



# Bounded Multi-agent Reasoning: Actualizing Distributed Knowledge

Anthia Solaki<sup>(✉)</sup>

ILLC, University of Amsterdam, Amsterdam, The Netherlands  
a.solaki2@uva.nl

**Abstract.** The idealizations resulting from the use of Kripke semantics in Epistemic Logic are inherited by formalizations of group epistemic notions. For example, *distributed knowledge* (DK) is often taken to reflect the potential knowledge of a group: what agents *would* know if they had unbounded means of communication and deductive ability. However, this does not specify whether/how this potential can be actualized, especially since real people are not unbounded reasoners. Inspired by experiments on group reasoning, we identify two dimensions of actualizing DK: communication and inference. We build a dynamic framework with effortful actions accounting for both, combining *impossible-worlds semantics* and *action models*, and we provide a method for extracting a sound and complete axiomatization.

## 1 Introduction

Epistemic Logic (EL), seen as a normal modal logic (usually S5), has been used in the study of multi-agent systems, modelling not only the individual knowledge of each agent, but also collective epistemic notions. For example, a group is said to have *common knowledge* (CK) of  $\phi$  whenever everybody knows that everybody knows (ad infinitum) that  $\phi$ , and *distributed knowledge* (DK) of  $\phi$  whenever agents can deduce  $\phi$  by pooling their knowledge together. With the tools of *Dynamic Epistemic Logic* (DEL), we can further capture the communicative actions giving rise to them, e.g. actions actualizing DK and converting it into CK.

However, EL is often criticized on grounds of idealization: its predictions are practically unattainable by real agents. This has implications for collective notions. It can well be that members of a group do not know all logical consequences of their knowledge (e.g. because of memory overload) or do not take all necessary communicative actions (e.g. because of time pressure). The same constraints apply to higher-order reasoning as agents cannot ascribe knowledge to others to an infinite modal depth. Group reasoning is a dynamic, mixed task that requires actions of both inference and communication. These are not always affordable by human agents, given their cognitive limitations. Therefore, the evolution of reasoning is bounded by agents' resources. Even from a normative viewpoint, it makes sense to study what can be *feasibly* asked of them.

This is corroborated by empirical findings. In deductive reasoning tasks people often have trouble applying certain inference rules. Perhaps the best known task is the *Wason selection task* [34]:

Four cards are given to the participants. Each card has a number on one side and a letter on the other. The visible sides of the cards read A, K, 4, and 7. The participants are asked which cards need to be turned to check whether the following holds: *if a card has a vowel on one side, it has an even number on the other.*

Individuals do notoriously bad in the task, although it involves just applications of Modus Ponens and Modus Tollens. This has given rise to theories in psychology of reasoning, explaining the asymmetry between the cognitive difficulty of different inferences [24, 28]. Other findings study the difficulty of reasoning about others [32]. Group variants of deductive tasks similarly reveal limits in group reasoning. Nonetheless, they also allow us to track which actions underlie successful performance and the effort they require. Its distribution among members often yields better performance compared to the individual case [21, 30].

In light of this, we can revisit group epistemic notions from the perspective of non-ideal agents. Using DEL, we can specify the intertwined effortful actions (communicative and inferential) that refine group knowledge, in accord with empirical facts. Revisiting DK is a first step because of the implicit flavour underlying its understanding as what *would* be known, *if* the agents were to pool their knowledge and deduce information on its basis. In revisiting DK, we need to specify (i) which actions may “actualize” it, i.e. turn it into (explicit) mutual knowledge of the group, and (ii) to what extent these can be undertaken, given that agents are bounded.

The first type of actions is *communicative actions*. Subtleties underpinning the understanding of DK as the outcome of some (unlimited) communication among group members have been discussed in [18, 29, 36]. The latter consider the formula  $p \wedge \neg K_{a_1} p$ :  $p$  is true but  $a_1$  does not know it. The formula  $D_G(p \wedge \neg K_{a_1} p)$ , where  $G$  is a group including  $a_1$ , is consistent in extensions of EL with  $D_G$  operators standing for DK. Yet no communication could render this mutual knowledge of  $G$ . The problem lies in that the formula is evaluated in a model that does not explicitly encode the effect of information pooling taking place. The operation introduced by the authors to fill this gap is called *resolution* and it is similar to operations in [6, 11].

Since our goal is to do justice to non-ideal agents, we should further account for the extent to which resolution can be undertaken. This has implications for the second type of actions too, namely *inferential actions*. There is more than pooling information together that occurs in group deliberations, but unlike communication, the deductive reasoning of group members is usually neglected in multi-agent EL, whereby agents automatically know all consequences of their knowledge. As with communication, we want to encode explicitly the inferential actions of group members, and the extent to which these can be undertaken.

**Outline.** In Sect. 2, we present our framework accounting for how agents actualize DK under resource-bounds, using a novel combination of *impossible-worlds semantics* [12] and *action models* [5]. We illustrate its workings in Sect. 3 and provide a method for the extraction of a sound and complete axiomatization in Sect. 4.

## 2 The Framework

### 2.1 Syntax

The logical language of our framework extends that of standard multi-agent epistemic logics. Given a non-empty set of agents  $Ag$ , it includes:

- i Quantitative comparisons between terms that are introduced to capture cognitive costs of actions (communicative, inferential) with the cognitive capacities of agents.
- ii Operators  $D_G$ , standing for the *distributed knowledge* of group  $G \subseteq Ag$ .
- iii Operators  $A_j$ , where  $j \in Ag$ , that indicate the inference rules that agent  $j$  has acknowledged as truth-preserving (similar to [16, 31, 35]).
- iv Operators  $\langle R_G \rangle$ , standing for *resolution* of group  $G$ , i.e. actions of communication through which members pool their knowledge together (in the spirit of operations appearing in [6, 11, 36]).
- v Operators of the form  $\langle C, e \rangle$ , where  $e$  is an *event* in *action model*  $C$  designed to capture applications of inference rules in a multi-agent setting.

In order to define the language formally we need the following two prerequisites. Given the propositional language  $\mathcal{L}_P$  based on a set of atoms  $P$ :

**Definition 2.1 (Rule).** *An inference rule  $\rho$  is of the form  $\{\phi_1, \dots, \phi_n\} \rightsquigarrow \psi$  where  $\phi_1, \dots, \phi_n, \psi \in \mathcal{L}_P$ .*

Inference rules should be read as *whenever every formula in  $\{\phi_1, \dots, \phi_n\}$  is true, so is  $\psi$*  (as in [31, Chapter 2]). We use  $pr(\rho)$  and  $con(\rho)$  to abbreviate, respectively, the set of premises and the conclusion of  $\rho$ . The set of rules is denoted by  $\mathcal{L}_R$ .

**Definition 2.2 (Terms).** *The set of terms  $T$  is defined as  $T := \{c_\rho \mid \rho \in \mathcal{L}_R\} \cup \{c_G \mid G \subseteq Ag\} \cup \{cp_j \mid j \in Ag\}$ . It contains elements for (i) the cognitive costs of rule applications (of the form  $c_\rho$ ), (ii) cognitive costs of resolution among members of groups (of the form  $c_G$ ), (iii) cognitive capacities of agents (of the form  $cp_j$ ).*

**Definition 2.3 (Language).** *With the above in place, language  $\mathcal{L}$  is given by:*

$$\phi ::= p \mid z_1 s_1 + \dots + z_n s_n \geq z \mid \neg \phi \mid \phi \wedge \psi \mid A_j \rho \mid D_G \phi \mid \langle R_G \rangle \phi \mid \langle C, e \rangle \phi$$

where  $p \in P$ ,  $z_1, \dots, z_n \in \mathbb{Z}$ ,  $z \in \mathbb{Z}^r$ ,  $s_1, \dots, s_n \in T$ ,  $\rho \in \mathcal{L}_R$ . The dynamic operators are,  $\langle R_G \rangle$  for resolution, and  $\langle C, e \rangle$ , where  $C$  is an action model and  $e$  an

event of  $C$ . We will specify the effect of dynamic operators later when presenting the semantics; for now they should be thought as operators for communication and inference respectively.<sup>1</sup>

**Examples of Formulas.** The formula  $(cp_j \geq c_\rho) \wedge A_j\rho$  says that (i) the cognitive capacity of agent  $j$  (to which the term  $cp_j$  corresponds) is greater or equal than the cognitive cost of a rule  $\rho$  (to which the term  $c_\rho$  corresponds), and (ii) the agent  $j$  has acknowledged rule  $\rho$  as truth-preserving. Individual knowledge of an agent  $j$  is defined in terms of DK as  $K_j := D_{\{j\}}$ . A formula like  $\langle C, e \rangle K_j \phi$  says that after the event  $e$  of the action model  $C$  takes place, the agent  $j$  knows that  $\phi$ .

## 2.2 Resource-Sensitive Epistemic Models

In order to interpret these formulas, we define a resource-sensitive epistemic model and suitable model updates, induced by actions of resolution and inference, corresponding to the effect of our dynamic operators  $\langle R_G \rangle$  and  $\langle C, e \rangle$ .

Our models supplement Kripke models with *impossible worlds* and cognitive components. Impossible worlds, unlike possible ones, are not closed under logical consequence, to do justice to the fallibility of agents as real people might entertain some inconsistent/incomplete scenarios. Yet by taking reasoning steps, to the extent they can cognitively afford them, they can gradually eliminate some of them. To start with, we impose *Minimal Consistency*: we rule out explicit contradictions, in line with the literature on *Minimal Rationality* [14].

For the other components, we first need to parameterize our models by *Res*, denoting the set of resources (*time, memory, attention* etc.) we want to consider. Then  $r := |Res|$  is the number of these resources. Another parameter concerns the cognitive effort of the agents w.r.t. each resource. The *cost function*  $c : \mathcal{L}_R \cup \mathcal{P}(Ag) \rightarrow \mathbb{N}^r$  assigns a *cognitive cost* to (i) each inference rule, (ii) each group, w.r.t. each resource. That is, cost is a vector (as in [2]), used to indicate the units consumed per resource for actions of inference and resolution. We use the notation  $c_k$ ,  $k = 1, \dots, r$  to refer to the value of the  $k$ -th element of the vector and we assume that the first resource, hence the first element of the vector, concerns *time*. Concrete assignments of costs rely on empirical research. This is because the cognitive difficulty of reasoning tasks is often explained in

---

<sup>1</sup> The choice of number  $r$  is discussed in the next subsection. Formulas involving  $\vee, \rightarrow, \leq, =, -$  can be defined in terms of the rest. This is why formulas like  $cp_j \geq c_\rho$  or  $cp_j \geq c_G$  are well-formed in this language.

terms of the number and the kind of the rules that have to be applied, also considering the different response times of people in different inferences [28] and memory constraints [15].<sup>2</sup>

With the above fixed, we introduce *cognitive capacity* to the model to capture the agents' available power w.r.t. each resource. As resources are depleted while reasoning evolves, capacity will not remain constant, but it may change as a result of actions of inference or resolution, that require effort by agents uncovering new information. This is because cognitive capacity w.r.t. certain resources, like memory, is correlated with deductive reasoning performance [7]. Overall:

**Definition 2.4 (Resource-sensitive model (RSM)).** *Given parameters  $Res$  and  $c$ , a RSM is a tuple  $M := \langle W^P, W^I, \{\sim_j\}_{j \in Ag}, V_P, V_I, R, \{cp_j\}_{j \in Ag} \rangle$  where:*

- $W^P$  and  $W^I$  are sets of possible and impossible worlds, respectively.
- Each  $\sim_j$  is an epistemic accessibility relation imposed on  $W := W^P \cup W^I$ , that is, a binary relation on  $W$ .
- $V_P : W^P \rightarrow \mathcal{P}(P)$  is a valuation function assigning to each possible world, the propositional atoms that are true there.
- $V_I : W^I \rightarrow \mathcal{P}(\mathcal{L})$  is a valuation function assigning to each impossible world, the formulas (atomic or complex) that are true there.
- $R : W \times Ag \rightarrow \mathcal{P}(\mathcal{L}_R)$  is a function that assigns to each pair of a world and an agent the rules that the agent has acknowledged there.
- $cp_j \in \mathbb{Z}^r$  stands for the cognitive capacity of each agent, i.e. what  $j$  can afford w.r.t. each resource. As a convention, we will consider that time is always a resource and the first component of the vector of  $cp_j$  refers to it.

Each RSM comes parameterized by  $Res$  and  $c$ , yet we will not explicitly write them down as components of the model. This is to serve simplicity of notation but also to emphasize that these, unlike  $cp$ , are not meant to be modified in the aftermath of our actions.

**Model Conditions.** To fulfill *Minimal Consistency* we ask:  $\{\phi, \neg\phi\} \not\subseteq V_I(w)$ , for any  $w \in W^I$  and  $\phi \in \mathcal{L}$ . To ensure that acknowledged inference rules are truth-preserving, we impose *Soundness of inference rules*: for  $w \in W^P, j \in Ag: \rho \in R(w, j)$  implies  $M, w \models tr(\rho)$  where  $tr(\rho) := \bigwedge_{\phi \in pr(\rho)} \phi \rightarrow con(\rho)$ .<sup>3</sup>

<sup>2</sup> Notice that different schools (e.g. *Mental Logic* [28], *Mental Models* [24]) point at different “measures” for the difficulty of deductive tasks; still, the very observation that not all inferences require equal effort is uncontroversial. Since this debate is not settled in the empirical realm, we have not committed to any view on cost assignments. Instead, we focus on providing the machinery to embed such features in formal logical modelling.

<sup>3</sup> We focus on truth-preserving rules because we accept the factivity of knowledge. Besides, this is one of the features distinguishing knowledge from belief, according to most theories of knowledge. However, notice that other notions, like belief, could be developed through non-truth-preserving inferences as well.

It is common in EL to ask that epistemic relations are reflexive, symmetric and transitive, properties that correspond to properties of knowledge: factivity, positive and negative introspection. In what follows, we will impose reflexivity (and thus factivity); still, we abstain from assuming unlimited introspection, thus from asking that relations are symmetric and transitive. In the context of resource-bounded agents, it is reasonable to extend considerations of non-ideal performance to higher-order reasoning as well.

Before we proceed to model updates, we define the truth clauses for the static fragment, i.e.  $\mathcal{L}$  without  $\langle R_G \rangle$  and  $\langle C, e \rangle$  operators. To do that, we first need to interpret the terms in  $T$ . The intuition is that those of the form  $c_\rho$  and  $c_G$  correspond to the cognitive costs of rules and group resolution (respectively), and those of the form  $cp_j$  to the cognitive capacities of agents. This is why  $cp_j$  is used both as a model component and as a term of the language. The use can be understood from the context. Notice that our intended reading of  $\geq$  is that  $s \geq t$  iff every  $k$ -th component of  $s$  is greater or equal than the  $k$ -th component of  $t$ .

**Definition 2.5 (Interpretation of terms).** *Given a model  $M$ , the terms of  $T$  are interpreted as follows:  $c_\rho^M = c(\rho)$ ,  $c_G^M = c(G)$  and  $cp_j^M = cp_j$ .*

**Definition 2.6 (Static truth clauses).** *Take  $\sim_G := \bigcap_{j \in G} \sim_j$ , for  $G \subseteq Ag$ .*

	$M, w \models p$	<i>iff</i>	$p \in V_P(w)$
	$M, w \models z_1 s_1 + \dots + z_n s_n \geq z$	<i>iff</i>	$z_1 s_1^M + \dots + z_n s_n^M \geq z$
<i>For <math>w \in W^P</math> :</i>	$M, w \models \neg \phi$	<i>iff</i>	$M, w \not\models \phi$
	$M, w \models \phi \wedge \psi$	<i>iff</i>	$M, w \models \phi$ and $M, w \models \psi$
	$M, w \models A_j \rho$	<i>iff</i>	$\rho \in R(w, j)$
	$M, w \models D_G \phi$	<i>iff</i>	$M, u \models \phi$ for all $u : w \sim_G u$
<i>For <math>w \in W^I</math> :</i>	$M, w \models \phi$	<i>iff</i>	$\phi \in V_I(w)$

In impossible worlds, formulas are evaluated directly (i.e. not recursively) by the valuation function. Notice that the clause for  $D_G$  is given through the intersection of relations of  $G$  members (as in DEL), but it now quantifies over possible and impossible worlds, hence leaving room for deductively imperfect agents and groups. A formula is said to be *valid in a model* iff it is true at all possible worlds.

### 2.3 Resolution

We use resolution as the action that captures how information is pooled by group members, thereby enhancing the group's knowledge. As in [36], resolution is understood as publicly known private communication among members.<sup>4</sup> The resolution of group  $G$  induces a model update such that an epistemic relation

<sup>4</sup> There are alternative understandings compatible with our framework, e.g. generalizations where agents share all they know with different sets of agents [6]. This might allow to break down the effect of resolution into the incremental sharing actions of the members and study their possibly asymmetrical contribution in actualizing DK. This is left for future work.

for a member of  $G$  is the intersection of relations of the members of  $G$ , and it remains intact for the rest. Moreover resolution might come at a cost. It can be that “pooling” is effortless, e.g. because information is shared within the group for “free”. However, it can be that adopting a piece of private information through a publicly known action requires effort, e.g. because the group is too big.<sup>5</sup> One way to formally account for this effort is as follows: resolution incurs a non-zero cost on cognitive capacity for members of  $G$ , but also a cost w.r.t. *time* (and only time) for agents outside  $G$  (as time passes while  $G$  deliberates). The model update of resolution is below:

**Definition 2.7 (Resolution).** *Given RSM  $M = \langle W^P, W^I, \sim_j, V_P, V_I, R, cp_j \rangle$ , the resolution of group  $G$  produces a new RSM  $M_G := \langle W^P, W^I, \sim'_j, V_P, V_I, R, cp'_j \rangle$  where:*

$$\sim'_j = \begin{cases} \bigcap_{i \in G} \sim_i, & \text{if } j \in G \\ \sim_j, & \text{otherwise} \end{cases} \quad cp'_j = \begin{cases} cp_j - c(G), & \text{for } j \in G \\ cp_j - (c_1(G), \dots, 0), & \text{otherwise} \end{cases}$$

The conditions of RSMs are preserved by this definition. Resolution formulas are interpreted as follows. For  $w \in W$ :

$$M, w \models \langle R_G \rangle \phi \text{ iff } M, w \models (cp_i \geq c_G) \text{ for all } i \in G \text{ and } M_G, w \models \phi$$

i.e. the “precondition” of resolving knowledge among the group  $G$  is that the action is cognitively affordable to everyone in the group.

## 2.4 Inference

**Action Models.** Action models are used in DEL to represent complex informational actions [5]. They usually include (i) a set of events  $E$ , (ii) a binary relation  $\approx_j$  on  $E$  for each agent, representing her uncertainty regarding events taking place, (iii) a *precondition function*  $pre$  assigning a formula to each event, to indicate what is required for the event to occur. A common example is (*semi-*) *private announcements*, whereby only some agents find out something while the rest do not. In this attempt, we design novel action models to represent the inferential steps of agents in a multi-agent context. For example, the events in our action models can represent rule applications. They will too contain relations  $\approx_j$  and a precondition function  $pre : E \rightarrow \mathcal{L}$ . However, we need additional components to capture the effect of inferential actions on RSMs, since the latter also have additional components compared to plain Kripke models. More specifically:

- a second type of precondition  $pre\_imp : E \rightarrow \mathcal{P}(\mathcal{L})$  that indicates which formulas should be represented by the impossible worlds entertained by the agent(s) acting in an event  $e$ . The rough idea is to impose a “measure” on the impossibilities they may entertain in order to qualify for a rule-application.

<sup>5</sup> It has been argued that there are two different kinds of such informational events, “implicit” and “explicit” [9, 10]. This fits well with distinctions in the philosophical and linguistic literature [8] between *bare seeing* (“naked infinitives”) and *seeing-that*, which additionally implies epistemic awareness of the fact described.

- a *postcondition* function  $pos : Ag \times \mathcal{P}(\mathcal{L}) \times E \rightarrow \mathcal{P}(\mathcal{L})$  that will allow us to capture the effect of each event on the valuation of impossible worlds.
- a *cognitive capacity postcondition*, of the form  $pos\_cp : Ag \times \mathbb{Z}^r \rightarrow \mathbb{Z}^r$ , that will allow us to capture the effect of actions on cognitive capacities of agents.
- for notational convenience, a *label* function assigning to each event which rule, if any, is applied and who the “actors” are. For example, if event  $e_1$  stands for an application of  $\rho$  only by agent  $a_1$ , its label is  $(\rho, \{a_1\})$  indicating that the applied rule is  $\rho$  and its actor is  $a_1$ . If the event represents that nothing happens, its label is  $(\emptyset, \emptyset)$ : no inferential step occurs and (naturally) no one undertakes it. The label function is of the form  $lab : E \rightarrow (\mathcal{L}_R \cup \{\emptyset\}) \times \mathcal{P}(Ag)$ .

**Definition 2.8 (Action model for inference).** *An action model  $C$  is a tuple  $\langle E, \approx_j, pre, pre\_imp, pos, pos\_cp, lab \rangle$ , with the components as above.*

Consider the group selection task; Modus Ponens is applied by all agents, as evinced by the reported dialogues of participants, e.g. in [21, p. 237], [30, p. 15–17]. We capture this type of inferential action with the action model below:

**Inference by all ( $C_{ALL}$ ).** This action model captures that *all* agents perform the same reasoning step, the application of a rule  $\rho$ , e.g. a Modus Ponens instance. It comprises one event  $e_1$ , and clearly  $lab(e_1) = \{\rho, Ag\}$ . The precondition is that everybody knows the premises of  $\rho$ , has it available and has enough cognitive capacity to apply it. The precondition of impossibility is such that impossible worlds should at least represent the premises of  $\rho$ . The postcondition is used to show that agents can add the conclusion in their epistemic state through this rule-application, while the postcondition on capacity reduces it by the cost of  $\rho$  (Fig. 1).

$$\begin{array}{l}
 pre(e_1) = \bigwedge_{j \in Ag} \bigwedge_{\phi \in pr(\rho)} K_j \phi \wedge \bigwedge_{j \in Ag} A_j \rho \wedge \bigwedge_{j \in Ag} (cp_j \geq c_\rho) \quad pos(j, X, e_1) = X \cup \{con(\rho)\} \\
 pre\_imp(e_1) = pr(\rho) \quad pos\_cp(j, n) = n - c(\rho)
 \end{array}$$

**Fig. 1.** The action model for an inference of  $\rho$  performed by all.

But back to the group selection task: not all agents apply Modus Tollens. In many groups, only one member applies it and figures out that 7 should be turned [21, p. 238, 241]. In [30, p. 18–20], some dyads succeed because there is a member with background in logic who has the rule available and affordable and thus applies it. This is captured by another type of action model:

**Inference by some ( $C_{SOME}$ ).** It is not uncommon that only *some* agents ( $G \subset Ag$ ) perform a rule unbeknownst to agents in  $Ag \setminus G$  who do not. For simplicity, we design the action model in the case where *one* agent applies a rule  $\rho$ , but the design can be generalized to other subsets of  $Ag$ . The action model comprises two events,  $e_1$  to represent the application of the rule by  $a$  (hence,  $lab(e_1) = (\rho, \{a\})$ )



and  $e_0$  to represent that nothing happens (hence,  $lab(e_0) = (\emptyset, \emptyset)$ ). The latter is needed to capture that agents other than  $a$  are uncertain about the content of their peer's action (the rule-application). The precondition for  $e_1$  is that  $a$  knows the premises of the rule, has the rule available and has enough cognitive capacity to apply it. For  $e_0$  it is just  $\top$ , as nothing happens. The precondition of impossibility in  $e_1$  is such that impossible worlds should at least represent the premises of the rule  $\rho$ , while for  $e_0$  it is the empty set. The postcondition will be used to show that the actor can add the conclusion of  $\rho$  in her epistemic state, while nothing changes for the other agents. The cognitive postcondition is such that only the cognitive capacity of the actor is reduced by the cognitive cost of applying  $\rho$ , while for the non-actors only time is consumed (Fig. 2).

$$\begin{aligned}
 pre(e) &= \begin{cases} \bigwedge_{\phi \in pr(\rho)} K_a \phi \wedge A_a \rho \wedge (cp_a \geq c_\rho), & \text{if } e = e_1 \\ \top, & \text{if } e = e_0 \end{cases} \\
 pre\_imp(e) &= \begin{cases} pr(\rho) & \text{if } e = e_1 \\ \emptyset, & \text{if } e = e_0 \end{cases} \\
 pos(j, X, e) &= \begin{cases} X \cup \{con(\rho)\}, & \text{if } j = a, e = e_1 \\ X, & \text{otherwise} \end{cases} \\
 pos\_cp(j, n) &= \begin{cases} n - c(\rho), & \text{if } j = a \\ n - (c_1(\rho), \dots, 0), & \text{otherwise} \end{cases}
 \end{aligned}$$

**Fig. 2.** The action model for an inference of  $\rho$  performed by  $a$  unbeknownst to the rest.

**Product Models.** We now define *product models*, i.e. the model updates induced by the inferential actions. Our RSMs have additional components compared to simple Kripke models, like the set of impossible worlds and the cognitive capacity, which should be also modified according to the effect of the actions. Roughly, impossible worlds entertained by actors of inference rules can be eliminated – if their inconsistency is uncovered by applying the rule – or become enriched because, through rule applications, actors come to know the conclusion of the rule. Moreover, cognitive capacities of agents are reduced by the suitable cost. We will describe the model transformations by actions of inference, i.e. the product models, component by component. First we need certain abbreviations:

**Abbreviations.** Given a RSM  $M$  and a world  $w$  in  $W^P$  we take:

$$[\sim_j(w)] := \{u \in W^I \mid w \sim_j u\} \quad [\sim(w)]_G = \bigcup_{j \in G} [\sim_j(w)] \quad [\sim]_G = \bigcup_{u \in W^P} [\sim(u)]_G$$

These abbreviations capture, respectively, which impossible worlds are accessible from  $w$  for agent  $j$ , for group  $G$ , and the ones overall entertained by  $G$ . Given a model  $M$  and a rule  $\rho$  we also need an abbreviation to talk about impossible worlds that *will* become inadmissible, given Minimal Consistency,

once  $\rho$  is applied:  $[MC]^\rho := \{w \in W^I \mid \neg \text{con}(\rho) \in V_I(w) \text{ or } \text{con}(\rho) = \neg\phi, \text{ for some } \phi \in V_I(w)\}$ . Next, given a model  $M$  and an action model  $C$ :

$$[MC]^e = \begin{cases} [MC]^\rho \cap [\sim]_{\text{lab}_2(e)}, & \text{if } \text{lab}_1(e) = \rho \\ \emptyset, & \text{otherwise} \end{cases}.$$

This allows us to talk about the impossible worlds that will be uncovered as inadmissible by an occurrence of  $e$ . For example, if the event represents a  $\rho$ -application, then this set of worlds will contain those worlds susceptible to Minimal Consistency that are also entertained by the actors (those who do apply the rule).<sup>6</sup> The components of the product model are then built as follows:

- the new set  $(W^P)'$  consists of pairs of possible worlds and events, such that the world satisfies the precondition of the event.
- the new set  $(W^I)'$  consists of pairs of impossible worlds and events, such that the world satisfies the precondition of impossibility and it is not ruled out by Minimal Consistency. That is, if an event  $e$  represents a rule-application, the impossible worlds which are paired with it are the ones that survive the rule-application. If an impossible world lies in the epistemic state of an actor who by applying the rule unveils that she initially entertained an inconsistency, then that world will not give rise to such a pair.
- The valuation  $V_P'$  is simply  $V_P$  restricted to the surviving possible worlds.
- The valuation  $V_I'$  is given as follows with the help of the postcondition function: if the pair  $(w, e) \in (W^I)'$  lies in the epistemic state of an actor, who applies  $\rho$ , then its valuation is extended by the conclusion of  $\rho$ : the agent came to know the conclusion via the rule-application. Otherwise, the valuation should not be extended, since the epistemic states of non-acting agents should not change: *they* do not come to know the conclusion.
- $R'$  is simply  $R$  restricted to the surviving worlds.
- The new cognitive capacity is given through the capacity postcondition. That is, the capacities of non-actors remain unchanged as they did not make any cognitive effort, with the exception of time (which is consumed anyway). However, actors' capacity is reduced by the cost of the rule-application.

**Definition 2.9 (Product model).** *Let  $M$  be a RSM and  $C$  an action model. The product model  $M \otimes C$  is a tuple  $\langle (W^P)', (W^I)', \sim'_j, V_P', V_I', R', cp'_j \rangle$  where:*

$$\begin{aligned} (W^P)' &= \{(w, e) \in W^P \times E \mid M, w \models \text{pre}(e)\} \\ (W^I)' &= \{(w, e) \in W^I \times E \mid \text{pre\_imp}(e) \subseteq V_I(w) \text{ and } w \notin [MC]^e\} \\ (w, e) \sim'_j (w', e') &\text{ iff } w \sim_j w' \text{ and } e \approx_j e' \\ V_P'(w, e) &= V_P(w), \text{ for } (w, e) \in (W^P)' \\ V_I'(w, e) &= \begin{cases} \text{pos}(j, V_I(w), e) \text{ with } j \in \text{lab}_2(e), & \text{if } w \in \bigcup_{(u,x) \in (W^P)'} [\sim]_{\text{lab}_2(e)} \\ \text{pos}(j, V_I(w), e), & \text{with } j \notin \text{lab}_2(e), \text{ otherwise} \end{cases} \text{ for } (w, e) \in (W^I)' \\ R'((w, e), j) &= R(w, j), \text{ for } (w, e) \in W' \text{ where } W' = (W^P)' \cup (W^I)' \\ cp'_j &= \text{pos\_cp}(j, cp_j) \end{aligned}$$

<sup>6</sup> For example, if an impossible world  $w$  represents  $p, p \rightarrow q, \neg q$  and is entertained by all agents, and event  $e_1$  represents the application of  $MP = \{p, p \rightarrow q\} \rightsquigarrow q$ , then  $w$  will be contained in  $[MC]^{e_1}$ . This world will become inadmissible by an  $e_1$  occurrence, because its inconsistency is uncovered by the application of the rule.

Then the semantic interpretation for operators  $\langle C, e \rangle$  is given below. For  $w \in W$ :

$$M, w \models \langle C, e \rangle \phi \text{ iff } M, w \models \text{pre}(e) \text{ and } M \otimes C, (w, e) \models \phi$$

### 3 Discussion

We now see these constructions in action and discuss features of the framework.

*Example 3.1 (Dyad selection task).* For this variant of the task, we focus on two agents, each knowing the visible side of *one* card. The first ( $a_1$ ) sees the letter card  $A$ , and the second ( $a_2$ ) sees the number card 7.<sup>7</sup>

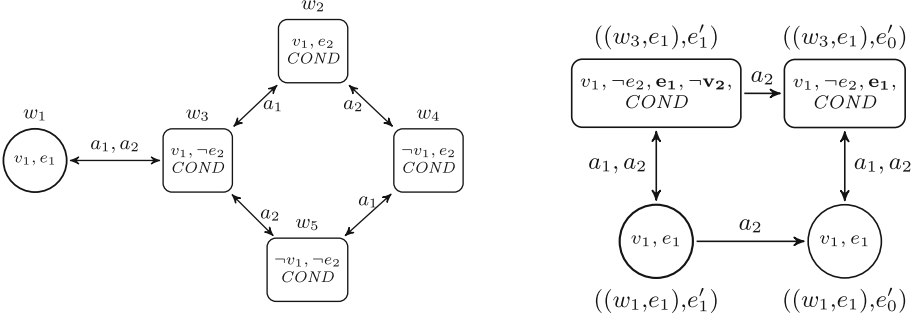
**Language.** Denote “card 1 has a vowel” with  $v_1$  and “card 1 has an even number” with  $e_1$ . Likewise,  $v_2$  (respectively,  $e_2$ ) stand for “card 2 has a vowel (even number)”. Abbreviate the formulas  $v_i \rightarrow e_i$  for  $i = 1, 2$  with  $COND$ . Also,  $MP := \{v_1 \rightarrow e_1, v_1\} \rightsquigarrow e_1$  and  $MT := \{v_2 \rightarrow e_2, \neg e_2\} \rightsquigarrow \neg v_2$ .

**Initial Model.** The model representing that  $a_1$  knows the content of the letter card and  $a_2$  knows that of the number card is Fig. 3 (left). The formulas of  $COND$  are true throughout all worlds. Since agents are fallible, at the beginning they only know what they see (the visible sides) – they have not immediately put their observations together nor have they inferred immediately what lies in the back of the cards. The impossible (incomplete) worlds representing the combinations of letter and number on the first and the second card are:

- ▶  $w_2$ : the first card depicts a vowel and the second card an even number.
- ▶  $w_3$ : the first card depicts a vowel and the second card an odd number.
- ▶  $w_4$ : the first card depicts a consonant and the second card an even number.
- ▶  $w_5$ : the first card depicts a consonant and the second card an odd number.

We draw these worlds as rectangles and write down all formulas true there, to distinguish them from the real (possible) world ( $w_1$ ), where we write the atoms that are true there, namely  $v_1, e_1$  (thus  $\neg e_2, \neg v_2$  are also true as possible worlds are maximal consistent alternatives). The epistemic relations represent the uncertainty of agents w.r.t. the card they have not seen. There are also reflexive and transitive arrows, not drawn for simplicity. Moreover, for  $Res = \{time, memory\}$ , take  $cp(a_1) = (6, 6)$ ,  $cp(a_2) = (6, 4)$ . Both agents have acknowledged  $MP$ , but only  $a_1$  has acknowledged  $MT$ . Finally,  $c(MP) = (1, 2)$ ,  $c(MT) = (3, 2)$  as  $MT$  is provably more difficult than  $MP$ , and  $c(G) = (1, 1)$ , for the cost of resolution of  $G = \{a_1, a_2\}$ .

<sup>7</sup> The framework can also be applied to the *Shadow-Box experiment* [23], investigating the synthesis of disparate points of view. For example, [1] sees learning as a social process of belief revision of interacting agents and draws connections between the Shadow-Box and variants of selection tasks whereby agents have access to different part of the world (set of cards).



**Fig. 3.** The initial model  $M$  and the updated pointed model  $(M_{fin}, ((w_1, e_1), e'_1))$

**Actions.** Afterwards, both agents share their observations. This is captured via resolution. This can be undertaken because  $cp(a_i) \geq c(G)$ , for  $i = 1, 2$ . However, it reduces capacities to  $(5, 5)$  and  $(5, 3)$  respectively. Then, all agents apply  $MP$  (captured by an action model  $C_1$ ), since they both have the rule available and affordable, in accord with the experimental dialogues [21, 30]. Their capacities become  $(4, 3)$  and  $(4, 1)$ . However, only  $a_1$  applies  $MT$ , having the rule available and affordable. This is in accord with the dialogues and captured by an action model  $C_2$ . Her capacity becomes  $(1, 1)$ , while  $a_2$ 's becomes  $(1, 1)$  too.

**Final Model.** The final pointed model is depicted in Fig. 3. We have  $M_{fin} := (M_G \otimes C_1) \otimes C_2$ , resulting from a resolution update ( $M_G$ ) and then from product updates with  $C_1$  and  $C_2$ . As a result,  $M_{fin}, ((w_1, e_1), e'_1) \models K_{a_1} e_1 \wedge K_{a_1} \neg v_2 \wedge K_{a_2} e_1 \wedge \neg K_{a_2} \neg v_2$ , so  $M, w_1 \models \langle R_G \rangle \langle C_1, e_1 \rangle \langle C_2, e'_1 \rangle (K_{a_1} e_1 \wedge K_{a_1} \neg v_2 \wedge K_{a_2} e_1 \wedge \neg K_{a_2} \neg v_2)$ .

**Further Development.** After another resolution round,  $a_2$  will also come to know  $\neg v_2$ , since she can afford *that* action (pooling information  $a_1$  derived earlier). This corresponds naturally to the dialogues in [21, p. 238–240] and [30, p. 16, 19], where the member who figures out that 7 should be turned shares the newly deduced information. Notice that  $a_2$  could use resolution, but not  $MT$ ; at the end, she did not have to apply  $MT$  herself, because her teammate did so, and all she had to do is communicate with her. Had the group not shared their information they would not have reported the correct solution; had  $a_2$  reasoned alone, her resources would not have allowed her to reach the solution. This illustrates one way in which reasoning in groups facilitates performance in tasks that are more challenging on the individual level.

Our framework models the crucial interplay of resolution and inference, also evident in tasks like the *Shadow-Box experiment* or interdisciplinary research itself. One member might provide input information and another the means (e.g. a proof strategy) to reach a result that would not have been reached if members worked alone. Scientific quests largely depend on gathering suitable information and deriving more on its basis to actualize scientific potential. However, this process is effortful; resolving and deducing comes with a cognitive cost.

We present some validities and invalidities, as further basis of discussion. The full-fledged proofs for all theorems that follow are omitted for brevity.

**Theorem 3.1 (Some validities).**  $\models D_G \langle R_G \rangle \phi \leftrightarrow \langle R_G \rangle E_G \phi$  where  $E_G := \bigwedge_{j \in G} K_j$

$$\models \bigwedge_{j \in Ag} \bigwedge_{\phi \in pr(\rho)} K_j \phi \wedge \bigwedge_{j \in Ag} A_j \rho \wedge \bigwedge_{j \in Ag} (cp_j \geq c_\rho) \rightarrow \langle C_{ALL}, e_1 \rangle E_{Ag} con(\rho) \text{ where } lab_1(e_1) = \rho$$

$$\models \bigwedge_{\phi \in pr(\phi)} D_G \langle R_G \rangle \phi \wedge A_j \rho \wedge (cp_j \geq c_G + c_\rho) \rightarrow \langle R_G \rangle \langle C_{SOME}, e_1 \rangle K_j con(\rho) \text{ where } j \in G, lab_1(e_1) = \rho, j \in lab_2(e_1)$$

$$\models \bigwedge_{\phi \in pr(\phi)} D_{Ag} \langle R_{Ag} \rangle \phi \wedge \bigwedge_{j \in Ag} A_j \rho \wedge \bigwedge_{j \in Ag} (cp_j \geq c_{Ag} + c_\rho) \rightarrow \langle R_{Ag} \rangle \langle C_{ALL}, e_1 \rangle E_{Ag} con(\rho) \text{ where } lab_1(e_1) = \rho$$

The first validity pertains to the effect of resolution on the understanding of DK (as in [36]) showing that after a group resolves their knowledge,  $\phi$  is known by its members. The second captures the effect of actions of inference. The agents do not immediately know all logical consequences of their knowledge: they have to undertake effortful reasoning steps. The other validities encapsulate the interplay of communication and inference: once members resolve their knowledge and come to know the premises, then those who apply the rule, come to know the conclusion as well. Contrary to these, we escape features of idealized agents:

**Theorem 3.2 (Some invalidities).**

$$\not\models D_G \phi \rightarrow \langle R_G \rangle E_G \phi \quad \not\models D_G \langle R_G \rangle \phi \rightarrow \langle R_G \rangle E_G E_G \phi \quad \not\models \bigwedge_{\phi \in pr(\rho)} D_G \phi \rightarrow D_G con(\rho)$$

$$\not\models \bigwedge_{\phi \in pr(\rho)} K_j \phi \wedge (cp_j \geq c_\rho) \wedge A_j \rho \rightarrow \langle C_{SOME}, e_1 \rangle K_j con(\rho) \text{ where } \rho = lab_1(e_1) \text{ and } j \notin lab_2(e_1)$$

The first invalidity unveils the problem behind the traditional understanding of DK (recall Sect. 1) and it is also identified in [36]. The second invalidity shows that higher orders of knowledge require additional reasoning steps that might not follow from attaining mutual knowledge alone. This departs from literature viewing actualizations of DK as CK, because our attempt focuses on *resource-boundedness*: higher-order knowledge, and hence CK, need extra effort that should not be taken for granted. The third invalidity shows that DK is not logically closed, therefore actualizing knowledge of logical consequences is not trivial. The fourth invalidity shows that non-acting agents might not come to know logical consequences, even if some of their peers do. This might need yet another round of resolution, exemplifying the continuous and resource-consuming interplay of communication and inference that takes place in reality when non-ideal groups deliberate.

The theorems illustrate how DK is actualized by non-ideal agents. The use of impossible worlds as “witnesses” of agents’ fallibility did not result in a trivialized system where anything goes, due to the dynamics: agents come to know more via resolution and inference, provided that they can. Specifying these steps, monitoring their interplay and the effort they require allows us to track to which extent a group realizes its potential, instead of pre-setting an arbitrary bound.

This provides the “bridge” between the implicit notion of DK and the explicit knowledge real groups can achieve.

**Related Work.** Comparisons with related work concern: (a) the inferential aspect of knowledge, (b) the communicative aspect of actualizing group potential.

Consider aspect (a): this approach contributes to impossible-worlds semantics used against omniscience [27] by adding *dynamics* that avoid the extreme of trivial logics. Other approaches discern implicit (omniscient) and explicit (omniscience-free) attitudes through a syntactic “filter” (like an *awareness function* [17]). However, forms of the problem may persist and it not clear how resource-boundedness could fit in this picture. Closer to our view are [13, 16, 31, 35], yet our elaborate specification of reasoning processes is important in bridging logic with empirical facts, because these usually pertain to the difficulty of *individual* inference rules [14, 24, 28]. This also discerns our framework from others with multiple, non-ideal agents [3, 4] that too study the effect of communication and inference in multi-agent settings.

Consider aspect (b): [6, 36] propose actions following the observation in [11]: it takes more than communication of formulas expressible in the standard languages to actualize DK. Our resolution action is based on [36]’s, and is similar to a special case of [6]’s *tell-all-you-know* actions, and to [11]’s *communication core*. While this wider variety of actions is compatible with the framework,<sup>8</sup> our dynamics is tailored to bounded agents, explaining how far group reasoning can go. It is precisely this difference in scope that justifies our divergence from studying actualizations of DK as CK. It would also be interesting to connect this resource-sensitive attempt and another generalization of operations for pooling information given in [26]: the authors provide an epistemic modality relative to structured communication scenarios as an alternative to distributed knowledge.

Overall, our approach addresses the problem of logical omniscience, in a multi-agent context and in agreement with experimental results and philosophical proposals (e.g. towards a *theory of feasible inferences* [14]). Departing from this well-known problem, this approach demarcates the communicative and inferential actions underlying whether and how DK is actualized. As [10] argues, information goes hand in hand with the processes that create, modify, and convey it; this analysis naturally applies to deliberating groups, and importantly, to resource-bounded ones.

## 4 Reduction and Axiomatization

In this section, we reduce RSMs to possible-worlds structures with syntactic functions, resembling awareness structures [17]. A reduction from impossible worlds to syntactic structures follows the converse direction to [33], showing how various

---

<sup>8</sup> Their model transformations could be captured as in the aforementioned attempts (epistemic relation-wise), accompanied by the incurrance of a cost to those receiving the information.

structures, like awareness ones, can be reduced to impossible-worlds models validating precisely the same formulas (given a fixed background language). Besides, the division of responses to logical omniscience into syntactic and semantic ones is common in the literature. Syntactic approaches are claimed to lack the elegance of semantic (impossible-worlds) ones, yet the latter’s semantic rules do not adequately capture intuitions about knowledge formation [17]. Our framework is a semantic one, using impossible worlds to do justice to multiple non-ideal agents but nonetheless preserves explanatory power since agents can engage in knowledge-refining actions. While the model and its actions accommodate these intuitions, the reduction is instrumental in providing a sound and complete logic, as it allows for the use of standard DEL techniques. In this way, we wish to harvest both the benefits of impossible-worlds semantics and the more convenient technical treatment of syntactic approaches.

An outline of the reduction is as follows. First, we focus on the static part and we show that the effect of impossible worlds in the interpretation of  $D_G$  can be captured in a possible-worlds model, provided that suitable syntactic functions are introduced. Second, we obtain a sound and complete static axiomatization, through modal logic techniques. Third, we move to the dynamics. We explain why the common DEL procedure of giving *reduction axioms* is not straightforward but also how this issue can still be overcome.

#### 4.1 Reduction and Static Axiomatization

**Reduced (Static) Language.** We fix an appropriate language  $\mathcal{L}_{red}$  as the “common ground” to show that the reduction is successful, i.e. the same formulas are valid under the original and the reduced models. Take  $\sim_G(w) := \{u \in W \mid w \sim_G u\}$ , which denotes the set the truth clause for  $D_G$  quantifies over. Auxiliary operators ( $L_{D_G}, I_{D_G}$ ) are then introduced to the static fragment of  $\mathcal{L}$  to discern (syntactically) the effect of quantifying over (im)possible worlds in  $D_G$ -interpretations. Their semantic interpretations are given below. For  $w \in W^P$ :

$$\begin{aligned} M, w \models L_{D_G}\phi & \text{ iff } M, u \models \phi \text{ for all } u \in W^P \cap \sim_G(w) \\ M, w \models I_{D_G}\phi & \text{ iff } M, u \models \phi \text{ for all } u \in W^I \cap \sim_G(w) \end{aligned}$$

These essentially help us break down the  $D_G$  operator. We also use  $\perp$  as an auxiliary element of  $\mathcal{L}_{red}$ , that is never true in any world.

**Building the Reduced Model.** Towards interpreting the auxiliary operators  $I_{D_G}$  in a reduced model, we construct *awareness-like functions*:

- $I_{D_G} : W^P \rightarrow \mathcal{P}(\mathcal{L})$  such that  $I_{D_G}(w) = \bigcap_{v \in W^I \cap \sim_G(w)} V_I(v)$ . Intuitively,  $I_{D_G}$  takes a possible world and yields the set of formulas true at all impossible worlds in its quantification set (the set of worlds  $D_G$  quantifies over).

**Definition 4.1 (Awareness-like structure (ALS)).** *Given  $M = \langle W^P, W^I, \sim_j, V_P, V_I, R, cp_j \rangle$ , its ALS (reduced model) is  $\mathbf{M} := \langle W, \sim_j^r, V, R, cp_j, I_{D_G} \rangle$  with:*

$$\begin{array}{ll}
 W = W^P & w \sim_j^r u \text{ iff } w \sim_j u, \text{ for } w, u \in W \\
 R(w, j) = R(w, j) \text{ for } w \in W & cp_j \text{ is as in the original}
 \end{array}
 \quad
 \begin{array}{l}
 V(w) = V_P(w) \text{ for } w \in W \\
 I_{D_G} \text{ as explained before}
 \end{array}$$

The clauses based on ALSs are such that the  $I_{D_G}$ -operators are interpreted via the awareness-like functions. Due to the construction of awareness-like functions, *Minimal Consistency* is inherited by the reduced model: for no  $w \in W$ ,  $G \subseteq Ag$ , is it the case that  $\{\phi, \neg\phi\} \subseteq I_{D_G}(w)$ . *Soundness of inference rules* is also clearly preserved. Moreover, take  $\sim_j^r(w) := \{u \in W \mid w \sim_j^r u\}$  now based on the new ordering  $\sim_j^r$ . The interpretation of terms is as in the original, since the values of capacities and costs are unchanged. The semantic clauses, based on  $\mathbf{M}$ , are standard for the boolean connectives. The remaining:

$$\begin{array}{ll}
 \mathbf{M}, w \models p \text{ iff } p \in V(w) & \mathbf{M}, w \models L_{D_G}\phi \text{ iff } \mathbf{M}, u \models \phi \text{ for all } u \in \bigcap_{j \in G} \sim_j^r(w) \\
 \mathbf{M}, w \models z_1s_1 + \dots + z_ns_n \geq z \text{ iff } z_1s_1^{\mathbf{M}} + \dots + z_ns_n^{\mathbf{M}} \geq z & \mathbf{M}, w \models I_{D_G}\phi \text{ iff } \phi \in I_{D_G}(w) \\
 \mathbf{M}, w \models A_j\rho \text{ iff } \rho \in R(w, j) & \mathbf{M}, w \models D_G\phi \text{ iff } \mathbf{M}, w \models L_{D_G}\phi \text{ and } \mathbf{M}, w \models I_{D_G}\phi
 \end{array}$$

We now show that the definition of the ALSs indeed fulfills its purpose:

**Theorem 4.1 (Reduction).** *Given a RSM  $M$ , let  $\mathbf{M}$  be its ALS. Then  $\mathbf{M}$  is a reduction of  $M$ , i.e. for any  $w \in W^P$  and formula  $\phi \in \mathcal{L}_{red}$ :  $M, w \models \phi$  iff  $\mathbf{M}, w \models \phi$ .*

*Proof.* The proof goes by induction on the complexity of  $\phi$ .

Based on this, we provide the static axiomatization:

**Definition 4.2 (Static axiomatization).**  *$\Lambda$  is axiomatized by Table 1 and the rules Modus Ponens, Necessitation $_{D_G}$  (from  $\phi$ , infer  $L_{D_G}\phi$ ).*

**Table 1.** The static axiomatization

<i>PC</i>	All instances of classical propositional tautologies
<i>Ineq</i>	All instances of valid formulas about linear inequalities
<i>D-Distribution</i>	$L_{D_G}(\phi \rightarrow \psi) \rightarrow (L_{D_G}\phi \rightarrow L_{D_G}\psi)$
<i>D-factivity</i>	$L_{D_G}\phi \rightarrow \phi$
<i>D-Monotonicity</i>	$L_{D_G}\phi \rightarrow L_{D_B}\phi$ , if $G \subseteq B$ $I_{D_G}\phi \rightarrow I_{D_B}\phi$ , if $G \subseteq B$
<i>Minimal Consistency</i>	$I_{D_G}\perp \vee (\neg(I_{D_G}\phi \wedge I_{D_G}\neg\phi))$
<i>Soundness of inference rules</i>	$A_j\rho \rightarrow tr(\rho)$
<i>Reducing <math>D_G</math></i>	$D_G\phi \leftrightarrow L_{D_G}\phi \wedge I_{D_G}\phi$

*Ineq*, described in [19], is introduced to account for the linear inequalities. The axioms for  $L_{D_G}$  mimic the behaviour of  $D_G$ -involving axioms in the standard logics with DK [18, 20, 22] only now using the auxiliary operator quantifying over the possible worlds alone. *Soundness of inference rules* and *Minimal Consistency* take care of the respective model conditions. Finally, the last axiom reduces  $D_G$  in terms of the corresponding auxiliary operators.



**Theorem 4.2** (*A soundness/completeness*). *A is sound and complete w.r.t. ALSs.*

*Proof.* Soundness follows from the validity of axioms. Completeness follows [22, p. 65], as the crucial difference (auxiliary operators interpreted through syntactic functions) is accommodated from the construction of a suitable canonical model.

## 4.2 Dynamic Axiomatization

Moving to the dynamic part, consider a RSM  $M$  and its reduced ALS  $\mathbf{M}$ . If an update, either of resolution or inference, takes place, then we get an updated  $M'$  and thus an updated ALS  $\mathbf{M}'$  corresponding to it. We observe that  $\mathbf{M}'$  is such that an updated awareness-like function  $I'_{D_G}$  is given in terms of  $I_{D_G}$ , i.e the awareness-like function in  $\mathbf{M}$ . That is, the new values are set expressions of the original ones. We present the updated functions below.<sup>9</sup>

After resolution of  $G$ :  $I'_{D_B}(w) = \bigcap_{u \in (W^I) \cap \sim'_B(w)} V_I(u) = \begin{cases} I_{D_{G \cup B}}(w), & \text{if } G \cap B \neq \emptyset \\ I_{D_B}(w), & \text{if } G \cap B = \emptyset \end{cases}$

After  $C_{ALL}$ :  $I'_{D_G}(w, e_1) = \bigcap_{(w', e') \in (W^I)' \cap \sim'_G(w, e_1)} V'_I(w', e') = \begin{cases} I_{D_G}(w) \cup \{con(\rho)\}, & \text{if } (W^I)' \neq \emptyset \\ I_{D_G}(w) \cup \overline{I_{D_G}(w)}, & \text{if } (W^I)' = \emptyset \end{cases}$

After  $C_{SOME}$ , we have the cases below regarding  $e_1$  and  $e_0$ :

$$I'_{D_G}(w, e_1) = \begin{cases} I_{D_G}(w), & \text{if } (W^I)' \cap \sim'_G(w, e_1) \neq \emptyset \text{ and } a \notin G \\ I_{D_G}(w) \cup \{con(\rho)\}, & \text{if } (W^I)' \cap \sim'_G(w, e_1) \neq \emptyset \text{ and } a \in G \\ I_{D_G}(w) \cup \overline{I_{D_G}(w)}, & \text{if } (W^I)' \cap \sim'_G(w, e_1) = \emptyset \end{cases} \quad I'_{D_G}(w, e_0) = I_{D_G}(w)$$

In DEL, it is common to provide *reduction axioms* for the dynamic operators, in our case, for  $\langle R_G \rangle$  and  $\langle C, e \rangle$ . However, reducing dynamic formulas involving the auxiliary operator  $I_{D_G}$  (thus  $D_G$  too) cannot be straightforwardly obtained because the new sets obtained through the update of  $I_{D_G}$  cannot be described by means of the static language alone. Similar problems are encountered in [31, Chapter 5]; in that single-agent framework, there are syntactic functions which are expanded after certain actions. The focus is on actions that give rise to syntactic functions which are structured expressions of the original ones, in turn treatable with a specific static language. We follow a similar procedure, tailored to *our* syntactic functions  $I_{D_G}$ . This is because, as shown above, the updated values are too given in terms of the original ones, reflecting the refinement induced by each action. Just to sketch the idea, as in [31], we extend the static language, essentially re-expressing the auxiliary operators as set-expression operators, and provide reduction axioms that yield a full sound and complete axiomatization.

For reasons of brevity, we cannot present the full-fledged procedure and the reduction axioms here. Some remarks to give a flavor of the more important reduction axioms: for inequalities, they reflect, with the help of abbreviations, the resource consumption each action induces; for  $L_G$  operators, they reflect that these operators behave as  $D_G$  does in standard DEL; for  $I_G$ , making crucial use of the set-expression operators, they reflect that the awareness-like functions are updated in a principled way: as *specific* set expressions of the original ones.

<sup>9</sup> We get these results using Definition 2.7, Definition 2.9, Definition 4.1.

## 5 Conclusions and Future Work

The EL modelling of unbounded agents has repercussions for group reasoning and DK is instrumental in illustrating this because it presupposes that agents can undertake unlimited actions of communication and inference. We looked into actualizations of DK under bounded resources, using RSMs and actions for communication and inference. The combination of impossible-worlds semantics and action models might be of independent interest given the former's use in areas beyond epistemic logic and the latter's popularity in the study of multi-agent dynamics. We furthermore showed that our models can be reduced to syntactic structures. In doing so, we confirmed a pattern observed in the omniscience literature and offered a useful detour for a sound and complete logic.

One direction for future work concerns non-ideal higher-order reasoning, and hence connections of DK and CK. As with deductive reasoning, we envisage the introduction of effortful steps for introspection and reasoning about other agents, and the use of experimental results showing that groups usually act on a large, but finite, degree of mutual knowledge as if they had CK. On another note, group reasoning, in this attempt, can be better than individual in ways that agree with the distribution of skills observed in [21, 30] and the view that at the upper limit groups perform as their best member [25]. However, the former also emphasize the facilitative effect of dialogue in group performance, which may be captured via a combination of RSMs with dialogical/inquisitive models.

**Acknowledgments.** Thanks to the audience and reviewers of DaLí 2020 for their remarks, as well as to Sonja Smets and the members of the “Logic of Conceivability” seminar. Anthia Solaki is supported by the Dutch Organisation for Scientific Research (“PhDs in the Humanities”, grant No. 322-20-018).

## References

1. Addis, T., Gooding, D.: Learning as collective belief-revision: simulating reasoning about disparate phenomena. In: AISB Symposium, pp. 6–9 (1999)
2. Alechina, N., Logan, B., Nga, N.H., Rakib, A.: A logic for coalitions with bounded resources. In: IJCAI, pp. 659–664 (2009)
3. Alechina, N., Logan, B., Nguyen, H.N., Rakib, A.: Reasoning about other agents' beliefs under bounded resources. In: Meyer, J.-J.C., Broersen, J. (eds.) KRAMAS 2008. LNCS (LNAI), vol. 5605, pp. 1–15. Springer, Heidelberg (2009). [https://doi.org/10.1007/978-3-642-05301-6\\_1](https://doi.org/10.1007/978-3-642-05301-6_1)
4. Balbiani, P., Fernández-Duque, D., Lorini, E.: The dynamics of epistemic attitudes in resource-bounded agents. *Studia Logica* **107**(3), 457–488 (2018). <https://doi.org/10.1007/s11225-018-9798-4>
5. Baltag, A., Renne, B.: Dynamic epistemic logic. In: Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab. Stanford University (2016)
6. Baltag, A., Smets, S.: Learning what others know. In: LPAR, pp. 90–119 (2020)
7. Bara, B.G., Bucciarelli, M., Johnson-Laird, P.N.: Development of syllogistic reasoning. *Am. J. Psychol.* **108**(2), 157–193 (1995)
8. Barwise, J., Perry, J.: *Situations and Attitudes*. MIT Press, Cambridge (1983)

9. van Benthem, J.: Merging observation and access in dynamic epistemic logic. *Stud. Logic* **1**, 1–16 (2008)
10. van Benthem, J.: 'Tell it like it is': information flow in logic. *J. Peking Univ. (Human. Soc. Sci. Ed.)* **1**, 80–90 (2008)
11. van Benthem, J.: *Logical Dynamics of Information and Interaction*. CUP, Cambridge (2011)
12. Berto, F., Jago, M.: *Impossible Worlds*. Oxford University Press, Oxford (2019)
13. Bjerring, J.C., Skipper, M.: A dynamic solution to the problem of logical omniscience. *J. Philos. Logic* **48**, 501–521 (2018)
14. Chermiak, C.: *Minimal Rationality*. Bradford Book. MIT Press, Cambridge (1986)
15. Cowan, N.: The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* **24**, 87–114 (2001). <https://doi.org/10.1007/s10992-018-9473-2>
16. Elgot-Drapkin, J., Kraus, S., Miller, M., Nirkhe, M., Perlis, D.: Active logics: a unified formal approach to episodic reasoning (1999)
17. Fagin, R., Halpern, J.Y.: Belief, awareness, and limited reasoning. *Artif. Intell.* **34**(1), 39–76 (1987)
18. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.Y.: *Reasoning About Knowledge*. MIT Press, Cambridge (1995)
19. Fagin, R., Halpern, J.Y.: Reasoning about knowledge and probability. *J. ACM* **41**(2), 340–367 (1994)
20. Fagin, R., Halpern, J.Y., Vardi, M.Y.: What can machines know? On the properties of knowledge in distributed systems. *J. ACM* **39**(2), 328–376 (1992)
21. Geil, D.M.M.: Collaborative reasoning: evidence for collective rationality. *Think. Reason.* **4**(3), 231–248 (1998)
22. Gerbrandy, J.D., et al.: *Bisimulations on planet Kripke*. Inst. for Logic, Language and Computation, Univ. van Amsterdam (1999)
23. Gruber, H.: The cooperative synthesis of disparate points of view. *Leg. Solomon Asch: Essays Cogn. Soc. Psychol.* **1**, 143–158 (1990)
24. Johnson-Laird, P.N., Byrne, R.M., Schaeken, W.: Propositional reasoning by model. *Psychol. Rev.* **99**(3), 418–439 (1992)
25. Laughlin, P.R., Ellis, A.L.: Demonstrability and social combination processes on mathematical intellectual tasks. *J. Exp. Soc. Psychol.* **22**(3), 177–189 (1986)
26. Punčochář, V., Sedlár, I.: Substructural logics for pooling information. In: Baltag, A., Seligman, J., Yamada, T. (eds.) *LORI 2017*. LNCS, vol. 10455, pp. 407–421. Springer, Heidelberg (2017). [https://doi.org/10.1007/978-3-662-55665-8\\_28](https://doi.org/10.1007/978-3-662-55665-8_28)
27. Rantala, V.: Impossible worlds semantics and logical omniscience. *Acta Philosophica Fennica* **35**, 106–115 (1982)
28. Rips, L.J.: *The Psychology of Proof: Deductive Reasoning in Human Thinking*. MIT Press, Cambridge (1994)
29. Roelofsen, F.: Distributed knowledge. *J. Appl. Non-Classical Log.* **17**(2), 255–273 (2007)
30. Trognon, A., Batt, M., Laux, J.: Why is dialogical solving of a logical problem more effective than individual solving?: A formal and experimental study of an abstract version of Wason's task. *Lang. Dialogue* **1**(1), 44–78 (2011)
31. Velázquez-Quesada, F.R.: Small steps in dynamics of information. Ph.D. thesis, Universiteit van Amsterdam (2011)
32. Verbrugge, R.: Logic and social cognition. *J. Philos. Logic* **38**(6), 649–680 (2009). <https://doi.org/10.1007/s10992-009-9115-9>

33. Wansing, H.: A general possible worlds framework for reasoning about knowledge and belief. *Stud. Logica* **49**(4), 523–539 (1990). <https://doi.org/10.1007/BF00370163>
34. Wason, P.C.: Reasoning. In: Foss, B. (ed.) *New Horizons in Psychology*, pp. 135–151. Penguin Books, Harmondsworth (1966)
35. Ågotnes, T., Walicki, M.: Syntactic knowledge: a logic of reasoning, communication and cooperation. In: *Proceedings of the Second European Workshop on Multi-Agent Systems (EUMAS)*, Barcelona, Spain (2004)
36. Ågotnes, T., Wang, Y.N.: Resolving distributed knowledge. *Artif. Intell.* **252**, 1–21 (2017)