langle i{:}\alpha_{\mathsf{phys}}\rangle\varphi\end{aligned}$$

where p_{phys} and p_{inst} respectively range over the set of brute facts $\mathbb{P}_{\mathsf{phys}}$ and the set of institutional facts $\mathbb{P}_{\mathsf{inst}}$; α_{phys} and α_{inst} respectively range over the set of assignments of physical facts $\mathcal{A}_{\mathsf{phys}}$ and the set of assignments of institutional facts $\mathcal{A}_{\mathsf{inst}}$; r ranges over the set of roles $\mathbb{R}$; and i ranges over the set of agents $\mathbb{A}$.

The logical constants $\top$ and $\bot$ and the logical connectives $\wedge$, $\rightarrow$ and $\leftrightarrow$ have the usual meaning.

Remarks. A few remarks about our choices of language are in order. As our assignments only operate on $\mathbb{P}$, the truth values of $\alpha \Rightarrow \beta$ and $\alpha \overset{r}{\rightsquigarrow} \beta$ cannot be re-assigned. Assignments being the only means to change things in our logic, it follows that the 'causes' and the 'counts-as' relation do not evolve. We are aware that assuming that causality and counts-as relation are rigid facts of the world is a limitation of our formal theory. For instance, that flipping a switch causes the light to go on can be changed by an action of disconnecting the electric wires; and the counts-as relation in an institution can be changed by way of new agreements, laws, etc.

In contrast, we have modeled the ability relation and the role-playing relation by means of propositional variables $\mathrm{A}_i(\alpha)$ and $R_i(r)$, and these variables are elements of $\mathbb{P}$. Both are therefore contingent facts of the world, just like the physical fact that "the pen is on the table" [23], and their truth values can change due to the performance of assignments.

Let us have a closer look at the way these changes are brought about. First, agent i's ability to perform α being a (non-basic) physical fact of $\mathbb{P}_{\mathsf{phys}}$, that fact can be modified by an assignment event that is performed by some agent j; precisely, some agent j for which $\mathrm{A}_j(\mathrm{A}_i(\alpha){\leftarrow}\top)$ or $\mathrm{A}_j(\mathrm{A}_i(\alpha){\leftarrow}\bot)$ holds. The fact $\mathrm{A}_i(\alpha)$ can also be modified indirectly by the causality relation. For instance, my action of grabbing John's arm causes the loss of John's ability to raise his arm. Second, an agent playing a role being an institutional fact of $\mathbb{P}_{\mathsf{inst}}$, that fact cannot be modified directly by an agent's action: just as any institutional fact, it can only change via the application of the counts-as relation. This allows a very strict control over the institutional consequences of an event.

3.2 Models

A model is a tuple

$$M = (\mathbb{A}, \mathbb{R}, \mathbb{P}^0_{\mathsf{phys}}, \mathbb{P}^0_{\mathsf{inst}}, V_{\mathsf{phys}}, V_{\mathsf{inst}}, C_{\mathsf{phys}}, C_{\mathsf{inst}})$$

where the sets $\mathbb{A}$, $\mathbb{R}$, $\mathbb{P}^0_{\mathsf{phys}}$ and $\mathbb{P}^0_{\mathsf{inst}}$ are as detailed above; $V_{\mathsf{phys}} \subseteq \mathbb{P}^0_{\mathsf{phys}}$ and $V_{\mathsf{inst}} \subseteq \mathbb{P}^0_{\mathsf{inst}}$ are valuations describing which atomic fact are true; $C_{\mathsf{phys}} \subseteq \mathcal{A}_{\mathsf{phys}} \times \mathcal{A}_{\mathsf{phys}}$ is a relation between assignments of physical facts; and $C_{\mathsf{inst}} : \mathbb{R} \longrightarrow 2^{\mathcal{A} \times \mathcal{A}_{\mathsf{inst}}}$ is a function mapping roles to relations between assignments and institutional assignments. When $(\alpha_1, \alpha_2) \in C_{\mathsf{phys}}$ then the occurrence of action α_1 causes the occurrence of action α_2. When $(\alpha_1, \alpha_2) \in C_{\mathsf{inst}}(r)$ then the occurrence of action α_1 performed by an agent playing role r counts as the occurrence of action α_2. The relations C_{phys} and $C_{\mathsf{inst}}(r)$ are nothing but the relations of "causal generation" and "conventional generation" in Goldman's sense as described in Section 2.

3.3 Constraints on models

Models have to satisfy the following two constraints:

($Refl_{\mathsf{phys}}$) C_{phys} is reflexive:
for every $\alpha \in \mathcal{A}_{\mathsf{phys}}$, $(\alpha, \alpha) \in C_{\mathsf{phys}}$.

(Coh_{phys}) C_{phys} is coherent:
for every $\alpha \in \mathcal{A}_{\mathsf{phys}}$ and $q \in \mathbb{P}$, if $(\alpha, q \leftarrow \top) \in C_{\mathsf{phys}}$ then $(\alpha, q \leftarrow \bot) \notin C_{\mathsf{phys}}$.

($Trans_{\mathsf{phys}}$) C_{phys} is transitive:
$C_{\mathsf{phys}} \circ C_{\mathsf{phys}} \subseteq C_{\mathsf{phys}}$

(Coh_{inst}) C_{inst} is coherent:
for every $\alpha \in \mathcal{A}$, $q \in \mathbb{P}$, and $r_1, r_2 \in \mathbb{R}$,
if $(\alpha, q \leftarrow \top) \in C_{\mathsf{inst}}(r_1)$ then $(\alpha, q \leftarrow \bot) \notin C_{\mathsf{inst}}(r_2)$.

($Trans_{\mathsf{phys,inst}}$) C_{phys} and C_{inst} satisfy a mixed transitivity property:
for every $r \in \mathbb{R}$, $C_{\mathsf{phys}} \circ C_{\mathsf{inst}}(r) \subseteq C_{\mathsf{inst}}(r)$.

In the rest of the section we briefly discuss these properties in the light of our exposition in Section 2. (See also [28] for a review of properties of causality relations in the framework of artificial intelligence.)

The constraint ($Refl_{\mathsf{phys}}$) means that we consider that causality is reflexive. We are aware that this can be criticized because causes temporally precede their effects. It however simplifies the technicalities when updating models; the reader may wish to think of it as the reflexive closure of the causality relation. (Coh_{phys}) says that a physical action cannot have inconsistent causal ramifications. (Coh_{inst}) is a similar principle for the counts-as relation: for every assignment α there cannot be two roles leading to inconsistent consequences via the counts-as relations. Constraint ($Trans_{\mathsf{phys,inst}}$) is the institutional counterpart of the transitivity of causality as expressed by ($Trans_{\mathsf{phys}}$).

Reflexivity of event generation relations is rejected by Goldman [8, p. 5] on the simple ground that it is not intuitive to say that John turns on the light by turning on the light. In our proposal, the counts-as relation is not necessarily reflexive; we however allow that $\alpha \overset{r}{\rightsquigarrow} \alpha$ if α is an institutional action. Moreover, our modeling of the causality relation assumes reflexivity. This is essentially motivated by the fact that this way, our definitions are less cluttered, but since Goldman's argument against reflexivity is merely linguistic, we believe it is not a major conceptual transgression.

Goldman insists that an event generation relation should be antisymmetric. We neither preclude the symmetry of the causality relation nor of the counts-as relation since we could have for instance that a subset of events form an equivalence class in which all events causally generate all events.

It is worth noting that there is some disagreement in the literature whether the counts-as relation should satisfy transitivity. For a discussion on this matter see [17, 11, 21]. In our logic this is not necessarily the case.

Finally, some author argues that the counts-as should satisfy contraposition [11], while other authors have a different opinion on this matter [17]. Again, we remain uncommitted w.r.t. this point, and it may be the case that $(p{\leftarrow}\top, q{\leftarrow}\top) \in C_{\mathsf{inst}}(r)$ while $(q{\leftarrow}\bot, p{\leftarrow}\bot) \notin C_{\mathsf{inst}}(r)$.

3.4 Updating a model by an action

An agent's capability can be represented semantically by the valuations V' his actions can bring about. This is achieved by interpreting the agents' actions as model updates.

Definition 1. *Let*

$$M = (\mathbb{A}, \mathbb{R}, \mathbb{P}^0_{\mathsf{phys}}, \mathbb{P}^0_{\mathsf{inst}}, V_{\mathsf{phys}}, V_{\mathsf{inst}}, C_{\mathsf{phys}}, C_{\mathsf{inst}})$$

be a model and let $\alpha \in \mathcal{A}_{\mathsf{phys}}$. The update of M by the action $i{:}\alpha$ is defined as

$$M^{i:\alpha} = (\mathbb{A}, \mathbb{R}, \mathbb{P}^0_{\mathsf{phys}}, \mathbb{P}^0_{\mathsf{inst}}, V^{\alpha}_{\mathsf{phys}}, V^{i:\alpha}_{\mathsf{inst}}, C_{\mathsf{phys}}, C_{\mathsf{inst}})$$

where the updates $V^{i:\alpha}_{\mathsf{phys}}$ and $V^{i:\alpha}_{\mathsf{inst}}$ of the valuations V_{phys} and V_{inst} by $i{:}\alpha$ are defined as follows:

$$\begin{aligned} V^{\alpha}_{\mathsf{phys}} &= \Big(V_{\mathsf{phys}} \setminus \{q \ : \ (\alpha, q{\leftarrow}\bot) \in C_{\mathsf{phys}}\} \Big) \cup \ \{q \ : \ (\alpha, q{\leftarrow}\top) \in C_{\mathsf{phys}}\} \\ V^{i:\alpha}_{\mathsf{inst}} &= (\ V_{\mathsf{inst}} \setminus \{q \ : \ \exists r \in \mathbb{R} : R_i(r) \in V_{\mathsf{inst}}, (\alpha, q{\leftarrow}\bot) \in C_{\mathsf{inst}}(r)\} \ ) \\ &\qquad \cup \{q \ : \ \exists r \in \mathbb{R} : R_i(r) \in V_{\mathsf{inst}}, (\alpha, q{\leftarrow}\top) \in C_{\mathsf{inst}}(r)\} \end{aligned}$$

Actions therefore (1) directly affect the physical world (via the causality relation), and (2) affect the institutional world via the counts-as relation.[5] Suppose that $\alpha = p{\leftarrow}\top$. Then, due to reflexivity of the causality relation C_{phys}, the valuation $V^{p\leftarrow\top}_{\mathsf{phys}}$ contains p and $V^{p\leftarrow\bot}_{\mathsf{phys}}$ does not contain p. Note that the physical valuation is actually updated by the event α, not by the action $i{:}\alpha$.

Our constraints on models are clearly preserved under updates because neither the causal relation nor the counts-as relation can be modified.

Let us illustrate our definition by a couple of examples.

[5] Note that the order of the set theoretic operations in the definition seems to privilege positive facts; however, due to our two constraints (Coh_{phys}) and (Coh_{inst}) —and also because $\mathbb{P}_{\mathsf{phys}}$ and $\mathbb{P}_{\mathsf{inst}}$ have empty intersection— the ramifications of an assignment of a physical fact will never conflict.

Example 1. Suppose V_{inst} contains $R_i(r_1)$ and $p{\leftarrow}\top \overset{r_1}{\rightsquigarrow} q{\leftarrow}\top$, i.e. agent i plays role r_1, and in role r_1 making p true counts as making q true. Then $V_{\mathsf{phys}}^{p\leftarrow\top}$ contains p, and $V_{\mathsf{inst}}^{i:p\leftarrow\top}$ contains q. Hence, under the hypothesis that V_{phys} contains $\mathrm{A}_i(p{\leftarrow}\top)$ (that is that agent i is indeed able to make p true), agent i can achieve about $p \wedge q$ by doing $p{\leftarrow}\top$.

Example 2. Suppose V_{inst} contains $R_i(r_1)$ and $R_i(r_2)$, and $(p{\leftarrow}\top, q_1{\leftarrow}\top) \in C_{\mathsf{inst}}(r_1)$ and $(p{\leftarrow}\top, q_2{\leftarrow}\bot) \in C_{\mathsf{inst}}(r_2)$, i.e. agent i plays two roles r_1 and r_2, and in role r_1 this counts as making q_1 true, while in role r_2 this counts as making q_2 false. Then in $M^{i:p\leftarrow\top}$, the valuation $V_{\mathsf{phys}}^{p\leftarrow\top}$ contains p and $V_{\mathsf{inst}}^{i:p\leftarrow\top}$ contains q_1 and does not contain q_2. Hence, assuming that V_{phys} contains $\mathrm{A}_i(p{\leftarrow}\top)$ (that is agent i is indeed able to make p true), agent i can achieve $p \wedge q_1 \wedge \neg q_2$ by doing $p{\leftarrow}\top$.

3.5 Truth conditions

Let

$$M = (\mathbb{A}, \mathbb{R}, \mathbb{P}^0_{\mathsf{phys}}, \mathbb{P}^0_{\mathsf{inst}}, V_{\mathsf{phys}}, V_{\mathsf{inst}}, C_{\mathsf{phys}}, C_{\mathsf{inst}})$$

be a model. The truth conditions are as usual for the Boolean operators, and we only state those clauses that are not standard.

$$\begin{array}{ll}
M \models p_{\mathsf{phys}} & \text{iff } p_{\mathsf{phys}} \in V_{\mathsf{phys}} \\
M \models p_{\mathsf{inst}} & \text{iff } p_{\mathsf{inst}} \in V_{\mathsf{inst}} \\
M \models \alpha_1 \Rightarrow \alpha_2 & \text{iff } (\alpha_1, \alpha_2) \in C_{\mathsf{phys}} \\
M \models \alpha_1 \overset{r}{\rightsquigarrow} \alpha_2 & \text{iff } (\alpha_1, \alpha_2) \in C_{\mathsf{inst}}(r) \\
M \models \langle i{:}\alpha_{\mathsf{phys}}\rangle\varphi & \text{iff } \mathrm{A}_i(\alpha_{\mathsf{phys}}) \in V_{\mathsf{phys}} \text{ and } M^{i:\alpha_{\mathsf{phys}}} \models \varphi
\end{array}$$

The operator $\langle i{:}\alpha_{\mathsf{phys}}\rangle\varphi$ captures the notion of "achieving by doing" that has been sketched in Section 3.1 and in the two examples of the last sub-section.

4 Axiomatization and complexity

The logic is axiomatized as an extension of classical propositional logic with (1) a theory describing the constraints imposed on the counts-as and causality relations, (2) the reduction axioms of the dynamic operator, and (3) an inference rule of replacement of equivalents in the scope of a dynamic operator.

Theory of counts-as and causality.

$$\begin{array}{l}
\alpha \Rightarrow \alpha \\
(\alpha \Rightarrow p{\leftarrow}\bot) \rightarrow \neg(\alpha \Rightarrow p{\leftarrow}\top) \\
(\alpha \overset{r_1}{\rightsquigarrow} p{\leftarrow}\top) \rightarrow \neg(\alpha \overset{r_2}{\rightsquigarrow} p{\leftarrow}\bot) \\
((\alpha_1 \Rightarrow \alpha_2) \wedge (\alpha_2 \Rightarrow \alpha_3)) \rightarrow (\alpha_1 \Rightarrow \alpha_3) \\
((\alpha_1 \Rightarrow \alpha_2) \wedge (\alpha_2 \overset{r_1}{\rightsquigarrow} \alpha_3)) \rightarrow (\alpha_1 \overset{r_1}{\rightsquigarrow} \alpha_3)
\end{array}$$

Reduction axioms for the dynamic operator.

$$
\begin{array}{ll}
\langle i{:}\alpha\rangle\top & \leftrightarrow \mathrm{A}_i(\alpha) \\
\langle i{:}\alpha\rangle\bot & \leftrightarrow \bot \\
\langle i{:}\alpha\rangle(\alpha_1 \Rightarrow \alpha_2) & \leftrightarrow \mathrm{A}_i(\alpha) \wedge (\alpha_1 \Rightarrow \alpha_2) \\
\langle i{:}\alpha\rangle(\alpha_1 \overset{r}{\rightsquigarrow} \alpha_2) & \leftrightarrow \mathrm{A}_i(\alpha) \wedge (\alpha_1 \overset{r}{\rightsquigarrow} \alpha_2) \\
\langle i{:}\alpha\rangle p_{\mathsf{phys}} & \leftrightarrow \mathrm{A}_i(\alpha) \wedge (\ (\alpha \Rightarrow p_{\mathsf{phys}}{\leftarrow}\top) \vee (p_{\mathsf{phys}} \wedge \neg(\alpha \Rightarrow p_{\mathsf{phys}}{\leftarrow}\bot))\ ) \\
\langle i{:}\alpha\rangle p_{\mathsf{inst}} & \leftrightarrow \mathrm{A}_i(\alpha) \wedge (\ \bigvee_{r\in\mathbb{R}}(R_i(r) \wedge (\alpha \overset{r}{\rightsquigarrow} p_{\mathsf{inst}}{\leftarrow}\top)) \\
& \qquad\qquad \vee(p_{\mathsf{inst}} \wedge \neg \bigvee_{r\in\mathbb{R}}(R_i(r) \wedge (\alpha \overset{r}{\rightsquigarrow} p_{\mathsf{inst}}{\leftarrow}\bot)))\ ) \\
\langle i{:}\alpha\rangle\neg\varphi & \leftrightarrow \mathrm{A}_i(\alpha) \wedge \neg\langle i{:}\alpha\rangle\varphi \\
\langle i{:}\alpha\rangle(\varphi \vee \psi) & \leftrightarrow \mathrm{A}_i(\alpha) \wedge (\langle i{:}\alpha\rangle\varphi \vee \langle i{:}\alpha\rangle\psi)
\end{array}
$$

Inference rule.

$$\text{From } \varphi \leftrightarrow \psi \text{ infer } \langle i{:}\alpha\rangle\varphi \leftrightarrow \langle i{:}\alpha\rangle\psi$$

Proofs are defined in the standard way. For example, the rule of replacement of equivalents can be proved from our axiomatization (due to the inference rule).

Given a formula φ let $red(\varphi)$ be the formula obtained by iterating the application of the reduction axioms from the left to the right. Thanks to the rule of replacement of equivalents it is clear that $red(\varphi) \leftrightarrow \varphi$ is valid.

Proposition 1. *For every formula φ, $red(\varphi) \leftrightarrow \varphi$ is valid, and the length of $red(\varphi)$ is linear in the length of φ.*

Proposition 2. *Let φ be a formula without dynamic operators $\langle . \rangle$. φ is valid if and only if $\mathcal{T}_\varphi \rightarrow \varphi$ is valid in classical propositional logic, where $\mathcal{T}_\varphi$ is the conjunction of the axiom schemas of the theory of counts-as and causality instantiated by those assignments occurring in φ.*

As the length of $\mathcal{T}_\varphi$ is cubic in the length of φ we obtain a complexity result for our logic.

Corollary 1. *The problem of checking satisfiability of a formula is NP-complete.*

Our logic has therefore the same complexity as classical propositional logic. We however believe that it allows to express things in a more natural way. In the rest of the paper we give some arguments for this.

5 Institutional power and compact characterization¡

Let S be a finite set of assignments. We identify the concept "agent i has the capability to achieve outcome φ by possibly performing one of the actions in the set S" with the formula

$$\Diamond_{i:S}\varphi \overset{\mathrm{def}}{=} \varphi \vee \bigvee_{\alpha\in S} \langle i{:}\alpha\rangle\varphi$$

If φ belongs to the language of institutional facts $\mathcal{L}_{\mathsf{inst}}$ then i's capability to achieve φ can be rightly called i's *institutional power* to achieve φ by doing actions from S.

We now illustrate our logic by adapting an example of water management from [14]. In that paper, beyond abilities $\mathrm{A}_i(\alpha)$ there are also atomic facts $\mathrm{P}_i(\alpha)$ whose intended meaning is that agent i is permitted to perform α. Clearly, it seems natural to assume that such atomic facts are institutional facts from $\mathbb{P}_{\mathsf{inst}}$. Note that any combination of abilities and permissions is consistent: an agent might be able to perform α but not permitted to do so, etc.

Example 3. There are two farmers i_1 and i_2 working in a certain area close to a town called Thirstytown who need water in order to irrigate their fields. In this area there are three different exploitable water basins 1, 2 and 3. Only water basins 1 and 2 can be safely used by the farmers; basin 3 provides drinkable water to the population of Thirstytown, and if it is exploited for irrigation then Thirstytown will fall short of water. There are two other actors in this scenario: agent i_3 plays the role of chief of the Water Authority which has the jurisdiction over the area, and agent i_4 is a local policeman working in Thirstytown. Let wAuth denote the role of head of water authority, and let pol denote the role of policeman. We consider that $R_{i_3}(\mathsf{wAuth})$ and $R_{i_4}(\mathsf{pol})$ are both true.

The propositional variables $\{p_1, p_2, p_3\}$ indicate whether the level of water in a given basin is high or low: p_1 means that "the level of water in the basin 1 is high", $\neg p_1$ means that "the level of water in the basin 1 is low", etc. Furthermore, for every farmer $i_k \in \{i_1, i_2\}$ and for every propositional variable p_h with $h \in \{1, 2, 3\}$, $\mathrm{A}_{i_k}(p_h{\leftarrow}\bot)$ expresses that basin h is physically connected to the field of farmer i_k so that i_k is able to exploit the water of basin h and $\mathrm{P}_{i_k}(p_h{\leftarrow}\bot)$ expresses that i_k is authorized to exploit the water of basin h.

Let $\mathsf{prohibSign}_h$ mean that there are prohibition-to-pump signs at basin h. We suppose that the the counts-as relation is such that $\mathsf{prohibSign}_h{\leftarrow}\top \overset{\mathsf{pol}}{\rightsquigarrow} \mathrm{P}_{i_k}(p_h{\leftarrow}\bot){\leftarrow}\bot$ for every k and h: to make $\mathsf{prohibSign}_h$ true action while performing the policeman role counts as disallowing to anybody to pump water from that basin.

Our causality relation allows us to model indirect effects of actions. For example, basin 1 being close to basin 2, farmer pumping from 1 also lowers the water level in basin 2. This can be expressed by stating $p_1{\leftarrow}\bot \Rightarrow p_2{\leftarrow}\bot$.

Our definition of capability is only about performance of a single action from the set S. It can be generalized by allowing for arbitrary combinations of actions from S. Let us introduce a modal operator of *iterative capability* $\Diamond^*_{i:S}$ whose truth condition is:

$$M \models \Diamond^*_{i:S}\varphi \;\text{ iff }\; \text{there is a sequence } (\alpha_1, \ldots, \alpha_n) \text{ of assignments from } S \text{ such that } M \models \langle i{:}\alpha_1\rangle \ldots \langle i{:}\alpha_n\rangle\varphi$$

It is useful to first introduce the dynamic logic program operators `skip` and $\cup$. Their semantics requires to move from the functional interpretation of actions to a relational interpretation: now for every action $i{:}\alpha$, $R_{i:\alpha}$ relates models M to their updates $M^{i:\alpha}$. The recursive definition of R_π is as follows, where π is a program:

$$\begin{aligned} R_{i:\alpha} &= \{(M, M') \;:\; M' = M^{i:\alpha}\} \\ R_{\mathtt{skip}} &= \{(M, M') \;:\; M' = M\} \\ R_{\pi_1 \cup \pi_2} &= R_{\pi_1} \cup R_{\pi_2} \end{aligned}$$

We therefore have $\langle \texttt{skip} \rangle \varphi \leftrightarrow \varphi$ and $\langle i{:}\alpha \cup j{:}\beta \rangle \varphi \leftrightarrow \langle i{:}\alpha \rangle \varphi \vee \langle j{:}\beta \rangle \varphi$. The truth condition becomes:

$$M \models \langle \pi \rangle \varphi \text{ iff } M' \models \varphi \text{ for every } M' \text{ such that } MR_\pi M'$$

As a simple illustration, observe that the above definition of 'one shot capability' $\Diamond_{i:S}$ can now be written in an alternative and more elegant way. Let $S = \{\alpha_1, \ldots, \alpha_n\}$ be a set of assignments. We have:

$$\Diamond_{i:S} \varphi \leftrightarrow \langle \texttt{skip} \cup i{:}\alpha_1 \cup \ldots \cup i{:}\alpha_n \rangle \varphi$$

The next result states that in order to check whether $\Diamond^*_{i:S} \varphi$ it suffices to check whether φ can be obtained by performing some of the assignments in S once, in any order, provided that abilities are not used before being acquired, and are not abandoned before used.

Proposition 3. *Let $S = \{\alpha_1, \ldots, \alpha_{\mathsf{card}(S)}\}$ be a set of assignments. The formula*

$$\Diamond^*_{i:S} \varphi \leftrightarrow \langle \texttt{skip} \cup i{:}\alpha_1 \rangle \cdots \langle \texttt{skip} \cup i{:}\alpha_{\mathsf{card}(S)} \rangle \varphi$$

is valid, where $(\alpha_1, \ldots, \alpha_{\mathsf{card}(S)})$ is any ordering of the elements of S such that for all $\alpha_l \in S$, whenever $\alpha_k = \mathrm{A}_i(\alpha_l) \leftarrow \top$ then $k < l$, and whenever $\alpha_k = \mathrm{A}_i(\alpha_l) \leftarrow \bot$ then $l < k$.

Proof (sketch). *Suppose $M \models \Diamond^*_{i:S} \varphi$. Hence there is a sequence $(\alpha_1, \ldots, \alpha_n)$ of assignments from S such that $M \models \langle i{:}\alpha_1 \rangle \ldots \langle i{:}\alpha_n \rangle \varphi$. We can permute assignments and put them in the appropriate order by applying the following valid equivalences.*

$$\langle i{:}p \leftarrow \tau \rangle \langle i{:}q \leftarrow \tau' \rangle \varphi \leftrightarrow \begin{cases} \langle i{:}q \leftarrow \tau' \rangle \varphi & \textit{when } q = p \\ \langle i{:}q \leftarrow \tau' \rangle \langle i{:}p \leftarrow \tau \rangle \varphi & \textit{when } q \neq p \end{cases}$$

$$\langle i{:}\alpha \rangle \langle i{:}\mathrm{A}_j(\beta) \leftarrow \top \rangle \varphi \leftrightarrow \begin{cases} \langle i{:}\alpha \rangle \varphi & \textit{when } \beta = \alpha \textit{ and } j = i \\ \langle i{:}\mathrm{A}_j(\beta) \leftarrow \top \rangle \langle i{:}\alpha \rangle \varphi & \textit{when } \beta \neq \alpha \textit{ or } j \neq i \end{cases}$$

$$\langle i{:}\mathrm{A}_j(\beta) \leftarrow \bot \rangle \langle i{:}\alpha \rangle \varphi \leftrightarrow \begin{cases} \bot & \textit{when } \beta = \alpha \textit{ and } j = i \\ \langle i{:}\alpha \rangle \langle i{:}\mathrm{A}_j(\beta) \leftarrow \bot \rangle \varphi & \textit{when } \beta \neq \alpha \textit{ or } j \neq i \end{cases}$$

The first equivalence allows to eliminate multiple occurrences of the same assignment from S. The second equivalence allows to move ability gain assignments to the left, while the third equivalence allows to move ability loss assignments to the right. Together, these three equivalences allow to replace $\langle i{:}\alpha_1 \rangle \ldots \langle i{:}\alpha_n \rangle \varphi$ by the equivalent $\langle i{:}\beta_1 \rangle \ldots \langle i{:}\beta_m \rangle \varphi$ such that for all $\beta_l \in S$, whenever $\beta_k = \mathrm{A}_i(\beta_l) \leftarrow \top$ then $k < l$, and whenever $\beta_k = \mathrm{A}_i(\beta_l) \leftarrow \bot$ then $l < k$. It finally follows from the valid implication $\psi \to \langle \texttt{skip} \cup i{:}\beta \rangle \psi$ that those elements of S that are not in our sequence yet can be inserted, and that $M \models \langle \texttt{skip} \cup i{:}\beta_1 \rangle \ldots \langle \texttt{skip} \cup i{:}\beta_{\mathsf{card}(S)} \rangle \varphi$ where $(\beta_1, \ldots, \beta_{\mathsf{card}(S)})$ is an ordering of the elements of S satisfying the condition of the proposition.

The other direction of the proof is straightforward. ■

Note that unfolding the right-hand side of the equivalence in Proposition 3 yields a formula in $\mathcal{L}$ that is exponentially larger. In fact, extending the language $\mathcal{L}$ with the program construct $\cup$ increases the complexity of the logic from NP-complete to PSPACE-complete.[6] This seems to indicate that reasoning about the notion of iterative capability with the operator $\Diamond^*_{i:S}\varphi$ is computationally more expensive.

6 Related works

In the last decade several logicians have focused on a number of aspects of counts-as such as institutional power [17], defeasibility [3, 9], contextual and classificatory aspects [11], mental aspects [21, 20], the distinction between brute facts and institutional facts [10].

In their seminal paper [17], Jones and Sergot gave the status of an implication-like logical connective to the counts-as relation. The latter links two propositions φ_1 and φ_2 within a normative system (or institution) s. This is formally written $\varphi_1 \leadsto_s \varphi_2$ and reads "φ_1 counts as φ_2 in s". Jones and Sergot gave a possible worlds semantics for the counts-as connective together with an axiomatic characterization of the valid formulas of that logic. In order to capture the notion of institutional power they extended their logic with an action component: the 'bringing it about that' modal operator $E_i\varphi$ which has to be read "agent i brings it about that φ". $E_i\varphi_1 \leadsto_s E_i\varphi_2$ then expresses that "in s, i's action of bringing about φ_1 counts as i's action of bringing about φ_2".

In a more recent paper [11], Grossi and colleagues paved the way towards a substantial simplification of Jones and Sergot's logic. Contrarily to the latter they did not consider the counts-as relation as primitive: the basic logical operators of Grossi et col.'s logic are normal modal operators of the form $[s]\varphi$, reading "in normative system s, it is the case that φ". These operators can be combined with the standard Boolean operators. For example, $[s](\varphi \to \psi)$ is a formula of the language of Grossi et col., reading "in s, if φ then ψ". Based on the $[s]$ connectives, Grossi et col. then define the counts-as connective. First of all they argue that the formula $[s](\varphi \to \psi)$ is already an approximation of $\varphi \leadsto_s \psi$. Nevertheless, this approximation validates formulas such as $\varphi \leadsto_s \top$: in s, any φ counts as a tautology. This is felt to be counter-intuitive. Therefore, in order to better capture Jones and Sergot's $\leadsto_s$ connective, Grossi et col. introduce a so-called universal modality $[\forall]$, where $[\forall]\varphi$ reads "φ universally holds". The latter is used in order to strengthen the link between φ and ψ: in addition to $[s](\varphi \to \psi)$, Grossi et col. moreover require that for φ to count as ψ it should not be universally true that φ implies ψ. In formulas, they define a so-called *proper classificatory rule* $\varphi \leadsto^{cl+}_s \psi$ by stipulating:

$$\varphi \leadsto^{cl+}_s \psi \stackrel{\mathtt{def}}{=} [s](\varphi \to \psi) \land \neg[\forall](\varphi \to \psi)$$

In this way they guarantee that no φ counts as a tautology.

There are several novel aspects in our logical analysis of counts-as. First of all, differently from other approaches, our framework allows to explicitly represent physical actions and institutional actions, as well as the links between the two kinds of actions. By distinguishing in the object language a counts-as relation from a causal relation between events, our logic clearly opposes Goldman's notion of causal generation to that

[6] There is an easy reduction from QSAT, see e.g. [14].

of conventional generation. We have shown that these two relations are ontologically different for at least two reasons. While the former relates a physical action to another physical action, the latter relates a physical action to an institutional action. Moreover, while the causal relation is merely a relation between physical actions performed by an agent, counts-as is a relation between actions performed by an agent playing a certain role.

Furthermore, while previous logical accounts of counts-as were mainly conceptual and did not consider decidability issues, our work also focuses on the computational aspects of a logic of institutional action: we have provided in Section 4 a complete axiomatization of our logic based on reduction axioms and have characterized the complexity of the satisfiability problem.

We note that technically, our reduction axioms in terms of a causality relation are close to causality-based solutions to the ramification problem in reasoning about actions [18, 24, 25, 31, 28].

7 Conclusion

In the framework presented in this paper counts-as and causal relations are static, that is, there is no way to update models in order to modify these relations. An interesting direction of future research is to integrate into the framework a dynamic dimension of counts-as and causality in order to be able to model interesting phenomena such as: (1) the modification of causal connections between physical events (e.g. by disconnecting the electric wires, I can remove the causal relation *"flipping the switch" causes "turning on the lights"*); (2) norm promulgation (creating a new counts-as relation between events); (3) norm cancellation (removing a pre-existent counts-as relation between events).

Another interesting topic of future research is the creation of institutional facts. We intend to extend our logic in order to model scenarios such as the following one. Before 2000, it was not possible to assign a truth value to the sentence "he has a note of 50 Euro in his pocket", as the concept "Euro" was not an element of our vocabulary of institutional facts and objects. After its introduction the Euro became an element of our vocabulary of institutional facts and objects. This might be integrated into our framework by adapting approaches to awareness such as [13].

Finally, at the current stage our logic allows to clearly distinguish physical actions with physical effects from institutional actions with institutional effects. Nevertheless, it does not support reasoning about physical actions that an agent may decide to perform on the basis of his preferences. A further interesting direction of future research is to relate our framework with game and decision theory by introducing a notion of preference. This extended framework will allow to reason about situations in which agents desire that certain physical and/or institutional facts obtain, and choose strategically a given physical action in order to ensure these facts.

8 Acknowledgements

We are grateful to the reviewers of CLIMA XII. Nicolas Troquard is supported by a Marie Curie Trentino Fellowship.

References

1. C. Alchourrón and E. Bulygin. *Normative systems*. Springer, New York, 1971.
2. J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. *Information and Computation*, 204(204):1620–1662, 2006.
3. G. Boella and L. van der Torre. Regulative and constitutive norms in normative multiagent systems. In *KR 2004*, pages 255–266. AAAI Press, 2004.
4. E. Bulygin. On norms of competence. *Law and Philosophy*, 11(3):201–216, 1992.
5. D. Davidson. Agency. In *Essays on Actions and Events*. Oxford University Press, New York, 1980.
6. H. P. van Ditmarsch, W. van der Hoek, and B. Kooi. Dynamic epistemic logic with assignment. In *AAMAS'05*, pages 141–148. ACM Press, 2005.
7. J. Gerbrandy. Logics of propositional control. In *Proc. AAMAS'06*, pages 193–200, 2006.
8. A. Goldman. *A Theory of Human Action*. Prentice-Hall, Englewood Cliffs NJ, 1970.
9. G. Governatori, J. Gelati, A. Rotolo, and G. Sartor. Actions, institutions, powers: preliminary notes. In *Proceedings of the International Workshop on Regulated Agent-Based Social Systems: Theories and Applications (RASTA'02)*, pages 131–147, 2002.
10. D. Grossi. A note on brute vs. institutional facts: modal logic of equivalence up to a signature. In G. Boella, G. Pigozzi, P. Noriega, and H. Verhagen, editors, *Proceedings of Dagstuhl Seminar on Normative Multi-agent Systems*, 2009.
11. D. Grossi, J.-J. Ch. Meyer, and F. Dignum. Classificatory aspects of counts-as: An analysis in modal logic. *Journal of Logic and Computation*, 16(5):613–643, 2006.
12. H. L. A. Hart. *The concept of law*. Clarendon Press, Oxford, 1992. new edition.
13. A. Heifetz, M. Meier, and B. Schipper. Interactive unawareness. *Journal of Economic Theory*, 130(1):78–94, 2006.
14. A. Herzig, E. Lorini, F. Moisan, and N. Troquard. A dynamic logic of normative systems (regular paper). In T. Walsh, editor, *Proc. 22nd Int. Joint Conf. on Artificial Intelligence (IJCAI'11), Barcelona*. Morgan Kaufmann Publishers, 2011.
15. W. van der Hoek, D. Walther, and M. Wooldridge. Reasoning about the transfer of control. *JAIR*, 37:437–477, 2010.
16. W. van der Hoek and M. Wooldridge. On the logic of cooperation and propositional control. *Artificial Intelligence*, 164(1-2):81–119, 2005.
17. A. Jones and M. J. Sergot. A formal characterization institutionalised power. *J. of the IGPL*, 4:429–445, 1996.
18. F. Lin. Embracing causality in specifying the indirect effects of actions. In *Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pages 1985–1991, Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95), 1995.
19. E. Lorini and A. Herzig. A logic of intention and attempt. *Synthese*, 163(1):45–77, 2008.
20. E. Lorini and D. Longin. A logical account of institutions: from acceptances to norms via legislators. In *Proceedings of the International Conference on Principles of Knowledge Representation and Reasoning (KR 2008)*, pages 38–48. AAAI Press, 2008.
21. E. Lorini, D. Longin, B. Gaudou, and A. Herzig. The logic of acceptance: grounding institutions on agents' attitudes. *J. of Logic and Computation*, 19(6):901–940, 2009.

22. D. Makinson. On the formal representation of rights relations: remarks on the work of Stig Kanger and Lars Lindahl. *The Journal of Philosophical Logic*, 15:403–425, 1986.
23. C. Masolo, L. Vieu, E. Bottazzi, C. Catenacci, R. Ferrario, A. Gangemi, and N. Guarino. Social roles and their descriptions. In D. Dubois, C. A. Welty, and W.-A. Williams, editors, *KR*, pages 267–277. AAAI Press, 2004.
24. N. McCain and H. Turner. A causal theory of ramifications and qualifications. In *Proc. 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pages 1978–1984, 1995.
25. N. McCain and H. Turner. Causal theories of action and change. In *Proc. Nat. Conf. on Artificial Intelligence (AAAI'97)*, pages 460–465, 1997.
26. M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.
27. I. Pörn. *Action Theory and Social Science: Some Formal Models*. Synthese Library 120. D. Reidel, Dordrecht, 1977.
28. C. Schwind. Causality in action theories. *ETAI Electronic Articles in Computer and Information Science, www.ida.liu.se/ext/etai*, 3(A), 1999.
29. J. R. Searle. *Speech acts: An essay in the philosophy of language*. Cambridge University Press, New York, 1969.
30. J. R. Searle. *The Construction of Social Reality*. The Free Press, New York, 1995.
31. M. Thielscher. Ramification and causality. *Artificial Intelligence*, 89:317–364, 1997.
32. M.L. Tiomkin and J.A. Makowsky. Propositional dynamic logic with local assignments. *Theor. Comput. Sci.*, 36:71–87, 1985.
33. J. van Eijck. Making things happen. *Studia Logica*, 66(2):41–58, 2000.