Auton Agent Multi-Agent Syst manuscript No.
(will be inserted by the editor)

Reasoning about coalitional agency and ability in the logics of "bringing-it-about"

Nicolas Troquard

Received: date / Accepted: date

Abstract The logics of "bringing-it-about" have been part of a prominent tradition for the formalization of individual and institutional agency. They are the logics to talk about what states of affairs an acting entity brings about while abstracting away from the means of action. Elgesem's proposal analyzes the agency of individual agents as the goal-directed manifestation of an individual ability. It has become an authoritative modern reference. The first contribution of this paper is to extend Elgesem's logic of individual agency and ability to coalitions. We present a general theory and later propose several possible specializations. As a second contribution, we offer algorithms to reason with the logics of bringing-it-about and we analyze their computational complexity.

Keywords logic · bringing-it-about · coalitions · agency · ability · complexity

1 Introduction

We aim to contribute to the literature that views an action as the mere result of the activity of an agent.[1] It is generally acknowledged that this tradition dates back at least to St. Anselm who claimed that the phenomenon of an action is better explained by what is brought about. This is to be distinguished from

N. Troquard
Laboratory for Applied Ontology
Institute of Cognitive Sciences and Technologies (CNR)
via alla Cascata 56/C
Povo
38123 Trento, Italy
Tel.: +39 0461 314843
Fax: +39 0461 314875
E-mail: troquard@loa.istc.cnr.it

[1] We develop the research agenda presented in the earlier extended abstract published as [47].

other traditions of logic of action talking explicitly about action terms: for instance, Dynamic Logics ([21]) in computer science, or the study of action sentences in philosophy using first-order theories ([14]).

In the Anselmian tradition, a sentence "Jones pays his rent" is interpreted as "Jones *does that* his rent is paid". A logic equipped with a modality "does-that" is particularly adequate for representing and reasoning about *ex post acto* properties. In particular, it enables the reasoning about the responsibility of an acting entity for a presently achieved state of affairs. Logics typically excel in abstraction and modularity. This is no different for logics of the modality does-that. For instance, relations of influence between agents—coercive or not, positive or negative—are simply captured by combined statements as "Mary does-that Jones does-that his rent is paid" or "Mary does-that it is not the case that Jones does-that the rent is paid". (See e.g., [37].) The does-that modality can also be usefully combined with diverse other notions pertaining to agents and multiagent systems (MAS). In imperfect information settings, it allows to reason about epistemically uniform strategies when combined with a concept of *knowledge*: "Jones knows that he does-that the rent is paid". (See e.g., [23].) Combined with a concept of *obligation* the main application of Anselmian's actions probably resides in deontic logics and normative systems in general. While obligations alone allow to express "it is obligatory that the rent is paid", it is now possible to designate the subject of the obligation: "it is obligatory that Jones does-that the rent is paid". (See e.g., [35].)

The modality does-that has taken many forms in the literature. Chellas ([12]), von Kutschera ([29]), and Belnap and Perloff ([3,4]) have studied within elaborate models of branching-time the modality now well known as "seeing-to-it-that" (often referred as STIT). Belnap et al. have compiled in [4] and [24] many interpretations of the modality. Most of them support a game-theoretic interpretation of action: in a given situation, a group sees to it that something holds if this something would also hold no matter what the other agents do.

Kanger ([26]), Pörn ([37,38]), Lindahl ([31]), Elgesem ([15]), and others, on the other hand have studied the same sort of modality in a variety of frameworks, always abstracting away from time and from any game-theoretical interpretations. Nowadays, the term preferred to designate the modality is "bringing-it-about". The bringing-it-about modality, henceforth noted E_x, has been quite popular in the MAS community (e.g., [25,39,40,11,35]) where it has been used to model the actions and responsibilities of acting entities x: the formula $E_x\varphi$ traditionally reads "x brings it about that φ". In the MAS literature about bringing-it-about, x has been either an individual agent, a role in an institution, or an institutional agent. An institution can involve several agents, each playing a specific role in it. But institutions are not groups or coalitions.

It was mentioned in [11] to have x represent a set of agents. But only in much earlier work ([26,31]) was a group version of the bringing-it-about modality analyzed with some depth.[2] We borrow particularly from [31] in an early step of our analysis.

Elgesem's proposal has come up to be an authoritative modern reference to the logic of the modality of bringing-it-about.[3] On top of a thorough and penetrating analysis of agency, the proposal in [15] particularly distinguishes itself from the rest of the literature by studying the agency of individual agents as the goal-directed manifestation of an individual ability. This will be our starting point in this paper.

1.1 Contributions

The first contribution of this paper is an extrapolation of a theory of coalitional agency and ability from Elgesem's logic. That is, we study the logic of the operator E_x where x is a set of agents, along with an operator of coalitional ability C_x.

The second contribution is to provide algorithms to reason within these logics and study their complexity. Although we are using standard techniques of modal logics, this is, we believe, the first computational analysis of the logics of bringing-it-about.

1.2 Individual agency and ability

Elgesem's logic was a fresh look at a long tradition of philosophical logic of action, where the traditional modality of bringing-it-about noted E_x is studied alongside related modalities of action. The logic still admits the core principles that are generally assumed for agency (where N is the set of individual agents, and $x \in N$):

- all substitution instances of classical tautologies
- $\vdash \neg E_x \top$
- $\vdash E_x\varphi \wedge E_x\psi \to E_x(\varphi \wedge \psi)$
- $\vdash E_x\varphi \to \varphi$
- if $\vdash \varphi \leftrightarrow \psi$ then $\vdash E_x\varphi \leftrightarrow E_x\psi$

For any enumerable set of agents N, we can obtain a logic of coalitional agency and ability that is represented by the set of theorems determined by the previous Hilbert system. Let us note this logic BIAT^N.

Following Sommerhoff ([46]), Elgesem argues that agency is the actual bringing about of a goal towards which an activity is directed. An agent acts

[2] Maybe an exception is [34]. There, the agency of a coalition is defined as adequate combinations of individual agency of its members. The additional operators with coalitions are then merely syntactic sugar, leaving the expressivity of the logic unchanged. The logic also goes beyond the Anselmian tradition by allowing explicit action terms in the language.

[3] The main reference is [15] which unfortunately is not published. The reference [16] is essentially the Chapter 2 of [15].

to achieve a goal. But an agent is not necessarily aware of his goals, at least not in the sense that he is consciously committed to achieve them. Elgesem also leans on Frankfurt ([17, Chap. 6]) according to whom, the pertinent aspect of agency is the manifestation of the agent's guidance (or control) towards a goal; not necessarily the intentional action. Here, we understand intention in agency as a motivated goal, possibly long pondered and rational. Elgesem seeks a more general notion of goal that guides agency.

Elgesem observes that the manifestation of control is the exercise of a power to bring about something. Therefore, the notion of potential control of an agent for a goal should be integrated in a theory of agency. What is then particularly interesting of Elgesem's logic is that the modality of agency is studied alongside a modality to talk about this potential control: a modality of ability. Elgesem argues that we should not deny the possibility of abilities that are exercised only once, giving the example of Bob Beamon, who jumped 8.90 m (long jump) in the 1968 Olympics. If Beamon jumped that far it is that he was exercising control towards a goal. Even though this goal was probably not intentionally to jump 8.90 m, we would not take back from Beamon that on that day he brought about the fact that he jumped that far and that he had the ability to do it.

Elgesem then suggests that there is a more basic notion of ability than an intention-based one, and that this non-intentional notion of ability is a necessary condition for agency.[4] By bringing about something, an agent *shows* that he is indeed able to so. In this paper we advance an interpretation of *evidence-based* ability.

Elgesem's philosophy involves a net of notions related to agency but one may just focus on the two modalities E_x (agency) and C_x (ability) from which the others are derived. This is what Governatori and Rotolo do in [19]. They provide a criticism of Elgesem's logic by a careful logical analysis of these two operators. One of their main results is the completeness of the logic. Besides the core principles of agency listed above one needs to add the following to capture ability ($x \in N$):[5]

- the axiomatics of BIAT^N
- $\vdash \neg C_x \bot$
- $\vdash \neg C_x \top$
- $\vdash E_x \varphi \rightarrow C_x \varphi$
- if $\vdash \varphi \leftrightarrow \psi$ then $\vdash C_x \varphi \leftrightarrow C_x \psi$

For any enumerable set of agents N, let us note ELG^N the logic consisting of the set of theorems determined by the previous Hilbert system.

[4] Similar distinctions between different kinds of ability have later been made by Kapitan ([27]) and Mele ([32]). Mele calls them *simple ability to A* and *ability to A intentionally.* He writes: "an agent's A-ing at a time is sufficient for his having the simple ability to A at that time.", and "being able to A intentionally entails having a simple ability to A and the converse is false" ([32, p. 448]).

[5] Notice that $\neg E_x \top$ becomes redundant.

The logic of C_x is then rather weak.[6] The only certainty one can have about the presence of an ability to bring about φ is in the presence of an actual bringing about of φ. This is much different from the notion of ability that is captured by Coalition Logic ([36, 18]) or Alternating-time Temporal Logic ([2]). In Coalition Logic, an ability exists right before the action that would be the manifestation of the ability. G is able to bring about φ in Coalition Logic if G can select an action in his current repertoire that if chosen, would ensure φ at the next step, whatever the other agents do. Once the ability is realized, at the moment of the actual agency, nothing ensures that the ability still exists. This is also different from Kenny's view on ability ([28]) which is one where an agent has the ability to bring about φ if he could bring about φ whenever he tries. Hence, Kenny suggests that ability requires repeatability.

The notion of ability captured by Elgesem is nevertheless very appealing because it is one where *the observation of an evidence* induces the existence of an ability. Imagine a repository of web services that are acting in some way upon their environment and can be queried. Whenever a request is successfully fulfilled, the ability of a service for a particular query can be logged and the couple service/query can be offered as a *suggestion* for later use.

This *evidence-based* perspective of ability is strikingly weak in the individual case. We will present briefly some perspectives to refine further the notion with the integration of some form of induction in Section 6. Nevertheless, the main aim of this paper is to study a framework allowing coalitions of agents to work together towards a same goal. We will see that extending the logic to coalitions can offer more flexibility for the suggestion of potentially successful acting entities, even for *complex goals* that have never been brought about.

1.3 Joint actions

We will identify a group with an arbitrary subset of agents. Joint actions are a species of actions involving a group that acts towards a shared goal. Schmid calls these actions *plural*, with main characteristics "many participants, one goal" ([42]). Miller ([33]) says of a joint action that it involves two co-present agents each of whom performs simultaneously with the other agent one basic individual action, and in relation to a collective goal.

Despite resorting to the concept of a collective goal,[7] Miller argues that we-intentions ([49]) are not a necessary element of joint actions. When two scholars start chatting at a conference break and somewhat start to take a walk in the park, they respect their turn in the conversation, they synchronize their pace, and take a direction in the park without having previously agreed on it.[8] Similar to the individual case (Beamon's jump), this suggests that there is a more basic notion of coalitional goal-directed agency than an intentional one. Again in analogy with the individual case, that means that there is a

[6] See [52] for a similar and even weaker account of ability.

[7] Miller uses the terminology "collective end".

[8] This example is adapted from [5].

basic notion of coalitional ability that is a necessary condition for coalitional agency. In particular, at a given time and from the evidence of actual agency of some coalitions for some goals, we will be able to infer the potential ability of larger coalitions for more complex goals. To come back to our example of web services, this suggests a procedure for web service composition. At each time, we can use the logic to infer new complex coalitional abilities of the system. In turn, this evidence-based perspective may actually provide a practical alternative to the computationally costly orchestration procedures in web service composition.

Since there is a basic notion of coalitional agency, like Elgesem for individual agency and ability, we can therefore focus on the principles of pure agency and ability without having to struggle with the formation of we-intentions. In addition to their possible unawareness of acting towards a particular goal, agents will potentially be unaware of their membership in a group.

1.4 Outline

We will recognize the types of sufficient evidences to infer logically the existence of potential controls of groups, in a way that is consistent with Elgesem's philosophy. More generally, we will identify a variety of important principles pertaining to coalitional agency and ability and assess their admissibility. This is what we do in Section 2. Our presentation is first semantic, and in Section 3 we provide a complete syntactic characterization of the models of collective agency and ability. In Section 4 we study the complexity of the logics and provide algorithms for deciding the satisfiability of formulas.

In Section 5 we investigate possible strengthenings of our basic logic by adding application specific principles of agency. In Section 6 we refine the notion of ability by integrating some form of induction. Some considerations and further clarifications are presented in Section 7. Finally, we conclude in Section 8 and propose some perspectives.

2 Models of coalitional agency and ability

Effectivity functions are a mathematical tool used in social choice theory ([1]) to study the interaction of the powers of groups of agents. For an arbitrary acting entity, an *effectivity function* is a function $E : S \longrightarrow 2^{2^S}$ indicating at every state for what propositions (subsets of states) it is effective. There are often some common conditions on these functions but like in [36], we designate every function of this type as an effectivity function.

Effectivity functions have also been used in philosophical logic, as components of the so-called *neighborhood* or *minimal* models ([13]). In particular, Brown's modality of ability in [9] was already interpreted with an effectivity function.

Governatori and Rotolo ([19]) propose an alternative semantics to Elgesem's selection functions, where the models are equipped for every individual agent with an effectivity function for what the agent is able to do, but also with an effectivity function of what the agent is actually *agentive* for, that is, what the agent brings about. Taking this as a starting point we are going to extend these models with coalitions. One of the main objectives of this paper is to propose and justify adequate constraints on the effectivity functions of the models of coalitional agency and ability.

For now, the formula $E_G\varphi$ reads "acting as the group G, G is bringing about the goal φ", while the formula $C_G\varphi$ reads "acting as the group G, G is able to bring about the goal φ."

We assume a set of atomic facts P. A *model of agency and ability* is a tuple (S, N, EE, EC, v) where S is a set of worlds, and N is a set of agents. For every $w \in S$ and $G \subseteq N$, the objects $EE_w(G)$ and $EC_w(G)$ are effectivity functions. Hence, $EE_w(G)$ is a set of sets of worlds and so is $EC_w(G)$. If $X \in EE_w(G)$ we say that at w, acting as a coalition, the group G is bringing about their goal X. Analogously, if $X \in EC_w(G)$ we say that at w, acting as coalition, the group G is able to bring about the goal X. Finally, $v : S \longrightarrow 2^P$ is a valuation function, indicating for every world what atomic propositions are true in it.

To talk about these models we will use the language that is defined by the following BNF (where $p \in P$ and $G \subseteq N$):

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid E_G\varphi \mid C_G\varphi$$

For every agent $i \in N$, we will use the notations $E_i\varphi \stackrel{\mathtt{def}}{=} E_{\{i\}}\varphi$ and $C_i\varphi \stackrel{\mathtt{def}}{=} C_{\{i\}}\varphi$. We remind that $E_G\varphi$ stands for "coalition G brings it about that φ" and $C_G\varphi$ reads "coalition G is able to bring about that φ".

In a standard set-theoretic interpretation a proposition φ, or fact of the world is represented by the set of worlds where φ is true. The semantics of the formula φ in a model M is defined as $||\varphi||^M = \{w \mid M, w \models \varphi\}$. To give the meaning of our language it only remains to define what is $\models$:

$$\begin{array}{lll} M, w \models p & \text{iff} & p \in v(w) \\ M, w \models \neg\varphi & \text{iff} & \text{not } M, w \models \varphi \\ M, w \models \varphi \wedge \psi & \text{iff} & M, w \models \varphi \text{ and } M, w \models \psi \\ M, w \models E_G\varphi & \text{iff} & ||\varphi||^M \in EE_w(G) \\ M, w \models C_G\varphi & \text{iff} & ||\varphi||^M \in EC_w(G) \end{array}$$

Hence, at the state w of a model M, a group G brings about the goal φ when at w the group G is bringing about the set of states in M where φ holds. Analogously, at the state w, a group G is able to bring about the goal φ when at w the group G is able to bring about the set of states in M where φ holds.

We adopt the usual definitions for the constants $\bot$ and $\top$ and classical connectives $\vee$, $\rightarrow$, $\leftrightarrow$. The language and the models presented above are the most general that encompass all those we are going to study in this paper. Eventually, we will restrict the set G to be a singleton, and obtain the language and models for Elgesem's logic. A further restriction will eliminate the component EC and the operator C_x, and obtain the language and models for the standard logic of bringing-it-about.

For the moment, for every G our modalities E_G and C_G are the most general form of classical non-normal modalities of necessity ([44, 13]) and each modality is independent of the others. The remainder of this section presents the constraints that the models should satisfy in order to be useful for the study of coalitional agency and ability. That is, by specializing the models, we give to the language its most intuitive meaning.

2.1 Elementary constraints on the models

Rotolo and Governatori ([19]) have translated Elgesem's philosophy of *individual* agency and ability in the type of models that we have just introduced. While we extend the purpose of the framework to *coalitional* agency, we acknowledge that every constraint suggested by Elgesem for the individual case lifts naturally to the coalitional case. Here, we adopt straightforwardly the constraints for our models.

As usual in "bringing-it-about", we assume that no agent or coalition brings about the tautologies:

$$S \notin EC_w(G). \tag{1}$$

Also, we need to reflect that the elements of $EE_w(G)$ are the manifestation of a successful agency of the group G at w:

$$\text{if } X \in EE_w(G) \text{ then } w \in X. \tag{2}$$

Note that as a consequence, it is not possible for a group to bring about the impossible: if $X \in EE_w(G)$ then $X \neq \emptyset$, for w at least must be in X. However, as pointed out by Governatori and Rotolo, we also need to explicitly reject in these minimal models that one can exercise control towards the impossible. It yields the following constraint:

$$\emptyset \notin EC_w(G). \tag{3}$$

Two concomitant actions of the same agent or coalition are, in the terminology used by Elgesem, aggregating:

$$\begin{aligned}&\text{if } X_1 \in EE_w(G) \text{ and } X_2 \in EE_w(G) \text{ then}\\&(X_1 \cap X_2) \in EE_w(G).\end{aligned} \tag{4}$$

It means that when a coalition acts towards two goals at the same time, they also act towards the sum of the goals.

Finally, we need to model the relationship between agency and ability. By bringing about a goal, an individual or a coalition is just giving evidence of control over the goal. As such, actual agency implies ability.

$$EE_w(G) \subseteq EC_w(G). \tag{5}$$

By immediate evidence, a group is at least able to achieve what it is bringing about. Observe that together with Constraint 1, we have that

$$S \notin EE_w(G). \tag{1E}$$

It is a more standard principle of bring-it-about logics of agency that must be present in models that do talk about ability.

2.2 Constraints of coalitional agency and ability

So far, it is just like each coalition is nothing more than a fresh individual acting entity. We need to acknowledge in our models that a term G indeed represents a group of individuals.

In this section we argue that in a true coalitional setting, the specification of the models so far is insufficient. Therefore, we will propose to further constrain the models so as to capture an adequate notion of coalitional agency and ability that we believe is consistent with Elgesem's philosophy. Our methodology to finding the coalitional version of Elgesem's logic rather naïvely consists in thinking of a hypothetical principle and trying to show that it is not acceptable in some scenario. If no significant counterexample is found, we must accept it at that stage.

Our notions and our models are reminiscent of those studied in game and social choice theory. Therefore, we will spend a significant part of this section on explaining why we believe the usual constraints are not adequate in our setting, but that some ideas can be successfully imported in an adapted form.

2.3 Empty coalition

We first look at the empty group that is the simplest group, though degenerate. Our notion of agency is one that is goal-directed, and our notion of ability is one of potential control towards a goal. It would not be right to give to the empty group a status of true coalition with a goal and a potential control for it. We adopt:

$$EC_w(\emptyset) = \emptyset. \tag{6}$$

It naturally follows from Constraint 5 that $EE_w(\emptyset) = \emptyset$, too. That is, deprived of potential control, the empty coalition is not an agentive entity at all. Also, in the interest of clarity, observe that Constraint 6 is not in conflict with Constraint 3. Indeed, if the empty coalition cannot exercise control towards anything, then *a fortiori* it cannot exercise control towards the empty set.

On the other hand, agency and abilities of the grand coalition, that is the coalition N containing all the agents, do not obey a special set of constraints.

2.4 No coalition monotonicity

In [26], Kanger & Kanger explore the agency of parties and their parliamentary rights. They use what they call the *principle of joint parties*, stating that whenever a party P brings about a goal X, then every parliamentary group constituted of at least P also brings about X. In our formal semantics:

if $G_1 \subseteq G_2,$ and $X \in EE_w(G_1)$ then $X \in EE_w(G_2)$.

This is sometimes called *coalition monotonicty* in social choice theory ([36]) and it is a core principle in STIT theories ([24,8]). Kanger & Kanger do not strongly commit to this principle, and seem to adopt it as a way to simplify the exposition of their example. Instead, Lindahl ([31]) refutes the principle of joint parties in his framework. His argument appeals to a notion of *contribution* of an agent to the bringing about of a state of affairs. He argues that his conception of agency is such that when a group of agents G brings about a goal X, then every agent in G contributes to X. If the principle of joint parties were to hold it would imply that whenever an agent brings about X, then "anybody whosoever contributes to the bringing about" ([31, p. 225]) of X. Lindahl considers it unacceptable. We also consider his argument a sufficient reason to not adopt the principle as a general one.

2.5 Superadditivity

For an arbitrary effectivity function E consider the following property:

if $G_1 \cap G_2 = \emptyset, X_1 \in E_w(G_1)$ and $X_2 \in E_w(G_2)$ then
$(X_1 \cap X_2) \in E_w(G_1 \cup G_2)$.

If at w, G_1 is effective for X_1 and also at w, G_2 is effective for X_2, the constraint says that when G_1 and G_2 are disjoint groups one also have that at w, $G_1 \cup G_2$ is effective for $X_1 \cap X_2$.

It is a well-known constraint in social choice theory and is designated as *superadditivity.*

Effectivity functions were designed to characterize the powers of coalitions. Hence, the requirement that $G_1 \cap G_2 = \emptyset$ is fundamental. My power to be in Toulouse tomorrow at noon together with my power to be in Liverpool at noon do not imply that I have the power to be in both cities at noon.

2.5.1 No superadditivity of ability

Superadditivity is a foundational principle in Coalition Logic, a logic of powers of coalitions. It seems that superadditivity could also be compatible with the rather weak notion of ability that Elgesem tries to capture. He writes that "there is clearly no problem in assuming that someone can have abilities that are never exercised. A lion in the zoo, for example, is clearly able to catch zebras; it just never has the opportunity to do it." ([15, p. 36]). Hence, Elgesem's notion of ability does not rely on an opportunity to exercise a control, and instead idealizes the possibility of an opportunity. Ideally then for the coalitional setting, we could postulate that coalition formation can *a priori* always be achieved. Two independent groups who are able to bring about two goals are able to form one coalition (the union of the two existing coalitions) with one goal (the intersection of the two existing goals) and bring it about. In other words, abilities of smaller coalitions could be aggregated into abilities of larger coalitions. We would have a sophisticated way to discover potential control of coalitions over a goal.

But such a constraint would yield at least one very unfortunate consequence. In social choice theory, we say that the powers of coalitions are *regular* when for every group G, if G is effective for a proposition (a subset of worlds) then the complement of G (the set $\overline{G}$ of all the agents that are not members of G) is not a group effective for the negation of the proposition.

We argue that this is unacceptable for a notion of ability as a potential control that might never be exercised if the opportunity is never offered. Indeed, as Elgesem's notion of ability is independent of opportunity, it should be possible to concomitantly have an ability to bring about something, and an ability to bring about its contrary. More so, this should be even true for two different acting entities being deemed able to bring about two contradictory state of affairs. Yet, superadditivity of ability would imply regularity of ability. To see this, let $X \in EC_w(G_1)$ and assume for *reductio* that $\overline{X} \in EC_w(G_2)$ and $G_1 \cap G_2 = \emptyset$. By superadditivity we obtain that $X \cap \overline{X} \in EC_w(G_1 \cup G_2)$, that is $\emptyset \in EC_w(G_1 \cup G_2)$, a contradiction with Constraint 3.

Hence, if a group G_1 has a potential control over X it would be impossible for an independent group G_2 to have potential control over $\overline{X}$. To put it bluntly, superadditivity of ability would mean that as soon that Ann is potentially able to put the lights on, Bill could not be deemed able to put the lights off. It is not a general principle that we want to have for a basis to model multiagent systems. The type of ability we are concerned with is thus of much different nature than the one captured by Coalition Logic.

2.5.2 No superadditivity of agency

If we do not think that superadditivity is an adequate assumption for our notion of ability, we do not think that agency should obey a similar principle either. What would be the consequences, for our idea of agency, of the constraint if $G_1 \cap G_2 = \emptyset, X_1 \in EE_w(G_1)$ and $X_2 \in EE_w(G_2)$ then $(X_1 \cap X_2) \in EE_w(G_1 \cup G_2)$? At w, G_1 is bringing about the goal X_1 while also at w, G_2 is bringing about the goal X_2. When G_1 and G_2 are disjoint groups one could deduce in these models that at w, $G_1 \cup G_2$ is bringing about the goal $X_1 \cap X_2$. It does not seem right. G_1 and G_2 are bringing about at w a goal of their own. It would be presumptuous to say now that the two goals have just been achieved, that their sum is also a goal; especially a goal of a coalition that does not necessarily exist as such at the moment.

We assumed G_1 and G_2 to be non-overlapping groups and hence independent. How about when they are not?

When one group is strictly included in the other and we have $X_1 \in EE_w(G_1)$ and $X_2 \in EE_w(G_1 \cup G_2)$. In words, G_1 are acting to obtain their goal X_1, and in addition the members of G_1 are also acting together with the bigger group $G_1 \cup G_2$ to bring about X_2, a goal of $G_1 \cup G_2$. Should we be able to deduce something more about the agentivity of some coalition? It is easy to argue for a negative answer. If Page and Plant (G_1) write together "Immigrant Song" (X_1), and Page, Plant and Jones ($G_1 \cup G_2$) write together "Black Dog" (X_2), Jones (G_2) has nothing to do with the bringing about of X_1. It is not right, then, to say that Page, Plant and Jones have written these two songs together. So we should not have in general to infer that $(X_1 \cap X_2) \in EE_w(G_1 \cup G_2)$. From another standpoint, Page and Plant, as G_1, are not either bringing about their own goal to have these two songs written as this would undermine the contribution of Jones in the writing of "Black Dog". Thus, we should not in general infer that $(X_1 \cap X_2) \in EE_w(G_1)$. Even though it might present an appropriate picture in some situations, it is clearly not a general principle of agency.

On the other hand, when they are in fact the same group G, it is established at the moment of agency that G are indeed a coalition and are achieving both X_1 and X_2. Thus, following the tradition of bringing-it-about, we consider that they are also achieving $X_1 \cap X_2$ as the agglomerated goal of the same group. It is a principle that we want, but this case falls under the case of Constraint 4. We do not need to constrain the models further.

2.5.3 Superadditivity, kind of

However, our evidence-based perspective of ability suggests something somewhat in-between superadditivity of ability and superadditivity of agency.

$$\begin{array}{l}\text{if } X_1 \in EE_w(G_1) \text{ and } X_2 \in EE_w(G_2) \text{ then} \\ (X_1 \cap X_2) \in EC_w(G_1 \cup G_2).\end{array} \tag{7}$$

Notice that it involves both agency and ability. When two groups G_1 and G_2 successfully but independently bring about two goals X_1 and X_2, had they acted as the coalition $G_1 \cup G_2$ they would have collectively brought about $X_1 \cap X_2$. Constraint 7 acknowledges that this is enough evidence for the ability of the coalition $G_1 \cup G_2$ to bring about $X_1 \cap X_2$.

Quite remarkably, the condition $G_1 \cap G_2 = \emptyset$ of superadditivity is not necessary anymore. By their actual and concomitant agency, the groups G_1 and G_2 have shown that for every shared agent $a \in G_1 \cap G_2$, the action of a towards X_1 and the action of a towards X_2 are not in conflict.

It is readily seen that Constraint 7 is a generalization of Constraint 5. (It suffices to take $G_1 = G_2 = G$ and $X_1 = X_2 = X$.) As a further justification that we are heading towards the right direction in extending Elgesem, it is important to note that a similar principle of aggregation already exists in the standard logics of bringing-it-about. By Constraint 4 we have indeed that two goals X_1 and X_2 brought about at the same time by the *same acting entity* aggregate in a goal $X_1 \cap X_2$ that is brought about. In turn, this bringing about implies the existence of an ability of the agent for $X_1 \cap X_2$ by Constraint 5. We have merely generalized this inference to the case of *two possibly different acting entities.*

It is a powerful formal device for our theory of evidence-based ability since it allows to deduce potential abilities of coalitions of agents from smaller "successes" in the society of agents. We can use the information of actual agency and suggest that the group of agents $G_1 \cup G_2$ could potentially be solicited to bring about the goal $X_1 \cap X_2$.

3 Axiomatizations

Let us now define a family of classes of models. We are going to characterize them syntactically. This will provide us two equivalent ways to look at the logics: one semantic, one syntactic. This characterization is instrumental for establishing the upper-bound on the logics' computational complexity in Section 4.

Definition 1 We say that a tuple (S, N, EE, v) is a *model of individual agency* if it satisfies the Constraints 1E, 2 and 4.

We say that a tuple (S, N, EE, EC, v) is a *model of individual agency and ability* if it satisfies the Constraints 1 through 5.

We say that a tuple (S, N, EE, EC, v) is a *model of coalitional agency and ability* if it satisfies all the Constraints 1 through 7.

Naturally, the models of individual agency correspond to the core models of the logic bringing-it-about. The models of individual agency and ability correspond to the models of Elgesem's logic. The following result has already been established in [19]:

Theorem 1 *For a set of agents N:*

- *the logic BIAT^N is sound and complete wrt. the class of models of individual agency;*
- *the logic ELG^N is sound and complete wrt. the class of models of individual agency and ability.*

We still have to characterize the models of coalitional agency and ability syntactically.

For all groups G, G_1, and G_2 and formulas φ and ψ:

- **[Ax0]** $\vdash \varphi$, when φ is a tautology in propositional logic
- **[Ax1]** $\vdash E_G\varphi \wedge E_G\psi \to E_G(\varphi \wedge \psi)$
- **[Ax2]** $\vdash E_G\varphi \to \varphi$
- **[Ax3]** $\vdash E_G\varphi \to C_G\varphi$
- **[Ax4]** $\vdash \neg C_G\bot$
- **[Ax5]** $\vdash \neg C_G\top$
- **[Ax6]** $\vdash \neg C_\emptyset\varphi$

- **[Ax7]** $\vdash E_{G_1}\varphi \wedge E_{G_2}\psi \to C_{G_1\cup G_2}(\varphi \wedge \psi)$
- **[ERE]** if $\vdash \varphi \leftrightarrow \psi$ then $\vdash E_G\varphi \leftrightarrow E_G\psi$
- **[ERC]** if $\vdash \varphi \leftrightarrow \psi$ then $\vdash C_G\varphi \leftrightarrow C_G\psi$

For any finite set of agents N, we can obtain a logic of coalitional agency and ability that is represented by the set of theorems determined by the previous Hilbert system.[9] Let us note this logic COAL^N.[10] We will sometimes just note it COAL.

Theorem 2 *For a set of agents N, the logic COAL^N is sound and complete with respect to the class of models of coalitional agency and ability.*

Proof Soundness is readily checked. For completeness we start by defining a canonical model. The canonical model $M^c = (S^c, N^c, EE^c, EC^c, v^c)$ is defined as follows:

- S^c is the set of maximally COAL^N-consistent sets;
- $N^c = N$;
- For every $\Sigma \in S^c$: $EE^c_\Sigma(G) = \{|\varphi|^{M^c} \mid E_G\varphi \in \Sigma\}$;
- For every $\Sigma \in S^c$: $EC^c_\Sigma(G) = \{|\varphi|^{M^c} \mid C_G\varphi \in \Sigma\}$;
- $\Sigma \in v^c(p)$ iff $p \in \Sigma$.

where $|\varphi|^{M^c} = \{\Sigma \in S^c \mid \varphi \in \Sigma\}$.

The Truth Lemma states that for every φ and $\Sigma \in S^c$ we have $M^c, \Sigma \models \varphi$ iff $\varphi \in \Sigma$. Equivalently $||\varphi||^{M^c} = |\varphi|^{M^c}$. It can be proved by standard induction. In particular, $M^c, \Sigma \models E_G\psi$ iff $||\psi||^{M^c} \in EE^c_\Sigma(G)$ iff $|\psi|^{M^c} \in EE^c_\Sigma(G)$ (by induction hypothesis) iff $E_G\psi \in \Sigma$. The similar argument works for the induction case of $C_G\psi$. Atomic formulas and formulas whose main operator is a classical logical connective are trivial cases.

Let us show that EE^c and EC^c are well-defined. Suppose $|\varphi|^{M^c} = |\psi|^{M^c}$. Then $\vdash \varphi \leftrightarrow \psi$ and by rule ERE we have for every $G \subseteq N$ that $\vdash E_G\varphi \leftrightarrow E_G\psi$. Similarly by rule ERC we have for every $G \subseteq N$ that $\vdash C_G\varphi \leftrightarrow C_G\psi$. This means for all $\Sigma \in S^c$, that $E_G\varphi \in \Sigma$ iff $E_G\psi \in \Sigma$, and that $C_G\varphi \in \Sigma$ iff $C_G\psi \in \Sigma$.

We need to prove that M^c is indeed a model of coalitional agency and ability. This is a simple task. We present the argument for Constraint 6 and Constraint 7.

Constraint 6: Let $\Sigma \in S^c$. From axiom Ax6 we know that $\neg C_\emptyset\varphi \in \Sigma$ for all φ. So there is no φ such $|\varphi|^{M^c} \in EC^c_\Sigma(\emptyset)$. Hence, by definition, $EC^c_\Sigma(\emptyset) = \emptyset$.

Constraint 7: For some $\Sigma \in S^c$, G_1 and G_2, let $X_1 \in EE_\Sigma(G_1)$ and $X_2 \in EE_\Sigma(G_2)$. So there is a formula φ_1 such that $E_{G_1}\varphi \in \Sigma$ and a formula

[9] Observe that, as pointed out in Section 2.5 for Constraint 5 and Constraint 7, axiom Ax3 is redundant in presence of axiom Ax7. Given its importance in ELG^N, we conserve it in this axiomatization.

[10] Contrarily to BIAT^N and ELG^N, we need N to be finite. Our language refers to coalitions of agents that are subsets of N, and the set of subsets of an enumerable infinite set is not enumerable. Hence, a logic COAL^N with a possibly infinite enumerable set N of agents would not be finitely axiomatizable.

φ_2 such that $E_{G_2}\varphi \in \Sigma$, and $X_1 = |\varphi_1|^{M^c}$ and $X_2 = |\varphi_2|^{M^c}$. By axiom Ax7 we obtain $C_{G_1 \cup G_2}(\varphi_1 \wedge \varphi_2) \in \Sigma$. Hence $|\varphi_1|^{M^c} \cap |\varphi_2|^{M^c} \in EC^c_{\Sigma}(G_1 \cup G_2)$. We conclude that $(X_1 \cap X_2) \in EC_{\Sigma}(G_1 \cup G_2)$, which proves that M^c satisfies Constraint 7.

Hence, M^c is a model of coalitional agency and ability. Moreover, by definition, for every consistent formula φ there is a state in S^c that contains φ. By the Truth Lemma, we conclude that for every consistent formula there is model of coalitional agency and ability that satisfies it.

4 Algorithms for reasoning about agency and ability

In this section, we are going to devise algorithms to reason about agency and ability, and analyze their computational complexity. More precisely for a logic $\mathcal{L} \in \{\mathsf{BIAT}, \mathsf{ELG}, \mathsf{COAL}\}$ we want to solve and evaluate the complexity of the $\mathcal{L}$-sat decision problem.

Definition 2 The $\mathcal{L}$-sat decision problem is defined as follows:
input: a formula φ in the language of $\mathcal{L}$;
output: false if $\neg\varphi$ is a theorem of $\mathcal{L}$; true otherwise.

Our proof of correctness of the algorithms and of their computational complexity is going to be semantic. Hence, we are going to use the semantics defined in Section 2 and our determination results of Section 3. Figuring out whether a formula $\neg\varphi$ is not a theorem (φ is consistent) will translate into finding that a model in the class determining $\mathcal{L}$ satisfies φ.

As in [51], we provide a variant of "tree-less" tableaux (Section 4.1), that can also be seen as SAT-based procedures ([43]). This will provide a decision procedure to the problem of determining whether a formula is satisfiable. The analysis of the algorithm will give us an upper-bound for the complexity of the decision problem.

Remark 1 SAT-based procedures have significant appeal. Most of the computing work to determine the satisfiability of a modal formula is delegated to a classical SAT solver for propositional logic. Any SAT solver can be plugged in. In general, one can use any good off-the-shelf SAT solver. However, some SAT solvers are optimized to be faster on some restricted instances of propositional formulas. It is then possible to tweak the SAT-based procedure by heuristically selecting a SAT solver that is specialized for the kind of instance at hand[11], paving the way for efficient automated reasoning within the logics of bringing-it-about.

We present the complete analysis for the case of the logic BIAT in Section 4.1. In Section 4.2 we simply extend the algorithm for BIAT-sat to ELG-sat and to COAL-sat.

[11] See for instance the results of the last SAT competition `http://www.satcompetition.org/`.

4.1 Algorithm for BIAT-sat

In this section we prove that the problem of satisfiability checking within the logic of BIAT is in PSPACE. Our proof is adapted from [51].

If φ is a formula, then $sub^{\neg}(\varphi)$ is the set of all subformulas of φ and their negations (we identify the formula $\neg\neg\psi$ with ψ). A *semi-valuation* for φ is a function $\pi : sub^{\neg}(\varphi) \longrightarrow \{0, 1\}$ that satisfies the following conditions:

- $\pi(\psi) = 1$ iff $\pi(\neg\psi) = 0$;
- $\pi(\psi_1 \wedge \psi_2) = 1$ iff $\pi(\psi_1) = 1$ and $\pi(\psi_2) = 1$;
- $\pi(\varphi) = 1$.

The following theorem is at the center of the analysis. We relegated the proof to the Annex; It is rather involved but the techniques are not novel.

Theorem 3 *φ is BIAT-sat iff there is a semi-valuation π for φ such that:*

1. *if $E_a\varphi \in sub^{\neg}(\varphi)$ and $\pi(E_a\varphi) = 1$, then $\pi(\varphi) = 1$;*
2. *if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$ then $\neg\psi$ is BIAT-sat;*
3. *if $E_a\psi_1, \ldots, E_a\psi_k, E_a\psi \in sub^{\neg}(\varphi)$, with $\pi(E_a\psi_j) = 1$ for all j, and $\pi(E_a\psi) = 0$, then the formula*

$$\left(\bigwedge_j (\psi_j) \wedge \neg\psi\right) \vee \left(\bigvee_j (\neg\psi_j) \wedge \psi\right)$$

is BIAT-sat.

Now that we have characterized the satisfiability of formulas in terms of the satisfiability of its sub-formulas, devising an algorithm for the decision problem BIAT-sat is just straightforward. It justifies the following result about the upper-bound on the complexity of satisfiability within BIAT.

Corollary 1 *Checking whether a formula φ is BIAT-sat can be done using space polynomial in the size of φ.*

Proof Theorem 1 establishes that a formula $\neg\varphi$ is not a theorem iff φ is satisfiable. To determine the satisfiability of a formula φ, consider the algorithm in Figure 1. The correctness of the algorithm wrt. the decision problem BIAT-sat is a direct consequence of Theorem 3.

Every recursive call is done on formulas with a strictly decreasing modal depth. It means that the algorithm terminates. It also means that this is an alternating polynomial-time algorithm. It implies that there exists an equivalent deterministic polynomial space algorithm (ATIME = PSPACE).

1. non-deterministically guess a semi-valuation π for φ;
2. if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$ then check that $\pi(\psi) = 1$;
3. if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$, recursively check that $\neg\psi$ is BIAT-sat;
4. if $E_a\psi_1, \ldots, E_a\psi_k, E_a\psi \in sub^{\neg}(\varphi)$, with $\pi(E_a\psi_j) = 1$ for all j, and $\pi(E_a\psi) = 0$, then non-deterministically and recursively check that either:
   - $\bigwedge_j(\psi_j) \wedge \neg\psi$ is BIAT-sat;
   - $(\neg\psi_1) \wedge \psi$ is BIAT-sat;
   - ...
   - $(\neg\psi_k) \wedge \psi$ is BIAT-sat.

Fig. 1 Algorithm for solving BIAT-sat.

1. non-deterministically guess a semi-valuation π for φ;
2. if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$, then check that $\pi(\psi) = 1$;
3. if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$, recursively check that $\neg\psi$ is ELG-sat;
4. if $C_a\psi \in sub^{\neg}(\varphi)$ and $\pi(C_a\psi) = 1$, recursively check that both:
   - $\neg\psi$ is ELG-sat;
   - ψ is ELG-sat;
5. if $E_a\psi_1, C_a\psi_2 \in sub^{\neg}(\varphi)$, with $\pi(E_a\psi_1) = 1$ and $\pi(C_a\psi_2) = 0$ then recursively check that $(\psi_1 \wedge \neg\psi_2) \vee (\neg\psi_1 \wedge \psi_2)$ is ELG-sat;
6. if $E_a\psi_1, \ldots, E_a\psi_k, E_a\psi \in sub^{\neg}(\varphi)$, with $\pi(E_a\psi_j) = 1$ for all j, and $\pi(E_a\psi) = 0$, then non-deterministically and recursively check that either:
   - $\bigwedge_j(\psi_j) \wedge \neg\psi$ is ELG-sat
   - $(\neg\psi_1) \wedge \psi$ is ELG-sat;
   - ...
   - $(\neg\psi_k) \wedge \psi$ is ELG-sat.

Fig. 2 Algorithm for solving ELG-sat.

4.2 Algorithms for ELG-sat and COAL-sat

The algorithm in Figure 2 is to decide whether a formula φ is satisfiable (ELG-sat). The algorithm in Figure 3 is to decide whether a formula φ is satisfiable (COAL-sat).

The same analysis as the one for BIAT can be done for ELG and COAL. It can be proved that the algorithms are correct and that they allow us to reason about the satisfiability of BIAT and of COAL in polynomial space.

5 Refining agency: delegations

So far, we believe that the principles of agency that we proposed are adequate in every situation of coalitional agency. The logic COAL is what we consider the minimal logic of coalitional goal-directed agency and evidence-based ability. That is, the logic that minimally extends Elgesem's logic to the case of coalitional agency. In this section, we introduce more principles of social interaction. Ever since Chellas ([12]), the notion of delegation (or influence) has been a major domain for applying the logics of bringing-it-about. (See also,

1. non-deterministically guess a semi-valuation π for φ;
2. if $E_\emptyset\psi \in sub^\neg(\psi)$, then check that $\pi(E_\emptyset\psi) = 0$;
3. if $C_\emptyset\psi \in sub^\neg(\psi)$, then check that $\pi(C_\emptyset\psi) = 0$;
4. if $E_G\psi \in sub^\neg(\varphi)$ and $\pi(E_G\psi) = 1$, then check that $\pi(\psi) = 1$;
5. if $E_G\psi \in sub^\neg(\varphi)$ and $\pi(E_G\psi) = 1$, recursively check that $\neg\psi$ is COAL-sat;
6. if $C_G\psi \in sub^\neg(\varphi)$ and $\pi(C_G\psi) = 1$, recursively check that both:
   - $\neg\psi$ is COAL-sat;
   - ψ is COAL-sat;
7. if either:
   - $E_G\psi_1, \ldots, E_G\psi_k, E_G\psi \in sub^\neg(\varphi)$, with $\pi(E_G\psi_j) = 1$ for all j, and $\pi(E_G\psi) = 0$;
   - $E_{G_1}\psi_1, \ldots, E_{G_k}\psi_k, C_G\psi \in sub^\neg(\varphi)$, with $\pi(E_{G_j}\psi_j) = 1$ for all j, $\pi(C_G\psi) = 0$, and $G = G_1 \cup \ldots \cup G_k$;

   then non-deterministically and recursively check that either:
   - $\bigwedge_j(\psi_j) \wedge \neg\psi$ is COAL-sat
   - $(\neg\psi_1) \wedge \psi$ is COAL-sat;
   - ...
   - $(\neg\psi_k) \wedge \psi$ is COAL-sat.

Fig. 3 Algorithm for solving COAL-sat.

e.g., [16,34].) We then choose to investigate the mechanisms of delegation between agents and coalitions.

In Section 5.1, we study Chellas's constraint forcing the full responsibility of a delegating entity for the result brought about by the delegatee under his influence. Because it is not inherently pertaining to coalitional agency, we analyze it alongside a common assumption of logics of agency, namely *strict-joint agency.* In Section 5.2 we propose a new principle forcing that the responsibility of a delegator is shared with the delegatee.

The various principles that we offer in this section are not meant to be integrated in COAL. We do not acknowledge them as general principles of coalitional agency. Instead, we use the logical machinery to analyze their consequences, and let a modeler judge whether they should be adopted with an application at hand.

5.1 Two common principles

5.1.1 Strict-joint agency

It seems right in some situations that if G_1 acting as a coalition is bringing about a goal X, then all the members of G_1 are actually contributing in some way to X. (Recall Lindahl's argument against coalition monotonicity in Section 2.4.) They might be necessary for the performance of a bodily movement, or they might be necessary for the group attitude that is put into the goal X.[12] We will use the term *team* to capture this. The members of the team G_1, each with a group attitude towards the coalition G_1 with regard to X, when $G_2 \subset G_1$ the group G_2 cannot be agentive for X in the same way. At

[12] We understand group attitude only loosely in the sense of [49].

least the group G_2 is not bringing about the goal X as a team. This is what is captured by the following constraint:

$$\text{when } G_2 \subset G_1, \text{ if } X \in EE_w(G_1) \text{ then } X \notin EE_w(G_2). \quad (8)$$

Note that the contrapositive of the constraint is equivalent to: when $G_1 \subset G_2$, if $X \in EE_w(G_1)$ then $X \notin EE_w(G_2)$. This corresponds to the notion of *strict joint agency* in the theory of seeing-to-it-that ([4, 10]).

In Section 2.4, we discussed Kanger & Kanger's principle of joint parties, or coalition monotonicity, stating that whenever a group G brings about a goal X, then every group constituted of at least G also brings about X. We rejected it as a general principle. Strict joint agency and coalition monotonicity are clearly two incompatible ways of looking at agency. Strict joint agency makes the most sense when one considers that coalitions have to form.[13] Individual agents have incentives to participate in a coalition and a coalition has incentives to accept a new member and form a larger coalition. A free-rider is usually not welcome.

From an external point of view, that is from the perspective of a system engineer, this is a formal device that allows to optimize the use of the resources (the agents) of the system. If one needs a goal X to be achieved and the group G is able to bring it about, it would be a waste to involve a larger group containing all the members of G. Therefore, the agency of a goal should be attributed to the only group that is sufficient to bring about the goal and that contains the agents that are necessary for the goal to be brought about. Nevertheless, bear in mind that this does not rule out the existence of two or more distinct groups that bring about the same goal; even groups that share some members.

5.1.2 Responsibility of the delegator

Chellas ([12]) in his formalism of imperative sentences does look at principles of interaction between agents. He argues that like in the common law maxim "quid facit per alium facit per se", we should have that when agent a makes another agent bring about something, agent a is himself bringing about that something. Lifted up to groups of agents, that would translate in our models as:

$$\text{if } \{v \mid X \in EE_v(G_2)\} \in EE_w(G_1) \text{ then } X \in EE_w(G_1). \quad (9)$$

With this constraint, we accept a principle that the goal plus the means entails the goal: if the goal of a coalition G_1 is that another coalition brings about X, then we acknowledge that G_1 has the goal that X. Hence, coalitions are responsible for what they bring about, even if they do so through another acting entity.

Elgesem promptly rejected this constraint, explaining that "a person is normally not considered the agent of some consequence of his action if another agent interferes in the causal chain." ([15, p. 82]) . Santos et al.'s logic in [40]

[13] See for instance [45].

is equipped with two notions of agency: direct and indirect agency. They explicitly state that if agent a *directly* brings about that b *directly* brings about that φ then a *does not directly* bring about the φ. They nevertheless adopt the principle for *indirect* actions.

We did not commit to any direct or indirect interpretation of action in our framework. What we assumed however in this section, is that the agentive coalitions have some sort of identity; they are acting together as a team towards a shared goal.

What we believe will be interesting to observe, are the possible mechanisms of delegation, or lack thereof, between and within groups in this setting. When "making do" is the realization of a delegation, this suggests the presence of organized groups more on par with an *institution* than with a coalition. When "making do" is the realization of a command or of a strain on the actions of an agent or a group of agents, this means that there is some sort of authority between the two acting entities. This suggests that delegator and delegatee are not acting together as a team towards a shared goal. We will come back to this issue in Section 5.1.4 after presenting the syntactic characterization of the new models.

5.1.3 Completeness

Let us note Λ_1^N the logic obtained from combining the axioms of COAL^N with the axioms:

- **[Ax8]** $\vdash E_{G_1}\varphi \rightarrow \neg E_{G_2}\varphi$, when $G_2 \subset G_1$
- **[Ax9]** $\vdash E_{G_1}E_{G_2}\varphi \rightarrow E_{G_1}\varphi$

Theorem 4 *The logic Λ_1^N is sound and complete with respect to the models of coalitional agency and ability satisfying Constraint 8 and Constraint 9.*

Proof The proof extends the one of Theorem 2. Assume that M^c is now the canonical model built from Λ_1^N-mcs. Soundness of axiom Ax9 and axiom Ax8 is straightforward. For the completeness, we need to check that M^c satisfies the additional constraints.

Constraint 8: For some $\Sigma \in S^c$, G_1 and G_2 such that $G_2 \subset G_1$, let $X \in EE_\Sigma(G_1)$. So there is a formula φ such that $X = |\varphi|^{M^c}$ and $E_{G_1}\varphi \in \Sigma$. Then by axiom Ax8 we also have $\neg E_{G_2}\varphi \in \Sigma$. It means that $E_{G_2}\varphi \notin \Sigma$ and that $|\varphi|^{M^c} = X \notin EE^c_\Sigma(G_2)$. This proves that M^c satisfies Constraint 8.

Constraint 9: Let $\{\Gamma \mid X \in EE_\Gamma(G_2)\} \in EE_\Sigma(G_1)$ for some $\Sigma \in S^c$. So there is a formula φ such that $|\varphi|^{M^c} = X$ such that $\{\Gamma \mid E_{G_2}\varphi \in \Gamma\} \in EE_\Sigma(G_1)$. This is equivalent to $|E_{G_2}\varphi|^{M^c} \in EE_\Sigma(G_1)$. Hence $E_{G_1}E_{G_2}\varphi \in \Sigma$. By axiom Ax9, we also have that $E_{G_1}\varphi \in \Sigma$. This means that $X \in EE_\Sigma(G_1)$. This proves that M^c satisfies Constraint 9.

Remark 2 The reader more familiar with normal modal logics (roughly, logics over Kripke models) might wonder why Ax2 does not already imply Ax9. This is because our logics are non-monotonic. That is, the *rule of monotony*

(RM) does not preserve validities; It is not the case that if $\vdash \varphi \to \psi$ then $\vdash E_G\varphi \to E_G\psi$. Together, Ax2 and RM would imply Ax9.

5.1.4 Impossible intra-team commands

We investigate what are the consequences of the logic Λ_1^N for the relationship between commands and coalitional agency. We claim that it captures a notion of action of coalitions where the members of a coalition do identify with a team for in a collectively aware manner for the bringing about of a goal.

Consider the following derivation in Λ_1^N:

1. $\Lambda_1^N \vdash E_a\varphi \to \neg E_{\{a,b\}}\varphi$ (instance of axiom Ax8)
2. $\Lambda_1^N \vdash E_aE_b\varphi \to E_a\varphi$ (instance of axiom Ax9)
3. $\Lambda_1^N \vdash E_aE_b\varphi \to \neg E_{\{a,b\}}\varphi$ (from 1. and 2. by Propositional Logic)
4. $\Lambda_1^N \vdash E_aE_b\varphi \to E_b\varphi$ (instance of axiom Ax2)
5. $\Lambda_1^N \vdash E_aE_b\varphi \to E_a\varphi \wedge E_b\varphi$ (from 2. and 4. by PL)
6. $\Lambda_1^N \vdash E_a\varphi \wedge E_b\varphi \to C_{\{a,b\}}\varphi$ (instance of axiom Ax7)
7. $\Lambda_1^N \vdash E_aE_b\varphi \to \neg E_{\{a,b\}}\varphi \wedge C_{\{a,b\}}\varphi$ (from 3., 5. and 6. by PL)

So from item 3, we have that when something is achieved by a after a command of b, the group $\{a, b\}$ is not acting as a coalition. In fact, as item 2 states it, it is in the first place a bringing about of agent a. Item 7 on the other hand shows that a group might not be bringing about a goal, its members are nevertheless presently showing sufficient evidence to infer that it is able to do so.

Command or delegation within a group yields different results. The following formulas are theorems of Λ_1^N:

1. $\Lambda_1^N \vdash E_{\{a,b\}}E_b\varphi \to E_{\{a,b\}}\varphi$ (instance of axiom Ax9)
2. $\Lambda_1^N \vdash E_{\{a,b\}}E_b\varphi \to E_b\varphi$ (instance of axiom Ax2)
3. $\Lambda_1^N \vdash E_b\varphi \to \neg E_{\{a,b\}}\varphi$ (instance of axiom Ax8)
4. $\Lambda_1^N \vdash E_{\{a,b\}}E_b\varphi \to E_{\{a,b\}}\varphi \wedge \neg E_{\{a,b\}}\varphi$ (from 1., 2., and 3. by PL)
5. $\Lambda_1^N \vdash \neg E_{\{a,b\}}E_b\varphi$ (from 4. by PL)

Analogously:

1. $\Lambda_1^N \vdash E_bE_{\{a,b\}}\varphi \to E_{\{a,b\}}\varphi$ (instance of axiom Ax2)
2. $\Lambda_1^N \vdash E_bE_{\{a,b\}}\varphi \to E_b\varphi$ (instance of axiom Ax9)
3. $\Lambda_1^N \vdash E_b\varphi \to \neg E_{\{a,b\}}\varphi$ (instance of axiom Ax8)
4. $\Lambda_1^N \vdash E_bE_{\{a,b\}}\varphi \to E_{\{a,b\}}\varphi \wedge \neg E_{\{a,b\}}\varphi$ (from 1., 2., and 3. by PL)
5. $\Lambda_1^N \vdash \neg E_bE_{\{a,b\}}\varphi$ (from 4. by PL)

The two previous derivations establish that a command or delegation by a group towards a subgroup or towards a super-group is never effective. The point is that with the constraint of strict agency a coalition G is acting together as a team to achieve a goal of theirs. This seems in accordance with the idea that if the coalition has to split before the goal to be obtained it does not act as a team anyway. Therefore, when acting as a coalition, there is no sense for a group to command something to a subgroup or a super-group. The group only and as a whole must be solicited.

Remark 3 Of course one might object that analogous delegations happen all the time in societies. For instance when a company delegates a signing of a contract to the person holding the role of CEO in the company. However, the company is *not* a coalition, nor a team. It is not a group of agents, and the person signing the contract is not part of the company but plays a role in it, instead. Our theory of coalitional agency does not explain institutions, roles, and institutional action. Maybe the best way to capture an institution and a role is then to see them as fresh new individual agents. Therefore, we would have that $E_I E_{CEO}\varphi \to E_I\varphi \wedge E_{CEO}\varphi$ where the formula $E_I\varphi \wedge E_{CEO}\varphi$ is perfectly consistent. Since a role or an institution are ontologically very peculiar kinds of agents, we would have to give a special meaning to institutional agents at the level of the semantics. It is not our objective here. This has been investigated in a series of papers by Pacheco and others (e.g., [11,35]).

5.2 Shared responsibility of the delegating entity

Although, the logic Λ_1^N might be adequate in some situations, other strengthenings of COAL are possible.

Consider the following constraint:

$$\text{if } \{v \mid X \in EE_v(G_2)\} \in EE_w(G_1) \text{ then } X \in EE_w(G_1 \cup G_2). \tag{10}$$

Let us note Λ_2^N the logic obtained from combining the axioms of COAL^N with the axiom:

– **[Ax10]** $\vdash E_{G_1}E_{G_2}\varphi \to E_{G_1 \cup G_2}\varphi$

We state without proof[14] the following theorem:

Theorem 5 *The logic Λ_2^N is sound and complete with respect to the class of models of coalitional agency and ability satisfying Constraint 10.*

Constraint 10 reflects that if the group G_1 makes G_2 bring about that something holds, the combined coalition $G_1 \cup G_2$ is responsible for it.

This is a principle that could not be expressed for the previous logics of bringing-it-about, but that our new language with coalitions now allows to formulate in a logic of bringing-it-about. Moving from individual agency to coalitional agency might then be useful with regard to ending a long argument in the literature. Indeed, it appears as a compromise to the controversial Constraint 9.

The following is a very simple derivation in our new system that starts to illustrate the consequences of adopting this form of responsibility of the delegating entity:

1. $\Lambda_2^N \vdash E_aE_b\varphi \to E_b\varphi$ (instance of axiom Ax2)
2. $\Lambda_2^N \vdash E_aE_b\varphi \to E_{\{a,b\}}\varphi$ (instance of axiom Ax10)

[14] It would just consist in rewriting the case of Constraint 9 in the proof of Theorem 4.

3. $\Lambda_2^N \vdash E_a E_b \varphi \rightarrow E_b \varphi \wedge E_{\{a,b\}} \varphi$ (from 1., and 2. by PL)

Like Constraint 9, our Constraint 10 does confer some responsibility to the delegating entity. It might therefore satisfy some scholars like Chellas who think that the delegating entity should bear responsibility. But the responsibility of the delegating entity for the bringing about of a state of affairs is shared with the delegate. Although allowed by the theory, the delegator does not have to be individually responsible. It might then satisfy scholars like Elgesem who think that Constraint 9 is nonsense. Still, the delegatee is individually responsible. As a consequence, in presence of strict-joint agency (Constraint 8 or Ax8), the consequences would be even more dramatic than in Section 5.1.4. Suppose that G_1 and G_2 are different coalitions and G_1 is not the empty coalition, we have the following derivation:

1. $\Lambda_2^N, \text{Ax8} \vdash E_{G_1} E_{G_2} \varphi \rightarrow E_{G_1 \cup G_2} \varphi$ (instance of axiom Ax10)
2. $\Lambda_2^N, \text{Ax8} \vdash E_{G_1} E_{G_2} \varphi \rightarrow E_{G_2} \varphi$ (instance of axiom Ax2)
3. $\Lambda_2^N, \text{Ax8} \vdash E_{G_2} \varphi \rightarrow \neg E_{G_1 \cup G_2} \varphi$ (instance of axiom Ax8)
4. $\Lambda_2^N, \text{Ax8} \vdash \neg E_{G_1} E_{G_2} \varphi$ (from 1., 2., and 3. by PL)

The only admissible kind of "delegation" compatible with strict-joint agency would then be one of a coalition towards itself. In formula, it is $E_G E_G \varphi \rightarrow E_G \varphi$, which is already a theorem of BIAT as an instance of Ax2.

6 Refining ability: evidences and inductive reasoning

Our interpretation of Elgesem's modality of ability as an evidence-based ability of the group G is fine in its relationship with the modality of actual agency. As such, however, it is arguably not strikingly useful. One can have the certainty that the group G is able to bring about that φ in only two circumstances: (i) G is actually bringing about φ (axiom Ax3), or (ii) some subgroups of $G_1, \ldots, G_k \subseteq G$ such that $G = G_1 \cup \ldots \cup G_k$ are bringing about some goals $\varphi_1, \ldots, \varphi_k$ such that $\varphi \leftrightarrow \varphi_1 \wedge \cdots \wedge \varphi_k$ (axiom Ax7).

The theory does not say what this ability becomes *after* the actual manifestation of the evidence. It might be true at some state where G is not bringing about that φ, but it might as well be false. As a matter of fact, "after" has no meaning in our theory, or in any logic of bringing-it-about in the literature. It is in practice of limited value, e.g., if we intend to use the existence of an ability as a suggestion of groups of services that are possibly effective for a goal.

But an evidence-based notion of ability hints that it would benefit from being investigated along with some form of induction. Induction is what allows us, at least for any practical purpose, to capitalize on the fact that a popular Internet checkout is able to bring about an online transaction. As far as we can tell, it never failed to work when we expected it to (or it was fixed soon after).

Our notion of evidence-based ability is calling for a similar kind of reasoning; one that could in fact somewhat reconcile Elgesem's and Kenny's views on ability. One can acknowledge G's ability for the goal φ at any time the coalition indeed brings it about, and then maintain this evidence-based ability *until* some further evidence falsifies it. That is, as a way to encompass Kenny's view, to possibly drop the ability when the group fails to repeat their control *when the 'appropriate opportunity' arises.*

Introducing inductive reasoning means that we need to transgress the tradition of bringing-it-about logics which is the one of abstracting away from time.[15] Indeed, we propose to introduce a linear temporal logic. In this brief account we can content ourselves with an 'until' operator U. The formula $\varphi \mathsf{U} \psi$ reads that "φ is true until ψ holds".

What would constitute an evidence for falsification of an ability? We need to introduce an elementary notion of explicit objective. For our framework, it is certainly helpful to see an objective of an agent (or of a coalition) as what Sommerhoff calls a *Focal Condition* and that is at the core of Elgesem's proposal. (See [16, Sect. 3] for the relationship of the Focal Condition with the modalities E_G and C_G.) Let us note $O_G\varphi$ to mean that "G as a group, has the objective that φ". Note that the notion of goal-directed agency justifies that:

$$E_G\varphi \rightarrow O_G\varphi.$$

We assume O_G to be primitive and do not try to explain the process of negotiation and preference aggregation to form a group objective.

Remark 4 Let us digress a moment from our discussion on ability. In Section 2.5.2, we have argued that when a group G_1 brings about X_1 and a group G_2 brings about X_2 we had no sufficient evidence that $G_1 \cup G_2$ brought about $X_1 \cap X_2$. The reason was that the group $G_1 \cup G_2$ might simply not have the goal that $X_1 \cap X_2$. However, now that we have extended our language with an explicit way to talk about the objective of coalitions, we could admit the axiom

$$E_{G_1}\varphi \wedge E_{G_2}\psi \wedge Obj_{G_1 \cup G_2}(\varphi \wedge \psi) \rightarrow E_{G_1 \cup G_2}(\varphi \wedge \psi).$$

That is, under the additional condition that $G_1 \cup G_2$ indeed has the objective that $X_1 \cap X_2$, then we can infer that $G_1 \cup G_2$ brings about $X_1 \cap X_2$.

We say that a situation disproves that G is able to achieve the goal φ when G has the objective that φ holds but does not bring about that φ. In formula, this suggests the following principle:

$$C_G\varphi \rightarrow (C_G\varphi)\mathsf{U}(O_G\varphi \wedge \neg E_G\varphi).$$

Hence, it is only by 'contextual repeatability' that one can infer that an agent or a group has the ability to exercise a certain control. The context here, is the adequate Focal Condition for the goal φ.

To illustrate this, consider again a repository of web services and a database listing their abilities for different goals. The Focal Condition $O_G\varphi$ can for

[15] Nevertheless, it should be noted that bringing-it-about logics are not inherently incompatible with a refined ontology of time, as argued in [15].

instance be instantiated by the existence of a certain query to the group G to bring about φ. The previous formula can be used to verify that the system satisfies the following property: a group of web services that are deemed able to bring about a goal, is considered able to bring about the goal as long as there is no evidence to the contrary. Notice that it does not rule out the possibility that a group has the ability and the objective to bring about a goal but fails to do so. It could be that the group failed to coordinate for some subjective reason.

After a counter-evidence, the database system (on behalf of its designer) is free (i) to drop the entry denoting the ability altogether, (ii) to flag the entry but keep it until further confirmation of inability, (iii) to ignore the counter-evidence and maintain the entry anyway. Options (i) and (ii) would be most in line with Kenny's view: who fails to repeat (possibly several times) the exercise of control when offered the opportunity is not deemed able.

7 Discussions

Before concluding this paper we present a few considerations and ultimate clarifications.

7.1 Logic of individuals, logic of coalitions

One possible concern with the logic COAL is that its restriction to the E_x modality (call it COAL(E)) does not offer a distinction and dependence between individual and group agency. In this sense, COAL(E) is not different from BIAT: they obey the same principles, and thus, it suffices to view each non-empty coalition in COAL(E) as a distinct individual in BIAT. The coalitional language gets its significance (i) with the addition of the modality of ability so as to obtain the fully fledged logic COAL; and/or (ii) with additional constraints like Constraint 10.

7.2 Seeing To It That, and action vs. interaction

In the individual case, the respective aims of the logics of bringing-it-about and those of seeing-to-it-that are not clearly different. In the coalitional case however, our extension of bringing-it-about clearly considers coalitional agency as a primitive concept, while in STIT the action of a coalition merely consists in a profile of individual actions. In STIT the set of outcomes of the actual choice of a coalition is completely determined as the set-theoretic intersection of the set of outcomes of the actual choices of the members. That is, like in non-cooperative game theory, the action of a group of agents in STIT is just a set of individual actions.

Where STIT provides the right logical tools for a *micro* analysis of interaction, bringing-it-about seems more appropriate for the formalization of a *macro* theory of interaction. It is exemplified for instance by some authors' use of bringing-it-about logics to study institutionalized agency ([25,11,41, 35]) while no similar work exist in STIT theories. On the other hand, STIT has been more prominently used in multiagent systems for the study of game-theoretical topics ([6,7,48]). It is less in tune with a macro theory of interaction because the choice of a coalition is ontologically completely determined as the intersection of the choices of the individuals (see [24,6,22]). The preponderant roles played in STIT by the empty coalition and the grand coalition (the coalition containing all the agents) are rather revealing. The potential manifestations of agency of these two special coalitions fulfill the function of two boundaries beyond which no coalition can act.

We think that the main conceptual difference is that our coalitional version of bringing-it-about looks at a notion of group agency that is not constructed from individual agency. There is no undeniable superiority of one tradition of logics of agency over the other. If we had to argue in favor of bringing-it-about, we could find some support in the criticisms by Turner of Parsons' theory of action: "The basic unit of a sociological analysis is not action, but *interaction*; and the presumption that one can begin with elementary conceptualizations of action and then progressively move up to the analysis of interaction and structure is highly questionable." ([50, p. 3]). This, we believe can be very significant for the formal representation and reasoning about multiagent systems.

A more systematic comparison of the logics of bringing-it-about with STIT theories is possible by confronting the axiomatizations of Section 3 with the axiomatizations presented in [8]. They are considerably different. The Chellas STIT operator for instance is a normal modality. In particular, that means that if an agent sees to it that the letter is sent, then he sees to it that the letter is sent or burnt. In the more goal-directed bringing-it-about however, agency is towards the goal and not towards its logical consequences. In the coalitional case, the Chellas STIT satisfies coalition monotonicity that we have rejected in Section 2.4 as well as superadditivity of agency that we have rejected in Section 2.5.2.

7.3 Coalitions and Elgesem's concepts of agency

Elgesem studied agency and ability through a net of derived concepts. We simply lift them to the coalitions.

$$\begin{aligned}
Does_G\varphi &\overset{\text{def}}{=} E_G\varphi \\
Ability_G\varphi &\overset{\text{def}}{=} C_G\varphi \\
Compatible_G\varphi &\overset{\text{def}}{=} \neg Does_G\neg\varphi \\
Unpreventable_G\varphi &\overset{\text{def}}{=} \neg Ability_G\neg\varphi \\
Independently_G\varphi &\overset{\text{def}}{=} \neg Does_G\varphi \wedge \varphi \\
Opportunity_G\varphi &\overset{\text{def}}{=} Does_G\neg\varphi \vee \varphi
\end{aligned}$$

The first two lines are simply a redefinition of our operators into Elgesem's notation. The exact meaning of these concepts for coalitional agency should not be different from the meaning in the individual case. We will not enter in the details and we simply refer to [16]. In general, one must resist the temptation to make up new apparently intuitive laws of coalitional agency. An intuitive reading of a logical language should not prevail over the insights of its semantics.

The seemingly useful formula

$$Does_{\{a,b\}}\varphi \rightarrow Compatible_{\{a\}}\varphi$$

(if a group G brings about that φ, then φ is compatible with the agency of every member of G) is already a theorem of BIAT because of Ax2. In fact it is true that $Does_{G_1}\varphi \rightarrow Compatible_{G_2}\varphi$ for any two arbitrary coalitions G_1 and G_2.

One might also consider useful the formula

$$Does_{\{a,b\}}\varphi \rightarrow \neg Independently_{\{a\}}\varphi.$$

It captures the fact that if a group G brings about that φ, then φ is not independent of the agency of any of the members of G. The issue is that the agency of the individual members $a \in G$ represented by a formula like $Does_a\psi$ captures exclusively the actions of the individual a acting alone towards an individual goal. A consequence of adopting the formula above would be $Does_{\{a,b\}}\varphi \rightarrow Does_a\varphi$, which should not be a general principle according to our interpretation of agency.

8 Conclusions and perspectives

We have lifted Elgesem's logic of individual agency and ability to the coalitional setting. We showed that it offers more flexibility in identifying potential ability of the multiagent system to bring about complex goals. We provided a complete axiomatization and studied the computational complexity for satisfiability checking of the logics of bringing-it-about. We offered alternating polynomial-time algorithms to solve the problem of satisfiability checking in the logics BIAT, ELG and COAL.

We then proposed several specializations of COAL. We illustrated the logic by showing that commands within a team is impossible when we assume strict-joint agency and full responsibility of the delegating entity, two common principles of agency. Capitalizing on our language with coalitions, we proposed a new principle that states the occurrence of a delegation implies a responsibility of the team consisting of the delegating entity and of the delegate. We believe this principle might reconcile an argument in the literature about the principle of individual responsibility of the delegating entity. We sketched the idea of integrating temporal features in the models of bringing-it-about.

A few perspectives emerge from the present research. The first one is technical and concerns the computational complexity of the logics. We only provided upper-bounds. Even though there are many reasons to believe that the results are tight,[16] we did not rule out the possibility that maybe our algorithms are not optimal. Proving lower-bounds for non-normal modal logics, however, is a very delicate problem in mathematical logic. For normal modal logics, the standard proof relies heavily on the existence of a semantics in terms of Kripke models ([30, 20]). We do not have this luck here.

Two other perspectives concern the extensions of COAL that we have proposed in Section 5.2 and in Section 6, and of which we only scratched the surface.

A more empirical study could be done concerning the adequateness of Constraint 10 in some application domains. Does it correspond to some real world situations? Does it reflect the decision in the attribution of responsibilities in some court set of laws? Our brief logical analysis falls short of providing answers to these questions, that are beyond the scope of this paper.

Finally, a rigorous framework integrating inductive reasoning with the concepts or agency and evidence-based ability could be enlightening. It would have the potential to bridge the gap between several notions of ability, power, opportunity, etc, that have been studied in the literature. In general, we believe that composing the very abstract models of bringing-it-about with simple temporal models is promising. Recently, the seeing-to-it-that logics seem to have been preferred by the community, arguably because of the temporal features of their models. It will be interesting to investigate further what new insights we can find for a number of topics studied in other multiagent theories: e.g., knowing how to play, strategically bringing about, etc.

Annex: proof of Theorem 3

Left-to-right direction

Suppose φ is BIAT-sat. This means that there is a model of individual agency $M = (S, N, EE, v)$, and a world $w \in S$, such that $M, w \models \varphi$. Define the function $\pi : sub^{\neg}(\varphi) \longrightarrow \{0, 1\}$ as follows: for all $\psi \in sub^{\neg}(\varphi)$,

$$\pi(\psi) = \begin{cases} 1 & \text{when } M, w \models \psi \\ 0 & \text{otherwise.} \end{cases}$$

It is readily checked that π is a semi-valuation. It remains to show that π satisfies all three conditions.

[16] A rule of thumb is that over a non-restricted language (all Boolean combinations, unbounded modal depth, etc), if a classical modal logic admits axiom Ax1 (axiom K) but does not admit $\neg E_x\varphi \to E_x\neg E_x\varphi$ (axiom 5), it is PSPACE-hard.

1. if $\pi(E_a\varphi) = 1$ then $M, w \models E_a\varphi$ by definition of π. So $||\varphi||^M \in EE_w(a)$. Since M is a model of individual agency we also have that $w \in ||\varphi||^M$ by Constraint 2. Hence, $M, w \models \varphi$, which means that $\pi(\varphi) = 1$.
2. As previously, $\pi(E_a\varphi) = 1$ implies that $||\varphi||^M \in EE_w(a)$. By Constraint 1E $||\varphi||^M \neq S$. It means that there is a world $u \in ||\neg\varphi||^M$. Hence, $M, u \models \neg\varphi$, which means that $\neg\varphi$ is BIAT-sat.
3. Again, for all j, $\pi(E_a\psi_j) = 1$ implies that $||\psi_j||^M \in EE_w(a)$. By Constraint 4 we obtain that $(\bigcap_j ||\psi_j||^M) \in EE_w(a)$. By hypothesis, $\pi(E_a\psi) = 0$. Thus, $||\psi||^M \notin EE_w(a)$ and clearly $||\psi||^M \neq \bigcap_j ||\psi_j||^M$. It implies that there must exist a world u such that either $u \in \bigcap_j ||\psi_j||^M \cap ||\neg\psi||^M$ or $u \in \bigcup_j ||\neg\psi_j||^M \cap ||\psi||^M$. It means that the formula $(\bigwedge_j(\psi_j) \wedge \neg\psi) \vee (\bigvee_j(\neg\psi_j) \wedge \psi)$ is indeed satisfiable: it is true in the world u.

Right-to-left direction

Suppose now that there is a semi-valuation π such that we have (1), (2) and (3) like in the formulation of Theorem 3. We can rewrite (3) as: if $E_a\psi_1, \ldots, E_a\psi_k, E_a\psi \in sub^{\neg}(\varphi)$, with $\pi(E_a\psi_j) = 1$ for all j, and $\pi(E_a\psi) = 0$, then there is a *witness* pointed model of individual agency $(M_{\psi_1,\ldots\psi_k,\psi}, w_{\psi_1,\ldots\psi_k,\psi})$ such that $M_{\psi_1,\ldots\psi_k,\psi}, w_{\psi_1,\ldots\psi_k,\psi} \models \Big(\bigwedge_j(\psi_j) \wedge \neg\psi\Big) \vee \Big(\bigvee_j(\neg\psi_j) \wedge \psi\Big)$. We can also rewrite (2) as: if $E_a\psi \in sub^{\neg}(\varphi)$ and $\pi(E_a\psi) = 1$ then there is a *witness* pointed model of individual agency (M_ψ, w_ψ) such that $M_\psi, w_\psi \models \neg\psi$.

Let $(M^1, w^1) \ldots, (M^{\mathsf{n}}, w^{\mathsf{n}})$ be an enumeration of all these witness pointed models of agency $M^i = (S^i, N, EE^i, v^i)$. This enumeration is finite because the length of φ is finite and only a finite number of $E_a\psi$ are in $sub^{\neg}(\varphi)$. We assume w.l.o.g. that $S^i \cap S^j = \emptyset$ when $i \neq j$.

To establish that φ is BIAT-sat, we are going to construct a model of individual agency that satisfies it. Suppose a new fresh world $\mathsf{w} \notin \bigcup_i S^i$. Let $M = (S, N, EE, v)$ such that: $S = \{\mathsf{w}\} \cup \bigcup_{i=1}^{\mathsf{n}} S^i$.

Instrumental for the definition of v and EE, we first introduce an intention assignment for formulas in $sub^{\neg}(\varphi)$. Let $V : sub^{\neg}(\varphi) \longrightarrow 2^S$, such that

$$V(\psi) = \{\mathsf{w} \mid \pi(\psi) = 1\} \cup \bigcup_{i=1}^{\mathsf{n}} ||\psi||^{M^i}.$$

Our model valuation v is simply defined as the projection of V on the set P of propositional variables (we can assume w.l.o.g. that $P \subseteq sub^{\neg}(\varphi)$). That is, $v(p) = V(p)$ for $p \in P$.

We finally define EE. For every $u \in S$, we let $X \in EE_u(a)$ iff there are $E_a\psi_1, \ldots, E_a\psi_k \in sub^{\neg}(\varphi)$, with $X = \bigcap_{j=1}^k V(\psi_j)$ and

$$\begin{cases} \pi(E_a\psi_j) = 1 \text{ for all } 1 \leq j \leq k & \text{when } u = \mathsf{w}; \\ M^i, u \models E_a\psi_j \text{ for all } 1 \leq j \leq k & \text{when } u \in S^i, 1 \leq i \leq \mathsf{n}. \end{cases}$$

This following claim will be useful later.

Claim 1 *If $V(\psi) \in EE_a(u)$ and $E_a\psi \in sub^{\neg}(\varphi)$ then either $u = \mathsf{w}$ and $\pi(E_a\psi) = 1$ or $u \in S^c$ and $M^c, u \models E_a\psi$ $(1 \leq c \leq \mathsf{n})$.*

Proof Case $u = \mathsf{w}$: suppose $V(\psi) \in EE_a(\mathsf{w})$. That is there $E_a\psi_1, \dots E_a\psi_k \in sub^{\neg}(\varphi)$ such that $V(\psi) = \bigcap_{j=1}^k V(\psi_j)$ and $\pi(E_a\psi_j) = 1$ for all j.

Now suppose that $E_a\psi \in sub^{\neg}\varphi$ and for contradiction, that $\pi(E_a\psi) = 0$. By hypothesis (3), $M_{\psi_1,\dots\psi_k,\psi}, w_{\psi_1,\dots\psi_k,\psi} \models \left(\bigwedge_j(\psi_j) \wedge \neg\psi\right) \vee \left(\bigvee_j(\neg\psi_j) \wedge \psi\right)$. This means that $||\psi||^{M_{\psi_1,\dots\psi_k,\psi}} \neq \bigcap_{j=1}^k ||\psi_j||^{M_{\psi_1,\dots\psi_k,\psi}}$. Consequently, $V(\psi) \neq \bigcap_{j=1}^k V(\psi_j)$—a contradiction.

Case $u \in S^c$: suppose $V(\psi) \in EE_a(u)$. So there are $E_a\psi_1, \dots, E_a\psi_k \in sub^{\neg}\varphi$ such that $V(\psi) = \bigcap_{j=1}^k V(\psi_j)$ and $M^c, u \models E_a\psi_j$ for all $1 \leq j \leq k$. So we have $||\psi_j||^{M^c} \in EE_a^c(u)$ for all $1 \leq j \leq k$. By Constraint 4 on M^c we deduce that $\bigcap_{j=1}^k ||\psi_j||^{M^c} \in EE_a^c(u)$. Since $V(\psi) = \bigcap_{j=1}^k V(\psi_j)$ and because $S^i \cup S^j = \emptyset$ when $i \neq j$, we also have that $||\psi||^{M^c} = \bigcap_{j=1}^k ||\psi_j||^{M^c}$. Hence, $||\psi||^{M^c} \in EE_a^c(u)$, which means that $M^c, u \models E_a\psi$.

We need to make sure that M is indeed a model of individual agency.

Claim 2 *M is a model of individual agency.*

Proof Constraint 1E: Clearly by construction and because $S^i \notin EE_u^i(a)$, we have $S \notin EE_u(a)$ for $u \neq \mathsf{w}$. Now, suppose for contradiction that $S \in EE_{\mathsf{w}}(a)$. It means that there are $E_a\psi_1, \dots, E_a\psi_k \in sub^{\neg}(\varphi)$, with $S = \bigcap_{j=1}^k V(\psi_j)$ and $\pi(E_a\psi_j) = 1$ for all $1 \leq j \leq k$. Hence, $V(\psi_j) = \{\mathsf{w} \mid \pi(\psi_j) = 1\} \cup \bigcup_{i=1}^{\mathsf{n}} ||\psi_j||^{M^i} = S$ for all $0 \leq j \leq k$. As we assumed that $S^i \cap S^j = \emptyset$ when $i \neq j$, it means that $||\psi_j||^{M^i} = S^i$ for all i. In particular, $||\psi_j||^{M^j} = S^j$. But since, $\pi(E_a\psi_j) = 1$, by hypothesis (2), it means that $M_j, w_j \models \neg\psi_j$, that is, $w_j \in \emptyset$—a contradiction.

Constraint 2: Now, suppose that $X \in EE_u(a)$. We need to show that $u \in X$.

By construction, there are $E_a\psi_1, \dots, E_a\psi_k \in sub^{\neg}(\varphi)$, such that $X = \bigcap_{j=1}^k V(\psi_j)$ and

$$\begin{cases} \pi(E_a\psi_j) = 1 \text{ for all } 1 \leq j \leq k & \text{if } u = \mathsf{w}; \\ M^i, u \models E_a\psi_j \text{ for all } 1 \leq j \leq k & \text{if } u \in S^i, 1 \leq i \leq \mathsf{n}. \end{cases}$$

If $u \in S^c$, we have $||\psi_j||^{M^c} \in EE_u(a)$ for all $1 \leq j \leq k$. Since M^c is a model of individual agency, we also have $u \in ||\psi_j||^{M^c}$ for all $1 \leq j \leq k$ (Constraint 2 on M^c). So $u \in \bigcap_{j=1}^k ||\psi_j||^{M^c}$. Then, $u \in \bigcap_{j=1}^k \bigcup_{i=1}^{\mathsf{n}} ||\psi_j||^{M^i}$. Then, $u \in \bigcap_{j=1}^k \{\mathsf{w} \mid \pi(\psi_j) = 1\} \cup \bigcap_{j=1}^k \bigcup_{i=1}^{\mathsf{n}} ||\psi_j||^{M^i}$ which is equivalent to $u \in X$.

If $u = \mathsf{w}$, by hypothesis (1), we have $\pi(\psi_j) = 1$ for all $1 \leq j \leq k$. So indeed, $u \in \bigcap_{j=1}^k V(\psi_j)$.

Constraint 4: $X_1 \in EE_u(a)$ and $X_2 \in EE_u(a)$. We need to show that $(X_1 \cap X_2) \in EE_u(a)$.

We do not give the details. Again, if $u \in S^c$ ($1 \leq c \leq \mathsf{n}$), it holds in virtue of M^a being a model of individual agency; this time we use Constraint 4 on M^a. If $u = \mathsf{w}$, the proof uses hypothesis (3).

Finally, we are going to prove that φ is true in M at w. To do this, we show that for all $\psi \in sub^{\neg}(\varphi)$, we have $V(\psi) = ||\psi||^M$.

Claim 3 *If* $\psi \in sub^{\neg}(\varphi)$ *then* $V(\psi) = ||\psi||^M$.

Proof The proof is by induction on the structure of ψ. This holds for $\psi = p$, $p \in P$, by definition of $v(p)$ and $||p||^M$. It also holds for $\psi = \neg\gamma$, and $\psi = \gamma_1 \vee \gamma_2$, because $V(\neg\gamma) = S \setminus V(\gamma)$ and $V(\gamma_1 \vee \gamma_2) = V(\gamma_1) \cup V(\gamma_2)$ respectively.

For $\psi = E_a\gamma$, we assume inductively that $V(\gamma) = ||\gamma||^M$ and show that for every $u \in S$, $u \in V(E_a\gamma)$ iff $u \in ||E_a\gamma||^M$.

Suppose $u \in V(E_a\gamma)$ then either $u = \mathsf{w}$ and $\pi(E_a\gamma) = 1$, or $u \in S^a$ and $M^a, u \models E_a\gamma$. In both cases, by definition of EE, we have $V(\gamma) \in EE_a(u)$. By induction hypothesis, we have $||\gamma||^M \in EE_a(u)$. Hence $M, u \models E_a\gamma$, that is, $u \in ||E_a\gamma||^M$.

The other way around suppose that $u \in ||E_a\gamma||^M$, that is, $M, u \models E_a\gamma$. It means that $||\gamma||^M \in EE_a(u)$. By induction hypothesis, $V(\gamma) \in EE_a(u)$. We conclude using Claim 1 that $u \in V(E_a\gamma)$.

To sum up, we have built a structure M that is a model of individual agency (Claim 2), and $||\varphi||^M = V(\varphi)$ (Claim 3). By definition of V and because $\pi(\varphi) = 1$ (π is a semi-valuation of φ), we have that $M, \mathsf{w} \models \varphi$. So indeed, φ is $\mathsf{BIAT\text{-}sat}$. This concludes the proof of Theorem 3.

Acknowledgements I would like to thank Andreas Herzig and Emiliano Lorini, as well as my colleagues at the Laboratory for Applied Ontology for inspiring and commenting earlier presentations of this work. It also greatly benefited from suggestions and corrections of reviewers on previous drafts. This research was funded by a Marie Curie Actions Fellowship FP7 PEOPLE COFUND Trentino.

References

1. Abdou, J., Keiding, H.: Effectivity functions in social choice. Kluwer Academic Publishers (1991)
2. Alur, R., Henzinger, T.A., Kupferman, O.: Alternating-time temporal logic. J. ACM **49**(5), 672–713 (2002)
3. Belnap, N., Perloff, M.: Seeing to it that: a canonical form for agentives. Theoria **54**(3), 175–199 (1988)
4. Belnap, N., Perloff, M., Xu, M.: Facing the Future (Agents and Choices in Our Indeterminist World). Oxford University Press (2001)
5. Bottazzi, E., Ferrario, R.: Critical situations from spontaneous to sophisticated social interactions (2011). In *New Trends in the Philosophy of the Social Sciences* (Madrid, 28-29/09 2011)
6. Broersen, J., Herzig, A., Troquard, N.: Embedding Alternating-time Temporal Logic in strategic STIT logic of agency. Journal of Logic and Computation **16**(5), 559–578 (2006)

7. Broersen, J., Herzig, A., Troquard, N.: Normal simulation of coalition logic and an epistemic extension. In: Proceedings of TARK 2007. ACM DL, Brussels, Belgium (2007)
8. Broersen, J., Herzig, A., Troquard, N.: What groups do, can do, and know they can do: an analysis in normal modal logics. Journal of Applied Non-Classical Logics **19**(3), 261–290 (2009)
9. Brown, M.A.: On the logic of ability. Journal of Philosophical Logic **17**, 1–26 (1988)
10. Carmo, J.: Collective agency, direct action and dynamic operators. Logic Journal of the IGPL **18**(1), 66–98 (2010)
11. Carmo, J., Pacheco, O.: Deontic and action logics for organized collective agency modeled through institutionalized agents and roles. Fund. Inform. **48**, 129–163 (2001)
12. Chellas, B.: The Logical Form of Imperatives. Perry Lane Press (1969)
13. Chellas, B.: Modal Logic: An Introduction. Cambridge University Press (1980)
14. Davidson, D.: The Logical Form of Action Sentences. In: N. Rescher (ed.) The Logic of Decision and Action. University of Pittsburgh Press (1967)
15. Elgesem, D.: Action theory and modal logic. Ph.D. thesis, Universitetet i Oslo (1993)
16. Elgesem, D.: The modal logic of agency. Nordic J. Philos. Logic **2**(2) (1997)
17. Frankfurt, H.: The Importance of what We Care About. Cambridge University Press (1988)
18. Goranko, V., Jamroga, W., Turrini, P.: Strategic Games and Truly Playable Effectivity Functions. Autonomous Agents and Multi-Agent Systems **26**, 288–314 (2013)
19. Governatori, G., Rotolo, A.: On the Axiomatisation of Elgesem's Logic of Agency and Ability. Journal of Philosophical Logic **34**, 403–431 (2005)
20. Halpern, J.Y., Moses, Y.: A guide to completeness and complexity for modal logics of knowledge and belief. Artif. Intell. **54**(2), 319–379 (1992)
21. Harel, D., Kozen, D., Tiuryn, J.: Dynamic Logic. MIT Press (2000)
22. Herzig, A., Schwarzentruber, F.: Properties of logics of individual and group agency. In: Advances in Modal Logic, pp. 133–149 (2008)
23. Herzig, A., Troquard, N.: Knowing how to play: uniform choices in logics of agency. In: Proceedings of AAMAS 2006, pp. 209–216. IFAAMAS (2006)
24. Horty, J.F.: Agency and Deontic Logic. Oxford University Press, Oxford (2001)
25. Jones, A., Sergot, M.: A formal characterization of institutionalised power. Journal of the IGPL **4**(3), 429–445 (1996)
26. Kanger, S., Kanger, H.: Rights and Parliamentarism. Theoria **32**, 85–115 (1966)
27. Kapitan, T.: Incompatibilism and ambiguity in the practical modalities. Analysis **56**(2), 102–110 (1996)
28. Kenny, A.: Will, Freedom and Power. Blackwell (1975)
29. von Kutschera, F.: Bewirken. Erkenntnis: an international journal of analytic philosophy **24**(3) (253–281)
30. Ladner, R.E.: The computational complexity of provability in systems of modal propositional logic. SIAM J. Comput. **6**(3), 467–480 (1977)
31. Lindahl, L.: Position and Change – A Study in Law and Logic. D. Reidel (1977)
32. Mele, A.R.: Agent's Abilities. Noûs pp. 447–470 (2003)
33. Miller, S.: Social Action (A Teleogical Account). Cambridge University Press (2001)
34. Norman, T.J., Reed, C.: A logic of delegation. Artificial Intelligence **174**, 51–71 (2010)
35. Pacheco, O., Carmo, J.: A Role Based Model for the Normative Specification of Organized Collective Agency and Agents Interaction. Autonomous Agents and Multi-Agent Systems **6**, 145–184 (2003)
36. Pauly, M.: A Modal Logic for Coalitional Power in Games. J. Log. Comput. **12**(1), 149–166 (2002)
37. Pörn, I.: The Logic of Power. Blackwell, Oxford (1970)
38. Pörn, I.: Action Theory and Social Science: Some Formal Models. Synthese Library 120. D. Reidel, Dordrecht (1977)
39. Santos, F., Carmo, J.: Indirect Action, Influence and Responsibility. In: Proceedings of DEON 1996, pp. 194–215. Springer-Verlag (1996)
40. Santos, F., Jones, A., Carmo, J.: Responsibility for Action in Organisations: a Formal Model. In: G. Holmström-Hintikka, R. Tuomela (eds.) Contemporary Action Theory, vol. 1, pp. 333–348. Kluwer (1997)
41. Santos, F., Pacheco, O.: Specifying and reasoning with institutional agents. In: Proceedings of ICAIL 2003, pp. 162–170. ACM (2003)

42. Schmid, H.B.: Plural Action. Philosophy of the Social Sciences **38**(1), 25–54 (2008)
43. Sebastiani, R., Tacchella, A.: SAT Techniques for Modal and Description Logics. In: A. Biere, M. Heule, H. Van Maaren, T. Walsh (eds.) The Handbook of Satisfiability, vol. II, chap. 25. IOS Press (2009)
44. Segerberg, K.: An essay in Classical Modal Logic. Filosofiska Studier, Uppsala Universitet (1971)
45. Shapley, L.S., Shubik, M.: The assignment game I: The core. Int. J. Game Theory **1**(1), 111–130 (1971)
46. Sommerhoff, G.: The Abstract Characteristics of Living Systems. In: F.E. Emery (ed.) Systems Thinking: Selected Readings. Penguin, Harmonsworth (1969)
47. Troquard, N.: Coalitional Agency and Evidence-Based Ability. In: Proceedings of AAMAS 2012, pp. 1245–1246. IFAAMAS (2012)
48. Troquard, N., van der Hoek, W., Wooldridge, M.: Model checking strategic equilibria. In: Proceedings of MoChArt 2008, *Lecture Notes in Computer Science*, vol. 5348, pp. 166–188. Springer Berlin / Heidelberg (2009)
49. Tuomela, R.: The Importance of Us. Stanford University Press (1995)
50. Turner, J.H.: A Theory of Social Interaction. Stanford University Press, Stanford, CA (1988)
51. Vardi, M.: On the Complexity of Epistemic Reasoning. In: Proceedings of LICS 1989, pp. 243–252. IEEE Computer Society (1989)
52. Wobcke, W.: Agency and the logic of ability. In: Agents and Multi-Agent Systems Formalisms, Methodologies, and Applications, pp. 31–45 (1997)