

# Coalitional Agency and Evidence-Based Ability

## (Extended Abstract)

Nicolas Troquard<sup>\*</sup>  
Laboratory for Applied Ontology (ISTC-CNR)  
Trento, Italy  
troquard@loa.istc.cnr.it

### ABSTRACT

The logics of “bringing it about” have been part of a prominent tradition for the formalization of individual and institutional agency. Our objective here is to extend Elgesem’s logic of individual agency and ability to coalitions.

### Categories and Subject Descriptors

I.2.11 [Distributed Artificial Intelligence]: Multiagent systems; I.2.4 [Knowledge representation formalisms and methods]: Modal logic

### General Terms

Theory

### Keywords

logic, bringing-it-about, coalitions, agency, ability

## 1. EXTENDED ABSTRACT

This extended abstract aims to contribute to the literature that views an action as the mere result of the activity of an agent. 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 other traditions of logic of action talking explicitly about action terms: for instance, Dynamic Logics in computer science, or the study of action sentences in philosophy using first-order theories.

Pörn, Elgesem ([1]), and others, have studied the modality of agency in the Anselmian tradition. The bringing-it-about modality  $E_x$  has been quite popular in the MAS community. (E.g., [4]). 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  $\varphi$ ”. In the literature,  $x$  has been either an individual agent, or an institutional agent. An institution can involve several agents, each playing a specific role in it. But institutions are not groups or coalitions. Our contribution is an extrapolation of a theory of coalitional agency and ability from Elgesem’s account of

<sup>\*</sup>Supported by a Marie Curie fellowship.

**Appears in:** *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

Copyright © 2012, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

individual agency. That is, we study the logic of the operator  $E_x$  where  $x$  is a set of agents, along with a coalitional operator of ability.

**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 is studied alongside related modalities of action. The logic still admits the core principles that are generally assumed for agency:

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

Following Sommerhoff, Elgesem argues that agency is the actual bringing about of a goal towards which an activity is oriented. An agent acts 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 according to whom, the pertinent aspect of agency is the manifestation of the agent’s guidance 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.

He observes that the manifestation of guidance is the exercise of a power to bring about something. Therefore, the notion of potential guidance, or ability, of an agent for a goal should be integrated in a theory of agency. Elgesem argues, much contradicting Kenny, 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 guidance 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. By bringing about something, an agent *shows* that he is indeed able to do so. We advance an interpretation of *evidence-based* ability.

Governatori and Rotolo proved the completeness of Elgesem’s logic ([2]). The principles are the following, where  $C_x\varphi$  reads “the acting entity  $x$  is able to bring about  $\varphi$ ”.

- all the previous principles
- $\vdash \neg C_x \perp$
- $\vdash \neg C_x \top$
- $\vdash E_x \varphi \rightarrow C_x \varphi$
- if  $\vdash \varphi \leftrightarrow \psi$  then  $\vdash C_x \varphi \leftrightarrow C_x \psi$

The logic of  $C_x$  is then rather weak. The only certainty one can have about the presence of an ability to bring about  $\varphi$  is in the presence of an actual bringing about of  $\varphi$ .

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 ways 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. Nevertheless, 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.

**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. Despite resorting to some notion of shared goal, Miller ([3]) argues that we-intentions 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. 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 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 an incremental procedure for web service discovery. 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.

**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 guidance 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 guidance for it. Hence

$$\vdash \neg C_\emptyset \varphi.$$

Together with the principle  $\vdash E_x \varphi \rightarrow C_x \varphi$  adopted above, it follows from it that  $\vdash \neg E_\emptyset \varphi$ , too.

**Evidence of coalitional ability.** If a coalition  $G_1$  brings about  $\varphi$  and a coalition  $G_2$  brings about  $\psi$ , had they acted as the coalition  $G_1 \cup G_2$  they would have together brought

about  $\varphi \wedge \psi$ . Our evidence-based perspective of ability suggests that as they showed evidence, they are deemed able. In formula:

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

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  $\varphi \wedge \psi$ , for instance in a context of web services orchestration.

**The logic of coalitional agency and ability.** Our methodology to finding the coalitional version of Elgesem’s logic rather naïvely consists in thinking of a principle and trying to show that it is not acceptable in some scenario. If no counterexample is found, we must accept it at that stage. We found only the previous two principles that we think are adequate with Elgesem’s philosophy and our analysis above.

The logic of coalitional agency and ability can be conveniently presented as a Hilbert system. For all groups  $G$ ,  $G_1$ , and  $G_2$  and formulas  $\varphi$  and  $\psi$ :

- Ax0**  $\vdash \varphi$  , when  $\varphi$  is a tautology in propositional logic
- Ax1**  $\vdash E_G \varphi \wedge E_G \psi \rightarrow E_G (\varphi \wedge \psi)$
- Ax2**  $\vdash E_G \varphi \rightarrow \varphi$
- Ax3**  $\vdash E_G \varphi \rightarrow C_G \varphi$
- Ax4**  $\vdash \neg C_G \perp$
- Ax5**  $\vdash \neg C_G \top$
- Ax6**  $\vdash \neg C_\emptyset \varphi$
- Ax7**  $\vdash E_{G_1} \varphi \wedge E_{G_2} \psi \rightarrow C_{G_1 \cup G_2} (\varphi \wedge \psi)$
- ERE** if  $\vdash \varphi \leftrightarrow \psi$  then  $\vdash E_G \varphi \leftrightarrow E_G \psi$
- ERC** if  $\vdash \varphi \leftrightarrow \psi$  then  $\vdash C_G \varphi \leftrightarrow C_G \psi$

From here, one can provide a class of models for which the logic is sound and complete. It can be proved that the decision problem of satisfiability checking within the logic can be solved in space polynomial.

**Towards stronger logics.** From this minimal logic, one can strengthen it and adapt it to more specific application domains. For instance, the language of our logic talking about coalitions, allows to formulate a variant to the controversial principle of law  $E_x E_y \varphi \rightarrow E_x \varphi$  that states that the delegating entity  $x$  is responsible for what the delegate  $y$  brings about. Elgesem rejected it. Instead we could adopt  $E_{G_1} E_{G_2} \varphi \rightarrow E_{G_1 \cup G_2} \varphi$ , which only attributes a shared responsibility to the delegating entity. Nevertheless,  $E_{G_1} E_{G_2} \varphi \rightarrow E_{G_2} \varphi$  remains true in virtue of **Ax2**.

## 2. REFERENCES

- [1] D. Elgesem. The modal logic of agency. *Nordic J. Philos. Logic*, 2(2), 1997.
- [2] G. Governatori and A. Rotolo. On the Axiomatisation of Elgesem’s Logic of Agency and Ability. *Journal of Philosophical Logic*, 34:403–431, 2005.
- [3] S. Miller. *Social Action (A Teleological Account)*. Cambridge University Press, 2001.
- [4] F. Santos, A. Jones, and J. Carmo. Responsibility for Action in Organisations: a Formal Model. In G. Holmström-Hintikka and R. Tuomela, editors, *Contemporary Action Theory*, volume 1, pages 333–348. Kluwer, 1997.