

# Group Knowledge in Interrogative Epistemology

A. Baltag, R. Boddy and S. Smets  
Institute for Logic, Language and Computation.  
University of Amsterdam

## Abstract

In this paper we formalize an approach to knowledge that we call Interrogative Epistemology, in the spirit of Hintikka’s “interrogative model” of knowledge. According to our approach, *an agent’s knowledge is shaped and limited by her interrogative agenda* (as defined by her fundamental questions or “epistemic issues”). The dynamic correlate of this postulate is our Selective Learning principle: the agent’s agenda limits her potential for knowledge-acquisition. Only meaningful information, that is relevant to one’s issues, can really be learnt. We use this approach to propose a new perspective on *group knowledge*, understood in terms of the *epistemic potential of a group* of agents: the knowledge that the group may come to possess in common (and thus act upon in a coordinated manner) after all members share their individual information. We argue that the standard notions of group knowledge studied in the literature, ranging from distributed knowledge to common knowledge, do not give us a good measure of a group’s epistemic potential. Common knowledge is too weak and too “static”, focusing on what the agents can coordinate upon only based on their actual, current knowledge (without any intra-group communication), thus disregarding testimonial knowledge. In contrast, the concept of distributed knowledge is too strong, being based on the assumption that agents can completely internalize all the testimonial evidence received from others, irrespective of the limitations posed by their own interrogative agendas. We show that a group’s true epistemic potential typically lies in between common knowledge and distributed knowledge. We propose a logical formalization of these concepts, which comes together with a complete axiomatization, and we use this setting to explain both the triumphs and the failures of collective knowledge, treating examples that range from “collective scientific knowledge” (Wray, 2007; Rolin, 2008; Fagan, 2012) to the so-called “curse of the committee”.

## 1 Introduction

The idea that groups can be treated as collective agents capable of knowledge and beliefs has gained increasing attention (List et al., 2011; List, 2005; Gilbert, 1987; Wray, 2007). Issues such as the social dimensions of knowledge, the various types and varieties of “collective knowledge”, the epistemology of testimony, the truth-tracking value of democracy etc., which were traditionally confined to Social Sciences, have recently become subject of philosophical reflection, as part of Social Epistemology (Goldman, 2004; Goldman and Whitcomb, 2011).

In a parallel evolution, logicians and computer scientists have perfected new formal tools for reasoning about knowledge and its dynamics. Various formal notions of group knowledge have been formalized by epistemic logicians, e.g. ‘mutual knowledge’, ‘distributed knowledge’ and ‘common knowledge’ (Fagin et al., 2003; Baltag et al., 2008; Barwise, 1988; Lewis, 1969), while Dynamic Epistemic Logic focused on the underlying dynamics of testimony, communication, deliberation and other social-informational interactions (Baltag et al., 1998; van Ditmarsch et al., 2007; Baltag et al., 2008; van Benthem, 2011). Yet, despite the increasing interest in the logical understanding of various forms of group knowledge, until now formal logical approaches to this topic have remained largely disconnected from philosophical discussions on collective knowledge.

In this paper, we present a formal analysis of the epistemic potential of a group, as motivated by the philosophical study of ‘collective scientific knowledge’ (Wray, 2007; Rolin, 2008; Fagan, 2012). We view group knowledge as a collective attitude, which has at least the essential feature that it is acquired through goal directed inquiry and information pooling. Our focus in this matter lies on the group’s ability to combine the knowledge that is present within the group after all empirical inquiry has already taken place. A group’s epistemic potential may depend on a number of factors, including the structure of the group and of its communication network, the degree of mutual trust etc. Here, we limit our attention to one of these aspects, namely the “interrogative agendas” of the group members, comprising their *epistemic issues*: the *fundamental questions* that they aim to address and get resolved.

By highlighting the role of questions and answers, we stay close to Hintikka’s interrogative approach to epistemology (Hintikka, 2007). We analyze known propositions—both individual and collective—as answers to an agent’s questions. To formalize ‘potential group knowledge’, we use logic and draw on resources from formal epistemology and formal approaches to modelling questions. Following (van Benthem and Minica, 2012; Minica, 2008) we add an *issue relation* to the standard models for knowledge from epistemic logic, and extend this language further with collective epistemic modalities. Using these formal tools, we develop a logical framework for group knowledge that internalizes the dynamic properties of potential knowledge acquisition through inter-agent communication, and that can be used to address issues regarding the relation of group knowledge to individual knowledge and to group decision-making.

## 1.1 Questions and Knowledge Acquisition

“All knowledge involves a question. To know is to know the answer” (Schaffer 2007: 401)

The inquiry-based approach in Epistemology amounts to shifting the focus from knowledge *possession* to knowledge *acquisition*. The shift, which fits well with the “dynamic turn” in Logic championed by the Dutch school (van Benthem, 2011) and with the learning-theoretic approach to epistemology (Kelly, 1996), is also closely aligned with the “inquisitive” move in Logic (Groenendijk, 1999, 2007; Groenendijk et al., 2009; Ciardelli and Roelofsen, 2009). This trend is very much in the spirit of the Interrogative Model of Inquiry initiated by Jaakko Hintikka in the 1970’s, and later developed by him into a ‘Socratic epistemology’ (Hintikka, 2007), which took the act of questioning as fundamental:

“Its basic idea is the same as that of the oldest explicit form of reasoning in philosophy, the Socratic method of questioning or *elenchus*. In it, all new information enters into an argument or a line of reasoning in the form of answers to questions that the inquirer addresses to a suitable source of information.” (Hintikka, 2007, p18)

Hintikka here adopts a broad view on what establishes an answer to a question by including as “sources of answers” witnesses providing testimonies, databases, observations, experiments, and also memory and tacit knowledge. Central to his approach is that “all information used in an argument must be brought in as an answer to a question” (Hintikka, 2007, p19).

But when taking interrogative inquiry as the basis of epistemology, what is the role that knowledge plays? Hintikka discusses this point in (Hintikka, 2007, p26) where he explains that “questions of knowledge do not play any role in the questioning process itself, only in evaluating its results”. He takes the process of inquiry (rather than any attempt to define the concept of knowledge) to be the main concern in epistemology:

“The criteria of knowledge concern the conditions on which the results of epistemological inquiry can be relied on as a basis of action” (Hintikka, 2007, p30).

In line with Hintikka’s model of inquiry as a goal-directed process, we mention the learning-theoretic approach to epistemology of (Kelly, 1996), who stresses the role of fundamental questions as long-term goals directing the learning process. Similarly, Interrogative Belief Revision stresses the role played by the agent’s research interests (her ‘research agenda’) in guiding and informing her belief-revision strategy (Olsson and Westlund, 2006). Schaffer’s ‘contrastive’ theory of knowledge (Schaffer, 2005, 2007; Aloni et al., 2013) is also based on the idea that all knowledge implies a question. According to this account, the knowledge that an agent can acquire is dependent on her questions, which determine what alternative states (or answers) she must eliminate in order to acquire knowledge (Schaffer, 2005, 2007; Aloni et al., 2013). Schaffer’s stance can be seen as a version of Dretske’s Relevant Alternatives approach to knowledge: one knows  $p$  only as an answer to a question, and this is known only by contrasting it with the other possible answers.

## 1.2 Questions as Epistemic Filters: the ‘Relevant Distinctions’ Theory of Knowledge

While agreeing with Hintikka, Kelly, Schaffer, Olsson and Westlund etc. on the general principle that all knowledge is question-relative, our conception gives a particular twist to this idea. For us, *questions act as “epistemic filters” that structure and limit what one can know*. So, although our conception can be said to perfectly fit the above-mentioned Schaffer quote when taken literally, this is only done by adopting an interpretation of that quote that is different from Schaffer’s own.

Our starting point is the same as Hintikka’s: inquiry is driven by questions. An agent’s questions shape her knowledge and guide her learning. But as a consequence, they also impose limits on the agent’s knowledge and ability to learn. In “normal” situations, agents can only learn information that fits their agendas: according to this view, *all we (can) know are answers to our own questions*. As a consequence, the agents’ divergent interrogative agendas can limit a group’s epistemic potential.

Our conception of the epistemic role of questions can be best understood as a “*Relevant Distinctions*” approach to knowledge (rather than focusing on Relevant Alternatives, like Dretske and Schaffer). Information that does not fit with the agent’s questions involves irrelevant or meaningless distinctions, and thus cannot be learned without breaking (or changing) the agent’s conceptual space. This would involve a change of paradigm (essentially, the raising of new fundamental questions and the abandoning of some old questions). In most of this paper, we deal with “normal” epistemic situations, in which the interrogative agenda remains fixed. In “normal” science, the investigator sticks with the same paradigm, based on the same fundamental issues, and thus she may be unable (or uninterested) to process information that is not relevant to her questions. This is the essence of our Selective Learning Principle: *One cannot find what one is not looking for*.

Our concept of potential group knowledge is formally related to other group attitudes, such as ‘distributed knowledge’ and ‘common knowledge’, that have been studied in epistemic logic. But in this paper we argue that these standard notions do not give us a good measure of a group’s epistemic potential. Common knowledge captures one feature of group knowledge, namely the fact that group knowledge is what allows coordinated action. Hence the stress on its “shared” nature: it has to be held *in common*. But common knowledge is too “static”, focusing on what the agents can coordinate upon only based on their *actual, current knowledge (without any intra-group communication)*, thus disregarding testimonial knowledge. This view gives a too modest assessment of a group’s epistemic powers: according to the common-knowledge account, a group knows almost always *less* than its members! So the concept of common knowledge is *too weak* to capture a group’s knowledge, giving only a *lower limit* for a group’s potential. In contrast, distributed knowledge (defined as the result of pooling together all the individual knowledge of the group members, and closing under logical consequence) may seem at first sight to accurately capture the potential knowledge of a group. But this is based on the unrealistic assumption that agents can completely internalize all the testimonial evidence received from others, irrespective of the limitations posed by their own interrogative agendas. This assumption contradicts the principle of Selective Learning (as well as everyday experience). In typical situations, not all the distributed knowledge can actually be “internalized” by all the members of the group. This is due to the discrepancy between different agents’ interrogative agendas. Hence, the concept of distributed knowledge is *too strong*, giving only an *upper limit* for a group’s potential.

In this paper, we propose a new logical formalization of the concept of group knowledge, based on our ‘Relevant Distinctions’ version of Interrogative Epistemology. We define the “epistemic potential of a group” as *the knowledge that the members of the group may come to jointly possess (and thus act upon as a group, in a coordinated manner) by (selectively) learning from each other all the available information that is relevant to their own issues*. We show that group knowledge typically lies in between common knowledge and distributed knowledge. We investigate the relationship between these concepts, and study their logic, which comes together with a complete axiomatization. We also look at the logical dynamics of information sharing in order to understand how the group’s potential knowledge is actualized (i.e. converted into common knowledge) by in-group communication. We use this setting to explain both the triumphs and the failures of collective knowledge, treating examples that range from “collective scientific knowledge” (Wray, 2007; Rolin, 2008; Fagan, 2012) to the so-called “curse of the committee”.

### 1.3 Examples

Our formal account in this paper is guided by the philosophical discussions on scientific knowledge, which in “normal” times is one of the most successful and most consensus-based forms of collective knowledge, but which sometimes may also lead to clashes between different conceptual frameworks and interrogative agenda. Our other motivation is provided by well-known common-day examples of failures of cooperative deliberation and information-sharing to maximize a group’s epistemic potential: the so-called “curse of the committee”.

**Example 1.** When Galileo announced his revolutionary astronomical discoveries, there was widespread skepticism. Many scholastics simply refused to look through his telescope. Today we find this odd, but at the time there was no good theory to support the usage of telescopes as instruments of discovery. Mirrors were notoriously unreliable, the instruments of magicians and illusionists. Nature, according to the Aristotelian philosophy, was accessible to direct observation by the unaided human eye. To presume that an eye can see better with the help of a mirror was regarded by many in the same way that we now regard Timothy Leary’s claim that we can penetrate deeper

into the essence of reality with the help of LSD. Questions such as “what new things can I see in the sky using mirrors, that nobody else ever saw before?”, or “why does the telescope show me mountains on the moon, which were never seen before?”, were not considered by many scholastics to be meaningful questions worthy of “scientific” investigation. Common wisdom said that with a cleverly-designed mirror you can make anybody see anything. Among other things, you might also see something real, but not necessarily so. And especially when they were showing things that went totally against common sense and tradition, there was no reason to believe in the reality of these visions. And so there was no reason to look for an explanation for them. As a result, Galileo’s facts were not accepted by many of his learnt contemporaries. A failure of collective scientific knowledge, if there ever was any.

Eventually, things did change. But we are not concerned here with the change: in this paper, we do not yet aim to formalize the paradigm shift that changed the interrogative agenda (though see the last section on how one may use dynamic epistemic logic and question-raising actions to model paradigm shifts). But we are simply concerned to model the situation at the time when the larger community did not share yet Galileo’s “facts” and his questions.

**Example 2.** Our second example refers to the theory of continental drift in the history of science. For centuries geographers noted the almost perfect fit between the coasts of Western Europe and North America, and between the ones of Africa and South America. But this piece of information did not answer any scientific question of the day: it simply did not fit with any scientific hypothesis or theory about the formation of the earth. In our terms, this fact, though noticed by scientists, was not really “known”; instead, it was disregarded as an irrelevant quirk of Nature, a mere coincidence or maybe a play of God. The question “why do the continents appear to fit together?” was not considered a scientifically relevant question, similarly to the way we would now consider a child’s question about “why does that cloud in the sky look like the face of an old lady?”. We assign no special meaning to the fact that some cloud resembles a human shape, and treat it as an irrelevant coincidence. The same applied for centuries to the apparent resemblance between the coasts on the two sides of the Atlantic. Occasionally, some lay people did speculate about the possibility that America and Europe were once closer, united into one continuous continent, but since there was absolutely no plausible explanation for how they may have drifted apart, no scientist took such speculation seriously.

In 1912 Alfred Wegener introduced the principle of continental drift as an “attempt to explain the origins of large Earth features, or the continents and ocean basins with a comprehensive principal” (Wegener, 2002). One of his pieces of evidence was the almost perfect fit of the coasts on the two sides of the Atlantic. Wegener stated that “wherever once continuous old land features are interrupted at the sea, we will assume continental separation and drift. The resulting picture of our Earth is new and paradoxical, but it does not reveal the physical causes.” He accumulated a large body of circumstantial evidence in favor of his hypothesis (e.g. similarity of plant and animal fossils on the two sides etc). In this way, the above-mentioned question came closer to becoming a “scientific” question. At this point, it *was* a valid question, at least for Wegener (and a few other scientists).

But not for everybody. Wegener almost totally lacked the support of the scientific community for a long time. The reason is that he wasn’t able to answer the question about the “cause of the drift”: there seemed no way to explain how continents could ‘float’. Continents are not icebergs floating on the sea, but huge solid masses of heavy rock. As a result, most scientists continued to consider continental drift to be a fanciful hypothesis. The fact that continents’ coasts fit together, far from being regarded as a ‘proof’ of continental drift, remained for a while still what it always was: a weird, meaningless coincidence.

Gradually though, as other evidence accumulated, geologists started to look again at that old quirky question “why do the continents appear to fit together?”. Slowly, this became a legitimate scientific question, its modern version being: “what caused the continental drift?”. As such, this was the basis of a research program that led to the contemporary theory of tectonic plates. And so in our times, that old coincidence is no longer meaningless. It became a solid scientific “fact”: a piece of knowledge, a partial answer to a good question, that fits well within our current framework and continues to fuel our research agenda.

Once again, in this paper we are not addressing the paradigm shift that converted the old irrelevant information into common knowledge. But we are simply concerned to model the situation facing young Wegener and his contemporaries. *His* relevant questions were not regarded as valid questions by many other scientists, and as a result his “facts” and his evidence were not generally accepted at the time (and for a long time after).

**Example 3.** Our third example deals with the deliberation in a hiring committee. Let us assume that John and Mary are two candidates for an open position and have just submitted their application. Both candidates claim to work at the intersection of Philosophy and Logic. Both candidates have asked their supervisors to write letters of reference for them and it now happens that John has better letters of reference (by far) than Mary. Our hiring

committee consists of two members, Alan and Betty. Alan is our philosophy expert, who doesn't understand much formal logic. He knows that John's philosophical writing is slightly better than Mary. But Alan can't judge the logic proofs of the candidates. Betty is a formal logician, she knows that Mary's logical work is really great, it fully backs her philosophical claims: indeed, Mary's work is much better than John's (John's proofs are full of mistakes). Both Alan and Betty can see that John has much better references than Mary, so this is *common knowledge* in the group  $G = \{\text{Alan}, \text{Betty}\}$ . We assume that these three facts together imply that *Mary is the best candidate overall*. Moreover, this fact is *distributed knowledge* in the group  $G = \{\text{Alan}, \text{Betty}\}$ . But is this really the decision that the committee will reach? As we'll see later on, the answer to this will depend on the issues that the committee members consider to be relevant for this hiring. The winner will be the candidate who performs better on (all and only) the issues which are *agreed to be relevant* by the committee members.

## 2 Background

**Questions and Answers.** Following (Groenendijk and Stokhof, 1985), we take a *question* to be just a partition<sup>1</sup> of the set  $S$  of all possible worlds into cells. This means that every cell corresponds to a possible answer to the question. And following (van Benthem and Minica, 2012), we *identify a question with the corresponding equivalence relation*  $\approx_Q$  (relating states lying in the same cell of the partition). We denote the set of questions over  $S$  as  $Quest_S := \{\approx_Q : \approx_Q \text{ is an equivalence relation on } S\}$ . A *binary question* is a special kind of question  $\approx^P$  given by a binary partition  $\{P, S \setminus P\}$ , where  $P$  is some proposition  $P \subseteq S$ . Formally, the equivalence relation  $\approx^P$  is given by

$$s \approx^P t \text{ iff either } s, t \in P \text{ or } s, t \notin P.$$

The cells of the partition  $Q$  coincide with the equivalence classes modulo  $\approx_Q$ , and represent the possible *full answers* to question  $Q$ . A *partial answer* to  $Q$  is any union of such  $\approx_Q$ -cells. A (partial or total) answer  $P$  to a question  $\approx_Q$  is “correct” at world  $s$  if it is true at  $s$  (i.e.  $s \in P$ ). Every family of questions  $Quest \subseteq Quest_S$  can be ‘compressed’ into one big ‘conjunctive’ question, whose cells (answers) are all the non-empty intersections of  $\approx_Q$ -cells for each question  $\approx_Q \in Quest$ : this is the least refined partition that refines every question in  $Quest$ . The corresponding equivalence relation is  $\approx_{Quest} = \bigcap \{\approx_Q : Q \in Quest\}$ .

**Agent's Issues: Conceptual Indistinguishability.** Following (van Benthem and Minica 2009), we assume a group  $G$  of “agents”, such that each agent is assumed to have a number of *fundamental questions* or “issues”, pertaining to all the world-properties or distinctions that she considers as relevant. As explained above, we can take the ‘conjunctive’ question  $\approx_a$  that collects all agent  $a$ 's fundamental questions (by taking the intersection of all the corresponding equivalence relations). van Benthem and Minica call  $\approx_a$  the agent  $a$ 's *issue relation*, or agent  $a$ 's *total question*; it essentially captures agent  $a$ 's *conceptual indistinguishability* relation, since it specifies the finest relevant world-distinctions that agent  $a$  makes. Together, the relevant distinctions form the basis of the agent's conceptual space: the way she represents the possible worlds. Two worlds  $s \approx_a s'$  are conceptually indistinguishable for agent  $a$  (since the answers to all  $a$ 's questions are the same in both worlds): one can say that  $s$  and  $s'$  will correspond to the *same world* in agent  $a$ 's own “subjective model”.

**Knowledge.** As common in epistemic logic, each agent  $a \in G$  is also endowed with an *epistemic indistinguishability relation*  $\rightarrow_a$ , assumed to be *reflexive*.  $s \rightarrow_a s'$  means that at world  $s$  agent  $a$  considers  $s'$  to be an epistemically-possible world (i.e. an un-excluded alternative to the actual world  $s$ ). Following (Hintikka, 1962), knowledge is then defined as truth in all the epistemically-possible worlds.

### 2.1 Epistemic Issue Models

**Definition 1 (Epistemic Issue Model)** (*van Benthem and Minica 2009*) Given a finite group (set)  $G$  of “agents”, and a set  $\Phi$  of atomic sentences (to be interpreted as “ontic facts”), an *epistemic issue model* over  $(G, \Phi)$  is a tuple  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$ , consisting of a finite set  $S$  of *possible worlds* (or “states”); a map  $\rightarrow$  associating to each agent  $a \in G$  some *reflexive* binary relation  $\rightarrow_a \subseteq S \times S$ , called *epistemic accessibility* relation (where read  $s \rightarrow_a t$  as saying that world  $t$  is an epistemically possible alternative to world  $s$ ); a map  $\approx$  associating to each agent  $a \in G$  some equivalence relation  $\approx_a \subseteq S \times S$ , capturing agent  $a$ 's *issue partition*; and a *valuation* which maps the atomic sentences  $p \in \Phi$  to sets of worlds  $\|p\| \subseteq S$ .

<sup>1</sup>Our restriction to work with a setting in which all answers to a question are mutually exclusive, can be lifted if one works with ‘covers’ instead of equivalence relations (Groenendijk, 2007), (Groenendijk et al., 2009). However, to keep matters simple, we choose a setting with equivalence relations.

A *proposition* is a set  $P \subseteq S$  of possible worlds.<sup>2</sup> We use the standard notations  $\neg P := S \setminus P$  for the *negation* (complement) of proposition  $P$ ,  $P \wedge Q := P \cap Q$  for the *conjunction* (intersection) of  $P$  and  $Q$ ,  $P \vee Q := P \cup Q$  for their *disjunction* (union),  $\top := S$  for the *tautologically true* proposition and  $\perp := \emptyset$  for the *inconsistent* proposition.

**Epistemic States.** For a given world  $s$ , we put  $s(a) := \{s' \in S : s \rightarrow_a s'\}$  for agent  $a$ 's *epistemic state at  $s$* : this is the set of all worlds that are epistemically possible for  $a$  at  $s$ .

**Current Issues: the agent's open questions.** The restriction  $\approx_a |s(a) := \approx_a \cap (s(a) \times s(a))$  of the issue relation  $\approx_a$  to agent  $a$ 's current epistemic state  $s(a)$  represents *agent  $a$ 's current (open) issue(s) at world  $s$* . This relation encodes a partition of the epistemic state  $s(a)$  into cells, corresponding to the answers to agent  $a$ 's open questions at  $s$ .

Since binary relations  $R \subseteq S \times S$  play an important role in epistemic logic, we also use the following notations for the usual operations on relations: *converse* of a relation  $R^{-1} = \{(s, t) \in S \times S : (t, s) \in R\}$ ; *union* of relations  $R_1 \cup R_2$ ; *intersection* of relations  $R_1 \cap R_2$ ; *relational composition*  $R_1 R_2 = \{(s, t) \in S \times S : \exists w \in S (s, w) \in R_1 \wedge (w, t) \in R_2\}$ ; the  $n^{\text{th}}$  *iteration* of a relation  $R^n$  (defined recursively by putting  $R^0 = id := \{(s, s) : s \in S\}$  to be the *identity* relation and  $R^{n+1} = R^n R$ ); the *transitive closure* of a relation  $R^+$ , defined as the least transitive relation  $R^+ \supseteq R$ ; and the *reflexive-transitive closure*  $R^*$ , defined as the least reflexive and transitive relation  $R^* \supseteq R$ . It is easy to see that we have:  $R^+ = \bigcup_{n \geq 1} R^n := R \cup R^2 \cup R^3 \cup \dots$ , and  $R^* = id \cup R = \bigcup_{n \geq 0} R^n$ . One should also note that  $R^* = R^+$  whenever  $R$  is itself reflexive.

**Definition 2 (Kripke Modalities)** *For any binary relation  $R \subseteq S \times S$  on the set  $S$  of all possible worlds, the corresponding Kripke modality  $[R]$  can be introduced, as a function mapping any proposition  $P \subseteq S$  to some proposition  $[R]P \subseteq S$ , given by:*

$$s \in [R]P \text{ iff } \forall t \in S (sRt \Rightarrow t \in P).$$

**Definition 3 (Knowledge)** (*Hintikka, 1962*)

*Given an epistemic issue model, the knowledge operator is defined as the Kripke modality for the epistemic indistinguishability relation:*

$$K_a P := [\rightarrow_a]P, \text{ for all propositions } P \subseteq S.$$

*Equivalently:  $K_a P = \{s \in S : s(a) \subseteq P\}$ .*

The reflexivity of  $\rightarrow_a$  means that our models validate the modal axiom (T), expressivity that *knowledge is factive* ( $K_a P \Rightarrow P$ ). But note that in general we do *not* require knowledge to be introspective, since the assumption of full introspection has long been attacked by many philosophers. However, we are also interested in those special situations in which knowledge *is* introspective, and so we will also consider two special classes of models:

An epistemic issue model is *positively introspective* if all the epistemic accessibility relations are transitive. A model is *fully introspective* if all the epistemic accessibility relations are transitive and symmetric. Note that the class of positively introspective models validates the modal axiom (4), encoding Positive Introspection ( $K_a P \Rightarrow K_a K_a P$ ). Similarly, the class of fully introspective models validates both the axiom (4) and the axiom (5), encoding Negative Introspection ( $\neg K_a P \Rightarrow K_a \neg K_a P$ ).

**Definition 4 (Common Knowledge)** (*Lewis, 1969*)

*The common knowledge operator  $Ck$  (for our group  $G$  of agents) is the Kripke modality for the transitive closure of the union of all epistemic relations:*

$$CkP := [\rightarrow_{Ck}]P, \text{ where } \rightarrow_{Ck} = \left( \bigcup_{a \in G} \rightarrow_a \right)^+.$$

*(where recall that  $R^+$  is the transitive closure of  $R$ ). Equivalently,  $Ck$  can be defined as the infinite conjunction of iterated knowledge (about others' knowledge...):*

$$Ck_G P = \bigwedge_{a \in G} K_a P \wedge \bigwedge_{a, b \in G} K_a K_b P \wedge \dots$$

<sup>2</sup>Also known as a "UCLA proposition" in logic, and as an "event" in probabilistic terminology.

Intuitively, the concept of common knowledge captures the knowledge that all the agents of the group *can act upon in a coordinated manner, without the need to communicate*<sup>3</sup>: e.g. if the rules of the road, as well as the fact that everybody respects them, are common knowledge, then everybody can safely drive; similarly, if it is common knowledge among the units of an army that they all need to attack the enemy at dawn, then they can rely on each other to attack simultaneously (at dawn), without any need for further communication.

**Definition 5 (Distributed Knowledge)** (Fagin et al., 2003)

The distributed Knowledge operator is the Kripke modality for the intersection of all knowledge relations:

$$DkP := [\rightarrow_{Dk}]P \quad \text{where } \rightarrow_{Dk} := \bigcap_{a \in G} \rightarrow_a .$$

Distributed knowledge is meant to capture the sum of all the information carried collectively by the group.

Next, we follow (van Benthem and Minica, 2012) by introducing an *issue modality*  $Q_a$ :

**Definition 6 (Intended Knowledge: the Issue Modality)** (van Benthem and Minica, 2012)

We say that proposition  $P$  is a true fact about  $a$ 's issues (or an  $a$ -relevant truth), and write  $Q_a P$  if it is entailed by the true answer(s) to (all)  $a$ 's question(s). Formally, this corresponds to taking the Kripke modality for the issue relation:

$$Q_a P = [\approx_a]P .$$

The issue modality captures the information carried by the answers to (all)  $a$ 's question(s). As we'll see, in our framework  $Q_a P$  expresses something akin to *intended knowledge*: if  $a$  would know the true answers to all her questions then she would know  $P$ . In the original framework of van Benthem and Minica, this role was played by a different modality, namely by the *resolution modality*  $R_a$ , given by  $R_a P := [\approx_a \cap \rightarrow_a]P$ . However, as we'll see in the next section, these two modalities will coincide in our more restricted framework (i.e. we will have  $R_a = Q_a$ ), so that  $Q_a$  will indeed capture agent  $a$ 's intended knowledge.

**Relevant Partial Information.** Besides full answers, agents may have some partial information relevant to another agents' question  $\approx_a$ . Such partial information of an agent  $b$  is given by the set of answers to  $\approx_a$  that she considers possible, *viz.* by the worlds  $t$  that are  $\approx_a$ -equivalent to any worlds that are considered by  $b$  to be epistemically possible alternatives of the actual world  $s$ . So the composite relation  $\rightarrow_b \approx_a$  encodes all knowledge of agent  $b$  about (the answers to)  $a$ 's question(s): *all the information in  $b$ 's possession that is relevant for  $a$ 's issues.*

### 3 Interrogative Epistemology and Selective Learning

**Knowledge and Issues.** Intuitively, *conceptual indistinguishability implies epistemic indistinguishability*: indeed, we will require that  $\approx_a \subseteq \rightarrow_a$ . In other words:  $s \approx_a s'$  will imply  $s \rightarrow_a s'$ . This is a natural condition, easily justifiable by counterposition: suppose that  $s \not\rightarrow_a s'$ ; then, at (actual) world  $s$ , the agent knows that  $s'$  is *not* the actual world, so there should exist some  $a$ -relevant property that distinguishes  $s'$  from the actual world ( $s$ ); i.e.  $s$  should differ from  $s'$  with respect to at least some of the answers to  $a$ 's questions (hence  $s \not\approx_a s'$ ). But this poses a fundamental restriction on an agent's knowledge: we will always have  $K_a \varphi \Rightarrow Q_a \varphi$ , i.e. *an agent can only know answers to her (implicit or explicit) question(s)*. We think of this as a natural limitation: we can truly 'know' (in the active sense, of being capable to act upon) only facts that are meaningful to us and relevant to our issues.

**To Know is to Know the Answer to a Question.** In fact, we will require a slightly *stronger condition*, namely that  $(\rightarrow_a \approx_a) \subseteq \rightarrow_a$ , in other words:  $s \rightarrow_a \approx_a s'$  will imply  $s \rightarrow_a s'$ . This condition is even closer in spirit to the Schaffer quote above: *to know  $P$  is to know that  $P$  is an answer to (one of) your question(s)*. This is indeed stronger than the above condition ( $\approx_a \subseteq \rightarrow_a$ ) as it implies it. In fact, the two conditions are equivalent for positively introspective epistemic agents. But the stronger condition can also be justified independently, again by counterposition: suppose that  $s \not\rightarrow_a s'$ ; then, at (actual) world  $s$ , the agent knows that  $s'$  is *not* the actual world, so she would be able to *point out* some  $a$ -relevant property that distinguishes  $s'$  from the actual world ( $s$ ). In other words, she should *know* that the true answer (at  $s$ ) to some of her questions is not the same as the answer at  $s'$  (hence we do *not* have  $s \rightarrow_a \approx_a s'$ ).

<sup>3</sup>Indeed, the notion of common knowledge traces back to the work of David Lewis (Lewis, 1969), who investigated it in the context of analyzing social conventions and their role in coordination.

### 3.1 Interrogative Epistemic Models

Putting this all together, we arrive at the following notion:

**Definition 7 (Interrogative Epistemic Model)** Given a set  $G$  of agents and a set  $\Phi$  of atomic sentences, an *interrogative epistemic model over*  $(G, \Phi)$  is an epistemic issue model  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$  satisfying the condition:

$$(*) \rightarrow_a \approx_a \subseteq \rightarrow_a .$$

**Observation** On positively introspective models, condition  $(*)$  can be simplified to:

$$\approx_a \subseteq \rightarrow_a .$$

Our additional requirement  $(*)$  is meant to capture our specific interpretation of the issue relation<sup>4</sup> (as encoding all the fundamental issues or conceptual distinctions that are relevant and meaningful for agent  $a$  and that define her research agenda). Using the issue modality, we can see that our semantic condition  $(*)$  corresponds to assuming the validity of the following statement:

$$K_a P \Rightarrow K_a Q_a P$$

This statement literally captures Schaffer’s quote, though only when interpreted according to our Relevant Distinctions conception: if an agent knows  $P$ , she knows that  $P$  is (entailed by) the answer to one of her questions. In other words: all one’s knowledge is knowledge *about one’s issues*.

### 3.2 The Hiring Committee: Example 3 reviewed

To analyze Example 3 above, let us make things precise by giving “scores” (on a scale from 1 to 10) to each candidate for each of the three areas (References, Philosophy, Logic). We represent the epistemic state and issues of each of the agents in the following tables, as well as the group’s common and distributed knowledge.

Alan’s Knowledge	References	Philosophy	Logic
John	10	10	?
Mary	6	8	?

Betty’s Knowledge	References	Philosophy	Logic
John	10	?	0
Mary	6	?	10

We assume that the candidate with a higher overall score (given by the sum of his/her three scores) is “better”. So, if somehow we could pull together the information of Alan and Betty on *each of the three questions*, then Mary should get the job!

Distributed Knowledge	References	Philosophy	Logic
John	10	10	0
Mary	6	8	10

But this is not always what happens in reality! What if Alan doesn’t care about Logic, and Betty doesn’t care about Philosophy? Suppose that Alan doesn’t consider the question about the Logic score to be relevant for the hiring issue; and dually, Betty doesn’t regard the Philosophy score as relevant. In this case, they won’t be able to learn anything from each other’s expertise: they’ll simply disregard the additional information possessed by the other! In order to reach any agreement at all, they will have to fall back on the only thing they agree: their common knowledge that John has better references. In this case, John might get the job after all!

Common Knowledge	References	Philosophy	Logic
John	10	?	?
Mary	6	?	?

<sup>4</sup> Note that an agent’s issue also includes both *answered* questions and her open questions.



To encode this example as an interrogative epistemic model, let a *scoring function* be a map  $\sigma : \{John.Mary\} \rightarrow \{0, 1, \dots, 10\}$  that assigns a score from 0 to 10 to each of the candidates. Let  $\Sigma$  be the set of all scoring functions. A *possible world/state* in our model is a map  $s : \{References, Logic, Philosophy\} \rightarrow \Sigma$  that associates a scoring function to each of the three topics; e.g.  $s(References)(Mary)$  is the answer to the question “what is Mary’s performance as far as References are concerned?”, etc. So, essentially, any such map  $s$  is a complete description of a table of numbers like the ones above. The set of worlds of our models is the set  $S$  of all these maps: so the model consists of all possible such tables. We have two agents, Alan (agent  $a$ ) and Betty (agent  $b$ ). As epistemic relation for Alan, we put  $s \rightarrow_a t$  iff:  $s(References) = t(References)$  and  $s(Philosophy) = t(Philosophy)$ . This encodes the fact that Alan knows the scores for references and for philosophy. His issue relation is the same ( $\approx_a = \rightarrow_a$ ), since the Logic score is not a relevant question for Alan. Dually, we put for Betty  $s \rightarrow_b t$  iff:  $s(References) = t(References)$  and  $s(Logic) = t(Logic)$ , which expresses the fact that Betty knows the scores for references and for logic; we also put  $\approx_b = \rightarrow_b$ , which means that she does not consider Philosophy scores to be relevant. None of them has any open question: they already know the answers to all their issues. One can easily see that it is *distributed knowledge* among Alan and Betty that Mary’s total score (the sum of all her scores) is higher than John’s, i.e. that Mary is overall better than Alan. But, as we’ll see, this fact is *not* potential group knowledge.

### 3.3 Dynamics: Selective Learning from Testimony

Moving on to the dynamics of information flow, let us consider *public testimony* (shared with the whole group  $G$ ), understood in the spirit of Dynamic Epistemic Logic (Baltag et al., 1998; van Benthem, 2011; van Ditmarsch et al., 2007). Suppose a proposition  $P \subseteq S$  is publicly announced in the above-mentioned sense. What will an agent  $a$  learn from this announcement? The above considerations lead us to the following:

**Principle of Selective Learning.** *When confronted with information, agents come to know only the information that is relevant for their issues.*

The usual semantics of public announcements does not fit with our framework, since it assumes that agents can automatically absorb any new information, irrespective of how it fits with their issues. So we need to modify this semantics. But we also aim to give our agents the benefit of the doubt by endowing them with the maximal learning capacity that is compatible with the Selective Learning Principle. So we assume that they don’t just disregard a testimony when it’s not directly relevant to them, but are able to extract from it the *maximal (i.e. strongest) relevant information that is entailed by the announcement* (and disregard the rest). For this, we need the following notion:

**Definition 8 (Strongest  $a$ -relevant proposition)** *For any proposition  $P \subseteq S$  and agent  $a \in G$ , we denote by  $P_a$  the strongest  $a$ -relevant proposition entailed by  $P$ , and we formally define it as:*

$$\mathbf{P}_a := \{s \in \mathbf{S} : s \approx_a s' \text{ for some } s' \in P\}.$$

We can now state our modified semantics for public announcements:

**Definition 9 (Selective Public Announcements)** *Given an interrogative epistemic model  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$  and some proposition  $P \subseteq S$ , a (selective) public announcement  $!P$  is an action that changes the model  $\mathbf{S}$  into an updated model  $\mathbf{S}^P = (S^P, \rightarrow^P, \approx^P, \|\bullet\|^P)$  in which each agent  $a$  has updated her knowledge only with  $P_a$ . Formally:*

1. *the set of possible worlds stays the same:  $S^P = S$ ;*
2. *the new epistemic relations  $\rightarrow_a^P := \rightarrow_a \cap \approx^{P_a}$  are formed by taking the intersection of the initial epistemic relations  $\rightarrow_a$  with the equivalence relation  $\approx^{P_a}$  for the binary question  $\{P_a, \neg P_a\}$ ;*
3. *the issue relations stay the same:  $s \approx_a^P t$  iff  $s \approx_a t$ .*
4. *the valuation stays the same:  $\|p\|^P = \|p\|$ , for  $p \in \Phi$ .*

**Observation** If  $\mathbf{S}$  is an interrogative epistemic model then so is  $\mathbf{S}^P$ .

**Definition 10 (Information Pooling: “Sharing All We Know”)** Given an interrogative epistemic model  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$ , the action  $!$  of information pooling is the one by which all the agents in the group share all their knowledge. This means that they simultaneously and publicly share all they know: formally, at each world  $s$  the conjunction of all their knowledge  $\bigcap_{a \in G} s(a)$  is publicly announced; so locally this is equivalent to a (selective) public announcement  $!(\bigcap_{a \in G} s(a))$ . As a global model-transformation though, this cannot be simulated by any specific announcement  $!P$ , so we need to define independently the resulting model  $\mathbf{S}^! = (S^!, \rightarrow^!, \approx^!, \|\bullet\|^!)$ , by putting:

1. the set of possible worlds stays the same:  $S^! = S$ ;
2.  $\rightarrow_a^! := \rightarrow_{Dk \approx_a}$  is the composition of the distributed knowledge relation and  $a$ 's issue relation. (Formally, if we use the local version of the definition of public announcement  $!(\bigcap_{a \in G} s(a))$ , we get  $\rightarrow_a^! := \rightarrow_a \cap \rightarrow_{Dk \approx_a}$ , but the intersection with  $\rightarrow_a$  is redundant here.)
3. the issue relations stay the same:  $s \approx_a^! t$  iff  $s \approx_a t$ .
4. the valuation stays the same:  $\|p\|^! = \|p\|$ , for  $p \in \Phi$ .

One can easily check that agent  $a$ 's new epistemic state  $s(a)^!$  in the model  $\mathbf{S}^!$  is the intersection of her old epistemic state with the  $a$ -relevant information carried by the group's distributed knowledge:  $s(a)^! = s(a) \cap (\bigcap_{a \in G} s(a))$ . This shows that at each world  $s$  the information pooling action  $!$  behaves indeed locally like a selective public announcement  $!(\bigcap_{a \in G} s(a))$  with all the distributed knowledge of the group  $G$ .

**Observation** If  $\mathbf{S}$  is an interrogative epistemic model then so is  $\mathbf{S}^!$ .

However, the above observations do not extend to Introspection properties.

**Proposition 1** Transitivity and symmetry are *not* preserved by (selective) public announcements  $!P$  or information pooling  $!$ . The update of a positive introspective model with the action  $!$  may fail to be positively introspective, and the same holds for negative introspection.

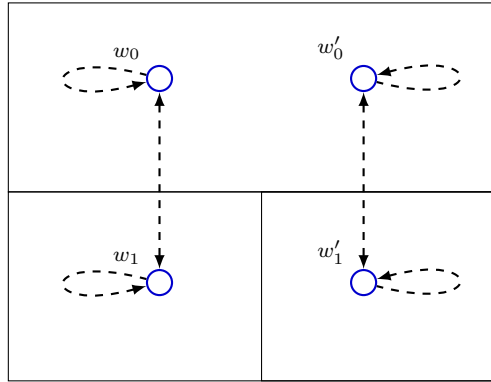


Fig. 2. An Epistemic Issue Model: the dashed lines represent agent  $b$ 's epistemic relation  $\rightarrow_b$ . We do not draw agent  $a$ 's epistemic relation, but we assume that agent  $a$  doesn't have any non-trivial knowledge: so  $\rightarrow_a$  is the total relation, connecting every two worlds. The three cells drawn as black squares represent agent  $a$ 's issue partition  $\approx_a$ . We do not draw here agent  $b$ 's issue partition (since it is not important), but one can just assume e.g. that  $\approx_b = \rightarrow_b$  (i.e.  $b$  knows the answers to all his questions).

We illustrate this proposition via Fig.2 and Fig.3. Note that both agents are fully introspective in the initial model drawn in Fig. 2, but this is no longer the case after information-sharing. Fig.3 illustrates the updated model after information sharing. Agent  $a$ 's new epistemic relation  $\rightarrow_a^!$  after information sharing is obtained by composing  $\rightarrow_{Dk \approx_a}$ , which in this case is the same as  $\rightarrow_b \approx_a$ . As depicted, this relation is *not* positively introspective: e.g.  $w_0 \rightarrow_b \approx_a w'_0 \rightarrow_b \approx_a w'_1$  holds but we do *not* have  $w_0 \rightarrow_b \approx_a w'_1$ .

So *Introspection can be lost by learning new facts!* We think that this conclusion is more natural than it may look at first sight. Some new information can be learnt in a non-introspective manner during selective learning; but this doesn't mean that the agent suddenly loses all introspection! She will still be introspective with respect to

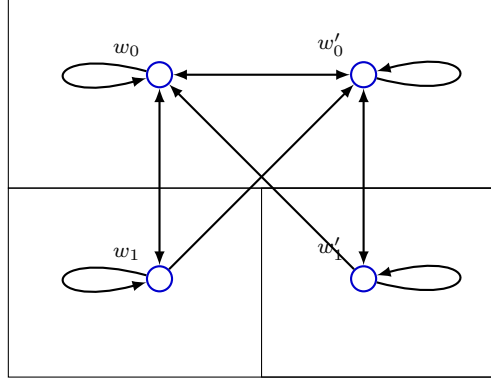


Fig. 3. Failure of positive introspection. The continuous arrows represent agent  $a$ 's relation after information-sharing:  $\rightarrow_a^! = \rightarrow_{Dk} \approx_a = \rightarrow_b \approx_a$ .

some of the information in her possession, possibly including some of the newly acquired information. It is a matter of philosophical choice or taste whether we use the term “knowledge” to cover *all* the information learnt by an agent, or on the contrary we restrict the use of this term to cover only the information possessed in an introspective manner. Formally, this corresponds to restricting the models to positively introspective, or to fully introspective, ones. If one chooses one of these options, then one can then do the same in the definition of update: focus only on the introspective knowledge acquired through this update. This leads us to the following alternative definitions:

**Definition 11 (Introspective Public Announcements)** *Given a (positively, or fully) introspective interrogative-epistemic model  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$  and some proposition  $P \subseteq S$ , the introspective public announcement  $!P^+$  is an action that changes the model  $\mathbf{S}$  into an updated model  $\mathbf{S}^{P+} = (S^P, \rightarrow^{P+}, \approx^P, \|\bullet\|^P)$ , where  $S^P, \approx^P, \|\bullet\|^P$  are the same as in the model  $\mathbf{S}^P$  above (for non-introspective public announcement  $!P$ ), and the epistemic accessibility relations  $\rightarrow_a^{P+}$  are given by the transitive closure  $(\rightarrow_a^P)^+$  of the relations  $\rightarrow_a^P$  from the model  $\mathbf{S}^P$ :*

$$\rightarrow_a^{P+} := (\rightarrow_a \cap \approx^{P_a})^+ = \rightarrow_a^P \cup \rightarrow_a^P \rightarrow_a^P \cup \rightarrow_a^P \rightarrow_a^P \rightarrow_a^P \cup \dots$$

**Observation** If an interrogative epistemic model  $\mathbf{S}$  is (positively, or fully) introspective, then so is  $\mathbf{S}^{P+}$ .

**Definition 12 (Introspective Information Pooling)** *Given a (positively, or fully) introspective interrogative epistemic model  $\mathbf{S} = (S, \rightarrow, \approx, \|\bullet\|)$ , the action  $!$  of introspective information pooling  $!^+$  is an action that changes the model  $\mathbf{S}$  into an updated model  $\mathbf{S}^{!+} = (S^!, \rightarrow^{!+}, \approx^!, \|\bullet\|^!)$ , where  $S^!, \approx^!, \|\bullet\|^!$  are defined the same way as in the model  $\mathbf{S}^!$  above (for non-introspective information sharing  $!$ ), and the epistemic accessibility relations  $\rightarrow_a^{!+}$  are given by the transitive closure  $(\rightarrow_a^!)^+$  of the relations  $\rightarrow_a^!$  from the model  $\mathbf{S}^!$ :*

$$\rightarrow_a^{!+} := (\rightarrow_{Dk} \approx^{P_a})^+ = \rightarrow_a^! \cup \rightarrow_a^! \rightarrow_a^! \cup \rightarrow_a^! \rightarrow_a^! \rightarrow_a^! \cup \dots$$

**Observation** If an interrogative epistemic model  $\mathbf{S}$  is (positively, or fully) introspective, then so is  $\mathbf{S}^{!+}$ .

Note that if we compare the updated models  $\mathbf{S}^!$  and  $\mathbf{S}^{!+}$  for the same initial model  $\mathbf{S}$ , then agents will typically *know less* after introspective information pooling  $!^+$  than after simple (non-introspective) pooling  $!$ . So, interestingly enough, *introspection can act as a barrier to learning!* Once again, we think this is more natural than it looks at first sight: by focusing on introspective knowledge, we “screen out” the information that was acquired in a non-introspective manner. Of course, on the other hand, the agent *knows more about* the things that she knows in this introspective manner: in addition to knowing them, she knows that she knows them, etc. But typically there are *less* things that are known in this way. So the conclusion is that “introspective agents know more about less”.

One very important observation is that (unlike knowledge or distributed knowledge), *the common knowledge after an update does not depend on whether the update is introspective or not:*

**Proposition 2** *The same propositions are common knowledge in  $\mathbf{S}^P$  as in  $\mathbf{S}^{P+}$ . Similarly, the same propositions are common knowledge in  $\mathbf{S}^!$  as in  $\mathbf{S}^{!+}$ .*

The reason is that *common knowledge is already introspective by definition*, so the choice of update does not make a difference. Formally:

$$\rightarrow_{Ck}^! = \left( \bigcup_{a \in G} \rightarrow_{Dk \approx a} \right)^+ = \left( \bigcup_{a \in G} (\rightarrow_{Dk \approx a})^+ \right)^+ = \rightarrow_{Ck}^{!+},$$

and the same for  $\rightarrow_{Ck}^P$  versus  $\rightarrow_{Ck}^{P+}$ .

## 4 The Epistemic Potential of a Group

In the philosophical literature, group knowledge is typically argued to be non-summativ—*not* fully reducible to the knowledge of the group members (Gilbert, 1987; Rolin, 2008; Wray, 2007).<sup>5</sup> In essence this implies a form of epistemic holism, of the type that one encounters in physical studies of correlated or entangled systems. In this context, the non-summativ ingredient refers to a complex interplay or the existence of a non-trivial correlation between the epistemic states of each of the individual members. Note that an analysis of group knowledge that is summative effectively eliminates group knowledge, and can thus do no additional explanatory work. Indeed, such a notion would merely be shorthand for ‘every body knows that’ or ‘someone in the group knows that’, but there would not be such a thing as the knowledge that is held by a group (as a group) over and above the knowledge of the individual group members. Specifically, philosophers who use the notion of group knowledge in order to explain scientific knowledge typically insist that group knowledge should at least be non-summativ. As such, the standard formal group-epistemic notion of distributed knowledge does not seem to fully meet the same requirements as typically envisioned in the philosophical literature; to compute distributed knowledge, it is enough to be given the knowledge of each of the agents in the group.<sup>6</sup>

Additionally, one can argue that *scientific group knowledge should imply potential individual knowledge*: although a scientific community as a group has knowledge that cannot be reduced to the knowledge of any of its members, these members must nonetheless be able to come to know it (through the group’s testimony) (Corlett, 1996; de Ridder, 2014; Rolin, 2008; Wray, 2007). Hence the correlation between the individual epistemic states has to be in principle accessible to the individuals.

To summarize, we propose a non-summativ notion of group knowledge that each individual member can potentially acquire. Moreover, one can make the non-summativ mechanism of epistemic correlations explicit, indicating how the group members’ potential individual knowledge can be combined to yield our concept of ‘potential group knowledge’. As we show formally later on, such a mechanism relies in essence on the role of the agents’ epistemic issues (questions). So, to compute a group’s knowledge, we have to be given both the knowledge and the issues of each of the agents.<sup>7</sup>

### 4.1 Potential Knowledge of an Individual or Group

Based on the above considerations, we first introduce the following informal definitions:

**Definition 13 (Potential Knowledge of an Agent in a Group)** We say that an agent  $a$  (belonging to a group  $G$ ) *potentially knows that  $P$  within the group  $G$* , and write  $K_a^G$ , if:

- (1)  $P$  is entailed by a true answer to (some of)  $a$ ’s question(s);
- (2) the fact that (1) holds is distributed knowledge in group  $G$ : i.e. this fact is entailed by some set of premisses, each of which is (individually) known by at least one member of the group.

In other words, an agent  $a$ ’s potential knowledge within a group  $K_a^G$ , is given by putting together  $a$ ’s (individual) knowledge with the group’s answers to all her questions. Formally,  $K_a^G$  is the Kripke modality for the group  $G$ ’s distributed-knowledge-about- $a$ ’s-question(s) relation:

$$K_a^G P := DkQ_a P = [\rightarrow_{G \approx a}]P.$$

<sup>5</sup>The term ‘summativism’ was introduced by Anthony Quinton in (Quinton, 1975) in an epistemic context.

<sup>6</sup>Though one should note that a fact can be distributed knowledge without being known to any member of the group: distributed knowledge encompasses everything that is logically entailed by the combination of all the individual pieces of knowledge.

<sup>7</sup>We agree that in fact this is still a rough idealization. Other factors should be taken into account when computing group knowledge, e.g. the structure of the group, the various relationships, the degree of mutual trust, each agent’s level’s of expertise etc. But for simplicity, in this paper we focus only on the role of questions.

This definition produces a not-necessarily-introspective notion of potential knowledge (even when the individual’s actual knowledge is introspective). An *alternative definition* (which pre-encodes the effect of Introspective Information Sharing) can be obtained by taking the “introspective” version of the above notion, i.e. putting

$$K_a^G P := [(\rightarrow_{Dk \approx_a})^+] P,$$

where recall that  $R^+$  is the transitive closure of  $R$ . This alternative definition gives a *different notion of potential knowledge*, which extracts only the “introspective part” of the concept of potential knowledge from the first definition: i.e. the information that is “introspectively known” according to the first notion. The fact that we use the same notation for two distinct notions may sound confusing. But this is motivated by the fact that they will provide two different semantics for *the same* syntactic construct ( $K_a^G$ ). The first is to be used when a non-introspective interpretation of knowledge is desired, while the second applies when we are talking about higher-level, “reflexive” knowledge (i.e. the introspective kind). Moreover, this systematic ambiguity is also justified by the fact that our focus in this paper is on the concept of (potential) *group* knowledge. And, as we’ll see, the two definitions above will make no difference for the group’s potential: *both notions of potential individual knowledge lead to the same concept of group knowledge*.

**Definition 14 (Potential Group Knowledge)** A group’s potential group knowledge  $Gk$  is given by the Kripke modality for the transitive closure of the union of all the potential individual knowledge relations:

$$GkP := [\rightarrow_{Gk}]P, \text{ where } \rightarrow_{Gk} = \left( \bigcup_{a \in G} \rightarrow_{Dk \approx_a} \right)^+,$$

where recall that  $R^+$  is the transitive closure of relation  $R$ .

**Observation.** *Potential group knowledge is common potential knowledge.* More precisely, *no matter which of the two above alternative definitions we adopt for potential individual knowledge  $K_a^G$ , we have that a group  $G$  potentially knows that  $P$  iff everybody in  $G$  potentially knows  $P$  within  $G$ , and everybody in  $G$  potentially knows within  $G$  that everybody potentially knows  $P$  within  $G$ , etc.:*

$$GkP = \bigcap_{a \in G} K_a^G P \cap \bigcap_{a, b \in G} K_a^G K_b^G P \cap \dots$$

So  $Gk$  is the “common” version of potential knowledge. But as we’ll see, it is also *potential common knowledge*: it is essentially given by what will become common knowledge after each agent in the group updates her knowledge with the group’s information concerning all her questions. This can already be seen from the fact that the accessibility relation for potential group knowledge in a model  $\mathbf{S}$  coincides with the accessibility relation for common knowledge in the updated model after information-sharing (in both versions  $\mathbf{S}^!$  and  $\mathbf{S}^{!+}$ ):

$$\rightarrow_{Gk} = \left( \bigcup_{a \in G} \rightarrow_{Dk \approx_a} \right)^+ = \left( \bigcup_{a \in G} \rightarrow_a^! \right)^+ = \rightarrow_{Ck}^! = \left( \bigcup_{a \in G} \rightarrow_a^{!+} \right)^+ = \rightarrow_{Ck}^{!+}.$$

So potential group knowledge is potential common knowledge: this is the motivation behind our definition. *Potential group knowledge “pre-encodes” the information that all the agents of the group can come to possess in common (and thus act upon in a coordinated manner) after they pool together all their knowledge.* As a consequence, group knowledge shares some of the properties of common knowledge:

**Proposition 3** *Potential group knowledge  $Gk$ , as well as common knowledge  $Ck$ , are always “positively introspective”:*

$$GkP = GkGkP, \quad CkP = CkCkP.$$

*Moreover, if all agents in  $G$  are fully introspective, then potentially group knowledge and common knowledge are also fully introspective (i.e.  $\neg GkP = \neg CkP = Ck\neg CkP$ ).*

Similar introspection properties hold for distributed knowledge *only* if we assume them for individual knowledge:

**Proposition 4** *If all agents in  $G$  are positively introspective, then distributed knowledge is as well (i.e.  $DkP = DkDkP$ ). Similarly, if all agents in  $G$  are fully introspective, then distributed knowledge is as well (i.e. we also have  $\neg DkP = Dk\neg DkP$ ).*

In fact, distributed knowledge and common knowledge can be seen as two ‘extreme’ versions of potential group knowledge, giving *the upper and lower limits for a group’s epistemic potential*:

**Proposition 5**  $CkP \subseteq GkP \subseteq DkP$ .

Moreover, these extremes can be reached in certain conditions:

**Proposition 6** If every agent in  $G$  is interested in the information possessed by all the other agents, then the potential group knowledge is *maximal* (and matches the group’s *distributed knowledge*):

$$\text{If } \approx_a \subseteq \rightarrow_{Dk} \text{ for every } a \in G \text{ then } GkP = DkP.$$

**Proposition 7** If each agent’s questions are ‘orthogonal’ to the other agents’ knowledge (i.e. no agent has any open questions that can be answered by the group), then the potential group knowledge is *minimal* (and reduces to the group’s initial *common knowledge*):

$$\text{If } \rightarrow_a \subseteq (\approx_a \cap \rightarrow_{Dk}) \text{ for every } a \in G \text{ then } GkP = CkP.$$

**Application: from the “Wisdom of the Crowds” to the “Curse of the Committee”.** What we can conclude from these results is that a group formed of agents with very dissimilar agendas will have a very poor joint epistemic power, whereas a group with very cohesive agendas will have a very strong epistemic potential. This is a consequence of our Selective Learning Principle, which we think to correctly reflect some features of real-life groups. Even in a scientific deliberation, researchers from different areas of science or philosophy will typically process successfully only the information that is relevant to the fundamental questions of their own area. Skeptical questions such as “How do I know that the universe exists, or that it’s all an illusion?” or “Do we have a logically infallible proof of this or that fact?” are typically regarded as irrelevant in Physics or Biology, though the first is essential in Epistemology and the second is essential in Mathematics. The situation gets even worse when scientists who adopted different paradigms meet each other (or read each other’s books): an Aristotelian physicist has very few things to learn from a Newtonian one, and vice-versa. Even if they agree on the actual observations, they cannot agree on what are the fundamental questions, and as a result some of one’s facts may look meaningless from the other’s perspective. The fact that a group’s potential can be severely reduced due to the agents’ very dissimilar agendas may also help us explain phenomena such as “the curse of the committee”, in which long episodes of deliberation and communication between apparently rational agents end up agreeing only on what was already common knowledge to start with. The group falls back onto the “lowest common denominator”, since the agents are unable to learn each other’s information (due to the extreme discrepancy between their agendas). The usual explanation is that people are ‘irrational’, but our analysis suggests that even ideally rational agents may end up in the same situation, if their agendas are ‘orthogonal’ to the others’ knowledge. The fact that a group may have a very low epistemic potential is not a matter of individual irrationality. Agents can be fully rational while still disagreeing on what are the relevant issues, or what are the fundamental questions that need to be asked.

**Illustrating the Curse (Example 3 continued).** We can for instance check now that in the model in Example 3, potential group knowledge equals common knowledge. This is because Alan’s information is irrelevant for Betty’s issues, and vice-versa. So, although it is distributed knowledge among them that Mary has overall a better score, this fact cannot be converted into common knowledge by any amount of information sharing.. At the end of their deliberation, the only thing that Alan and Betty can agree is what was common knowledge to start with: namely, that John had better references. The group fails to process all the available information. John gets the job: a “curse of the committee” indeed!

## 4.2 The Logic of Potential Group Knowledge

**Syntax and Semantics** In this section, we present an axiomatic system for the *logic of (potential) group knowledge (LGK)*, as well as variants for positively introspective knowledge  $LGK^+$  and fully introspective knowledge  $LGK^\pm$ , and some other extensions. We first fix a pair  $(G, \Phi)$  consisting of a finite set of agents with  $|G| \geq 2$  (i.e. we have *at least two distinct agents*) and a set  $\Phi$  of atomic propositions. The language of  $LGK$  is given in BNF form by the following recursive definition:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_a\varphi \mid Q_a\varphi \mid Dk\varphi \mid Ck\varphi \mid Gk\varphi$$

Intuitively,  $K_a\varphi$  means that *agent a knows  $\varphi$* ;  $Q_a\varphi$  means that *the answer to agent a's issue(s) entails  $\varphi$* ;  $Dk\varphi$  means that  $\varphi$  *is distributed knowledge*;  $Ck\varphi$  means that  $\varphi$  *is common knowledge*; and  $Gk\varphi$  means that  $\varphi$  *is potential group knowledge*.

The formal semantics is given by an *interpretation map* which associates with each sentence  $\varphi$  a proposition  $\|\varphi\|_{\mathbf{S}} \subseteq S$  in any given epistemic issue model  $\mathbf{S}$ . Intuitively,  $\|\varphi\|_{\mathbf{S}}$  is the set of all worlds in  $S$  satisfying  $\varphi$ . The definition is by induction, in terms of the obvious compositional clauses, starting by taking the valuation map  $\|p\|$  for atomic sentences and inductively using the Boolean operators on  $\mathcal{P}(S)$  and the epistemic operators defined above to compute the interpretation of complex sentences: e.g.  $\|K_a\varphi\| = K_a\|\varphi\|$ , etc.

**Proof system.** The axiomatic system  $LGK$  consists of the following:

- the rules and axioms of propositional logic,
- Necessitation Rules for all five modalities  $K_a, Q_a, Dk, Ck, Gk$ : e.g. from  $\varphi$  infer  $K_a\varphi$ , etc.
- Kripke's axiom for all five modalities  $K_a, Q_a, Dk, Ck, Gk$ : e.g.  $K_a(\varphi \Rightarrow \psi) \Rightarrow (K_a\varphi \Rightarrow K_a\psi)$ , etc.
- the  $T$  axioms (Factivity) for all five modalities  $K_a, Q_a, Dk, Ck, Gk$ : e.g.  $K_a\varphi \Rightarrow \varphi$ , etc.
- $Q_a\varphi \Rightarrow Q_aQ_a\varphi$
- $\neg Q_a\varphi \Rightarrow Q_a\neg Q_a\varphi$
- $K_a\varphi \Rightarrow K_aQ_a\varphi$
- $K_a\varphi \Rightarrow Dk\varphi$  (for all  $a \in G$ )
- $Ck\varphi \Rightarrow K_aCk\varphi$  (for all  $a \in G$ )
- $Ck(\varphi \Rightarrow \bigwedge_{a \in G} K_a\varphi) \Rightarrow (\varphi \Rightarrow Ck\varphi)$
- $Gk\varphi \Rightarrow DkQ_aGk\varphi$  (for all  $a \in G$ )
- $Gk(\varphi \Rightarrow \bigwedge_{a \in G} DkQ_a\varphi) \Rightarrow (\varphi \Rightarrow Gk\varphi)$

The system  $LGK^+$  is obtained by adding to  $LGK$  the following Positive Introspection Axioms:

- $K_a\varphi \Rightarrow K_aK_a\varphi$
- $Dk\varphi \Rightarrow DkDk\varphi$

(The corresponding positive introspection statements for  $Ck$  and  $Gk$  are provable in  $LGK^+$ .)

Finally, the system  $LGK^\pm$  is obtained by adding to  $LGK^+$  the following Negative Introspection Axioms:

- $\neg K_a\varphi \Rightarrow K_a\neg K_a\varphi$
- $\neg Dk\varphi \Rightarrow Dk\neg Dk\varphi$

(The corresponding negative introspection statements for  $Ck$  and  $Gk$  are provable in  $LGK^\pm$ .)

**Theorem 1** *Let  $G$  and  $\Phi$  be fixed, with  $|G| \geq 2$ . The system  $LGK$  is sound and complete over the class of interrogative epistemic models. Similarly,  $LGK^+$  is sound and complete over positively introspective models, and  $LGK^\pm$  is sound and complete over fully introspective models. Moreover, all these logics are decidable.*

**Proof:** See the Appendix.

One can also add to the syntax a “potential individual knowledge” modality  $K_a^G\varphi$  capturing the potential knowledge of agent  $a$  within group  $G$ . For the logic  $LaGK$  obtained by extending  $LGK$  with this modality, the semantics of  $K_a^G$  is given by the Kripke modality  $[\rightarrow_{Dk} \approx_a]$ , as explained in Section 4; while for the introspective versions  $LaGK^+$  and  $LaGK^\pm$ , obtained by extending  $LGK^+$  and respectively  $LGK^\pm$  with this modality, the semantics of  $K_a^G$  is given by the Kripke modality  $[(\rightarrow_{Dk} \approx_a)^+]$  (where  $R^+$  is the transitive closure of  $R$ ), as also explained in Section 4.

For this enriched syntax, the axiomatic system  $LaGK$  is obtained by adding to  $LGK$  the following axiom:

$$K_a^G\varphi \Leftrightarrow DkQ_a\varphi;$$

while the systems for the positively introspective version  $LaGK^+$ , and for the fully introspective version  $LaGK^\pm$ , are obtained by adding to the system  $LGK^+$ , and respectively to  $LGK^\pm$ , the following:

- Necessitation Rule for  $K_a^G$ : from  $\varphi$  infer  $K_a^G\varphi$ .
- Kripke’s axiom for  $K_a^G$ :  $K_a^G(\varphi \Rightarrow \psi) \Rightarrow (K_a^G\varphi \Rightarrow K_a^G\psi)$ .
- the  $T$  axiom (Factivity) for  $K_a^G$ :  $K_a^G\varphi \Rightarrow \varphi$ .
- $K_a^G\varphi \Rightarrow DkQ_aK_a^G\varphi$
- $K_a^G(\varphi \Rightarrow DkQ_a\varphi) \Rightarrow (\varphi \Rightarrow K_a^G\varphi)$ .

**Theorem 2** *Let  $G$  and  $\Phi$  be fixed, with  $|G| \geq 2$ . The system  $LaGK$  is sound and complete over the class of interrogative epistemic models over  $(G, \Phi)$ . Similarly,  $LaGK^+$  is sound and complete over the class of positively introspective models, and  $LaGK^\pm$  is sound and complete over the class of fully introspective models. Moreover, all these logics are decidable.*

**Proof:** See the Appendix.

### 4.3 Actualizing the Group’s Epistemic Potential

We use the concept ‘potential’ to indicate that we are dealing with epistemic states that can in principle be ‘actualized’, i.e. internalized by the agents. In this section we show how communication can convert potential group knowledge into common knowledge, and potential individual knowledge into actual individual knowledge.

In the spirit of Dynamic Epistemic Logic (Baltag et al., 1998; van Benthem, 2011; van Ditmarsch et al., 2007), we can extend our syntax with *dynamic modalities* for public announcements and information pooling. We only do this here for information pooling, since we are interested in formalizing the way in which potential knowledge is converted into actual knowledge by information pooling. So we extend each of our “static” logics  $L \in \{LGK, LaGK, LGK^+, LaGK^+, LGK^\pm, LaGK^\pm\}$  above to a dynamic version denoted by  $L!$ , and obtained by adding to the syntax a new construct operator  $[!]\varphi$ , whose intended meaning is that  $\varphi$  becomes true after information pooling within the group  $G$ . The semantics in the case of the general (non-introspective) logic is given by evaluating  $\varphi$  in the updated model  $\mathbf{S}!$ :

$$\| [!]\varphi \|_{\mathbf{s}} = \|\varphi\|_{\mathbf{s}!},$$

while for the (positively, or fully) introspective logics we evaluate  $\varphi$  in the model  $\mathbf{S}^{!+}$ :

$$\| [!]\varphi \|_{\mathbf{s}} = \|\varphi\|_{\mathbf{s}^{!+}}.$$

**Theorem 3** *The dynamic logics  $LGK!$ ,  $LaGK!$ ,  $LaGK^{!+}$ ,  $LaGK^{\pm!}$  have the same expressivity as the corresponding static logics (while the logics  $LGK^{!+}$  and  $LGK^{\pm!}$  appear to be more expressive<sup>8</sup> than their static fragments). Moreover, sound and complete axiomatizations are obtained for these logics by extending the above axiomatic systems with the following “Reduction Axioms”:*

- $[!]p \Leftrightarrow p$
- $[!]\neg\varphi \Leftrightarrow \neg[!]\varphi$
- $[!](\varphi \wedge \psi) \Leftrightarrow ([!]\varphi \wedge [!]\psi)$
- $[!]K_a\varphi \Leftrightarrow K_a^G[!]\varphi$
- $[!]K_a^G\varphi \Leftrightarrow K_a^G[!]\varphi$
- $[!]Dk\varphi \Leftrightarrow Dk[!]\varphi$
- $[!]Gk\varphi \Leftrightarrow Gk[!]\varphi$
- $[!]Ck\varphi \Leftrightarrow Gk[!]\varphi$

(where in  $LGK!$  the notation  $K_a^G$  is just an abbreviation for  $DkQ_a$ ).

<sup>8</sup>We don’t have a proof of this difference in expressivity, but this conclusion is suggestion by the fact that reduction laws do not exist for these logics



**Proof:** See the Appendix.

If we restrict ourselves to *ontic facts*, i.e. we take  $\varphi := p$  to be an atomic sentence in the above Reduction Axioms, we obtain nice characterizations of (both individual and group) potential knowledge in terms of its possible actualizations:

**Proposition 8** For atomic sentences  $p \in \Phi$ , we have:

$$K_a^G p \Leftrightarrow [!_G]K_a p.$$

In words: *an agent’s potential knowledge within a group is the knowledge that she can acquire (by extracting the relevant information) from the others’ testimony.* Similarly we have:

$$Gkp \Leftrightarrow [!_G]Ckp.$$

In words: *potential group knowledge is the information that can become common knowledge by information-sharing within the group* (given the limits posed by the agents’ different questions).

## 5 Related Work and Future Outlook

**Connections to Other Work in Dynamic Epistemic Logic.** On the formal level, our approach in this paper is fully in line with the dynamic-epistemic logic treatment of questions. The main goal of dynamic epistemic logics is to model informational actions, including the posing of questions and other acts of communication. In this context, it is natural to highlight the role of questions with respect to the potential or actual knowledge that may be obtained. In this paper, we built on the earlier work by Baltag (2001) on modelling queries and interrogative actions, on the formalization of the action of learning the answer to a question in (Baltag and Smets, 2009), and most importantly on the more recent work on epistemic issue models by (van Benthem and Minica, 2012; Minica, 2008). We took this last framework as the main building-block of our approach, while adding a semantic condition reflecting our view of questions as epistemic filters, and using this setting to formalize the Selective Learning Principle and to develop a new theory of group knowledge.

**Connections with Erotetic Logics and Inquisitive Semantics.** From a broader perspective, our investigation relates to the line of work on ‘erotetic logics’ including (Belnap Jr, 1966; Groenendijk and Stokhof, 1985; Wisniewski, 1995) as well as to the models for interrogation that recently culminated in the approach on inquisitive semantics (Groenendijk, 1999, 2007; Groenendijk et al., 2009; Ciardelli and Roelofsen, 2009). For details on the relations and differences between the dynamic-epistemic issue models and the inquisitive semantics account, we refer to (van Benthem and Minica, 2012; Minica, 2008). It is worth contrasting the aim of our proposal to that of Inquisitive Semantics, as given in (Ciardelli and Roelofsen, 2009) (Groenendijk, 2007) (Groenendijk and Stokhof, 1990). Inquisitive Semantics proposes a semantics to capture a layer of meaning that is not captured by the standard semantics of classical logic, namely, a layer of inquisitive meaning. On this picture, when an agent utters a sentence  $p$ , her utterance may have inquisitive content (raise a question) next to assertive content (assert a proposition). This second layer of meaning motivates new pragmatic principles that e.g. guide possible responses to an utterance. Here, however, we are not concerned with the role of questions as informational actions, but rather with the role of questions as a precondition for learning and the possibility of group knowledge.<sup>9</sup>

**Further Work: Testimonial Knowledge Conditions.** While we impose only one constraint (\*) on our interrogative epistemic models, in special cases one may want to impose two further ‘testimonial knowledge conditions’, which can be philosophically motivated as follows. There are restrictions on what agents can learn from the testimony of others (Coady, 1992). In particular, in order to rely on the testimony of a fellow researcher, and thus acquire knowledge from her, a scientist should (epistemically) trust the testifier. Testimonial knowledge is acquired from recognized experts or collaborators, as their testimonies are presumed to be trustworthy. In fact, Hardwig has argued that modern science rests (primarily) on trust (Hardwig, 1991).<sup>10</sup> Agents do not acquire knowledge from the testimony of others if they do not recognize them as experts. Furthermore, if an expert in the field sets the stage and looks for answers in a certain direction, then all his followers/collaborators will do the same.

<sup>9</sup>We further remark that there are essential technical differences between our setting and the work in inquisitive semantics.

<sup>10</sup>In his words, “the trustworthiness of members of epistemic communities is the ultimate foundation for much of our knowledge” (694).

The testimonial knowledge conditions are thus meant to capture the epistemic dependence among agents, i.e. the fact that agents explicitly rely on the knowledge (and questions) of others. In order to acquire knowledge through the testimony of group members, agents must recognize the other group members as experts, and hence be interested in their knowledge and questions. Thus, the first further condition we may want to impose is that for agent  $a$  to view the other group members as experts and collaborators, *what the others know about her issues must be a part of her issues*;  $a$ 's questions should include "what do others know about my questions?". Formally, this can be captured by the requirement that  $s \approx_a \rightarrow_b s'$  implies  $s \rightarrow_b \approx_a s'$ . This condition is important, not only for allowing an agent to (statically) represent the others' knowledge, but also for allowing her to dynamically enrich her knowledge, by learning the relevant facts entailed by the others' testimonies. Similarly, a second condition can require *an epistemic agent to be able to learn from the testimony of other agents only if she is interested in what (the answers to) the others' questions can tell about her own questions*. So  $a$ 's questions should include "what can the answers to the others' questions tell me about my questions?". Since this last question is encoded by the composite relation  $\approx_b \approx_a$ , we can formally capture this by the requirement that  $s \approx_a \approx_b s'$  implies  $s \approx_b \approx_a s'$ . This second "social rationality" condition is also important for learning: it allows the agent the possibility of learning relevant facts, not for only from the other agents' current testimonies (about their present state of knowledge), but also from any of their future testimonies (after they had learned more about *their* relevant questions).

In this paper we aimed at maximum generality, so we did not assume any of the special testimonial knowledge conditions mentioned above. But this line of inquiry is pursued in the third author's Master thesis (Boddy, 2014), and will form the topic of a future publication.

**Changing the Agenda: Paradigm shifts and Issue Negotiation.** In this paper, each agent's issues were fixed, playing only a directing role in the agent's learning. But one can also look at actions that *change the agents' conceptual space* by e.g. raising new questions, negotiating issues etc. For example, (van Benthem and Minica, 2012) study the action of *publicly raising a question*. They define it an action  $?_a \varphi$  that *adds question  $\varphi?$  to all agendas*, thus changing all the issue relations. The new issue relations are given by  $\approx_a^{? \varphi} := \approx_a \cap \equiv_{\varphi}$ . In other words,  $\varphi?$  becomes a relevant issue for everybody. Another example that we plan to investigate is the action  $?_b$ , by which *agent  $b$  convinces everybody to include all her questions into the group's agenda*. This changes the issue relations to  $\approx_a^{? b} := \approx_a \cap \approx_b$ . Even more complex actions of *negotiating the issues* can also be introduced. Based on these and other examples, we can make a general distinction between *two types of learning*, namely learning of new facts (propositions) versus learning of new issues (questions). The former type corresponds to eliminating worlds  $s$  from the model, as in the usual updates by public announcements; the latter type of learning entails the re-evaluation of the agents' conceptual frameworks. In on-going work, we investigate the behavior of group knowledge under issue-changing actions. We think this line of work can play a role in the logical-epistemological understanding of key topics in History and Philosophy of Science, such as *paradigm change* and the possibility of *interdisciplinary cooperation* (between members of different scientific communities, with initially different interrogative agendas).

## 6 Acknowledgments.

The contribution of Sonja Smets to this paper has received funding by the European Research Council under the European Community's Seventh Framework Programme ERC Grant no. 283963. We also want to thank Johan van Benthem, Wes Holliday, Kevin Kelly, Christian List and Erik Olsson for illuminating discussions on topics related to this paper. Early versions of this work were presented at the 2nd workshop on Questions and Inquiry LoQI 2013; the workshop on Game Theory, Epistemic Logic and Related Topics 2013; LORI IV 2013; the KNAW Dependence Logic workshop 2014; the Symposium on Epistemic Logic and Formal Epistemology at the Pacific APA meeting 2014; Methods and Modality 2014; Esslli 2014 workshop on Epistemic Logic for Individual, Social, and Interactive Epistemology; 16th Szklarska Poreba workshop 2015 and the 2nd Belgrade graduate conference in philosophy and logic 2015. We thank the organizers of these workshops for the opportunity to present this work. In particular we thank Julia Staffel for the commentary presented on this work at the APA meeting in 2014. A closely related (though more intricate) framework (going beyond partitional questions and involving a number of additional testimonial knowledge conditions that increase the epistemic compatibility of a group) was presented in the third author's ILLC Master thesis (Boddy, 2014). We thank to the members of her thesis committee, Maria Aloni, Johan van Benthem, Branden Fitelson, Wesley Holliday, Robert van Rooij and Martin Stokhof, for their useful comments and feedback.

## References

- M. Aloni, P. Egré, and T. De Jager. Knowing whether a or b. *Synthese*, 190(14):2595–2621, 2013.
- A. Baltag. Logics for insecure communication. In *Proceedings of the 8th conference on Theoretical aspects of rationality and knowledge*, pages 111–121. Morgan Kaufmann Publishers Inc., 2001.
- A. Baltag and S. Smets. Learning by questions and answers: From belief-revision cycles to doxastic fixed points. In *Logic, Language, Information and Computation*, pages 124–139. Springer, 2009.
- A. Baltag, L.S Moss, and S. Solecki. The logic of public announcements, common knowledge, and private suspicions. In *Proceedings of the 7th conference on Theoretical aspects of rationality and knowledge*, pages 43–56. Morgan Kaufmann Publishers Inc., 1998.
- A. Baltag, H. van Ditmarsch, and L. Moss. Epistemic logic and information update. *Handbook of the Philosophy of Information, Elsevier Science Publishers, Amsterdam*, pages 361–456, 2008.
- J. Barwise. Three views of common knowledge. In *Proceedings of the 2nd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 365–379. Morgan Kaufmann Publishers Inc., 1988.
- N. D Belnap Jr. Questions, answers, and presuppositions. *The Journal of Philosophy*, 63(20):609–611, 1966.
- P. Blackburn, M. De Rijke, and Y. Venema. *Modal Logic: Graph. Darst*, volume 53. Cambridge University Press, 2002.
- R. Boddy. Epistemic issues and group knowledge. Master’s thesis, University of Amsterdam., 2014.
- I. Ciardelli and F. Roelofsen. Generalized inquisitive logic: Completeness via intuitionistic kripke models. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge*, pages 71–80. ACM, 2009.
- C.A. Coady. *Testimony: A philosophical study*. Oxford University Press, 1992.
- A. J Corlett. *Analyzing social knowledge*. Rowman & Littlefield Publishers, 1996.
- J. de Ridder. Epistemic dependence and collective scientific knowledge. *Synthese*, 191(1):37–53, 2014.
- M. Fagan. Collective scientific knowledge. *Philosophy Compass*, 7(12):821–831, 2012.
- R. Fagin, J. Y. Halpern, and M. Y. Vardi. What can machines know?: On the properties of knowledge in distributed systems. *Journal of the ACM (JACM)*, 39, 1992.
- R. Fagin, Y. Moses, M.Y Vardi, and J.Y Halpern. *Reasoning about knowledge*. MIT press, 2003.
- M. Gilbert. Modelling collective belief. *Synthese*, 73(1):185–204, 1987.
- A. Goldman. Group knowledge versus group rationality: Two approaches to social epistemology. *Episteme*, 1(01): 11–22, 2004.
- A. Goldman and D. (Eds.) Whitcomb. *Social epistemology: essential readings*. Oxford University Press, 2011.
- J. Groenendijk. The logic of interrogation: Classical version. In *Semantics and linguistic theory*, volume 9, pages 109–126, 1999.
- J. Groenendijk. *Inquisitive semantics: Two possibilities for disjunction*. Springer, 2007.
- J. Groenendijk and M. Stokhof. On the semantics of questions and the pragmatics of answers. *Semantics: Critical Concepts in Linguistics*, page 288, 1985.
- J. Groenendijk and M. Stokhof. Partitioning logical space. *Annotated handout*, 1990.
- J. Groenendijk, F. Roelofsen, et al. Inquisitive semantics and pragmatics. In *workshop on language, communication, and rational agency at Stanford*, 2009.

- J. Hardwig. The role of trust in knowledge. *The Journal of Philosophy*, 88(12):693–708, 1991.
- J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, 1962.
- J. Hintikka. *Socratic epistemology: Explorations of knowledge-seeking by questioning*. Cambridge University Press, 2007.
- T. K. Kelly. *The Logic of Reliable Inquiry*. Oxford University Press, 1996.
- D. Lewis. *Convention: A philosophical study*. Harvard University Press., 1969.
- C. List. Group knowledge and group rationality: a judgment aggregation perspective. *Episteme*, 2(01):25–38, 2005.
- C. List, P. Pettit, et al. *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press Oxford, 2011.
- S. Minica. *Dynamic Logic of Questions. PhD Dissertation*. PhD thesis, ILLC, University of Amsterdam., 2008.
- E. J Olsson and D. Westlund. On the role of the research agenda in epistemic change. *Erkenntnis*, 65(2):165–183, 2006.
- A. Quinton. The presidential address: Social objects. In *Proceedings of the Aristotelian society*, volume 76, pages 1–viii. JSTOR, 1975.
- K. Rolin. Science as collective knowledge. *Cognitive Systems Research*, 9(1):115–124, 2008.
- J. Schaffer. Contrastive knowledge. *Berlin Studies in Knowledge Research*, page 357, 2005.
- J. Schaffer. Knowing the answer. *Philosophy and Phenomenological Research*, 75(2):383–403, 2007.
- J. van Benthem. *Logical dynamics of information and interaction*. Cambridge University Press, 2011.
- J. van Benthem and S. Minica. Toward a dynamic logic of questions. In *Journal of Philosophical Logic*, 41(4), pp.633-669, 2012.
- H. van Ditmarsch, W. van Der Hoek, and B. Kooi. *Dynamic epistemic logic*, volume 337. Springer Science & Business Media, 2007.
- A. Wegener. Die entstehung der kontinente, lecture held in 1912 for the general meeting of the geologische vereinigung in frankfurt. translated by roland von huene as 'the origins of continents' in 2002. *Int J Earth Sci (Geol Rundsch)*, 91:S4S17, 2002.
- A. Wisniewski. *The Posing of Questions, Logical Foundations of Erotetic Inferences*. Synthese Library. Kluwer., 1995.
- K B. Wray. Who has scientific knowledge? *Social Epistemology*, 21(3):337–347, 2007.

## APPENDIX: Completeness and Decidability (Proofs)

The proofs for Theorems 1, 2 and 3 follow the standard method used in the completeness proof for epistemic logic with common knowledge and distributed knowledge, as introduced in (Fagin et al., 1992). The proofs' plan is as follows:

1. **Completeness for the Canonical Structure.** First, we define a *canonical structure*. This is a Kripke structure, generated from the syntax in the usual way: possible worlds are maximally consistent theories. The canonical structure is *not* an interrogative epistemic model, since each modality has its own independent accessibility relation, that cannot be reduced to a combination of the other relations. Nevertheless, our axioms are sound for this structure, and moreover they are complete: every maximal consistent theory satisfies itself.
2. **Decidability via Finite Pseudo-Model Property.** We use our version of a Fischer-Ladner closure to filtrate the canonical structure, obtaining a *finite* “pseudo-model”: this is a structure in which only the individual epistemic and issue modalities and the distributed-knowledge modality have their own independent accessibility relations  $\rightarrow_a, \approx_a, \rightarrow_{Dk}$  (with  $\rightarrow_{Dk} \subseteq \rightarrow_a$ ), while the other modalities are defined exactly as in interrogative epistemic models. This gives us soundness and completeness of our logics with respect to finite pseudo-models, which implies the decidability of these logics.
3. **Unraveling.** In the third step, we *unravel* the pseudo-model obtained in the previous step: this means that we create *all* possible histories in the pseudo-model (all paths that can be taken when we follow these one-step relations  $\rightarrow_a, \approx_a, \rightarrow_{Dk}$ ), thus forming a tree. Then we redefine the relations  $\rightarrow_a$  in order to ensure that they include the relation  $\rightarrow_{Dk}$ ; if necessary (for the introspective versions), we also close the relations  $\rightarrow_a$  and  $\approx_a$  under transitivity (or transitivity and symmetry). Finally, we take the interrogative epistemic model obtained in this way (in which all the other relations are defined as in any epistemic issue model). We define a bounded morphism from this model into the finite pseudo-model obtained in the previous step. Completeness for interrogative epistemic models follows immediately.
4. **Completeness for the dynamic logics (proof of Theorem 3)** In the last step we show that the dynamic logics can be reduced to their static counterparts, using the reduction axioms and a translation mechanism, which gives us completeness for these logics.

We now give a sketch of the main definitions and results in each step.

### STEP 1: Soundness and Completeness for the Canonical Structure

The *canonical structure for LGK* is a Kripke structure

$$M = (S, \{\rightarrow_a : a \in G\}, \{\approx_a : a \in G\}, \rightarrow_{Dk}, \rightarrow_{Ck}, \rightarrow_{Gk}, \|\bullet\|), \text{ where}$$

- $S$  consists of all maximally consistent sets of formulas (“theories”) of *LGK*;
- $s \rightarrow_a t$  iff  $\forall \varphi (K_a \varphi \in s \Rightarrow \varphi \in t)$ ;
- $s \approx_a t$  iff  $\forall \varphi (Q_a \varphi \in s \Rightarrow \varphi \in t)$ ;
- $s \rightarrow_{Dk} t$  iff  $\forall \varphi (Dk \varphi \in s \Rightarrow \varphi \in t)$ ;
- $s \rightarrow_{Ck} t$  iff  $\forall \varphi (Ck \varphi \in s \Rightarrow \varphi \in t)$ ;
- $s \rightarrow_{Gk} t$  iff  $\forall \varphi (Gk \varphi \in s \Rightarrow \varphi \in t)$ ;
- $\|p\| = \{s \in S; p \in s\}$ .

It is easy to see that all the relations are reflexive. When the Positive Introspection axioms of  $LGK^+$  are added, one can show that all the relations are transitive; and when the Negative Introspection axioms of  $LGK^\pm$  are added as well, then all the relations are equivalence relations. For *LaGK* (and its introspective variants), the canonical structure has to be enriched with another independent accessibility relation  $\rightarrow_a^G$ , defined in a similar way to the others. The semantics is obvious: each modality is interpreted as the Kripke modality for its associated relation. Soundness and completeness with respect to the canonical structure are proved in the usual way, via the Truth Lemma<sup>11</sup>: each theory  $t \in S$  satisfies a formula  $\varphi$  iff  $\varphi \in t$ .

<sup>11</sup>See (Blackburn et al., 2002) for the definition and properties of canonical models.

## Step 2: Decidability via Finite Pseudo-Model Property.

A *pseudo-model for LGK* is a Kripke structure  $M = (S, \{\rightarrow_a : a \in G\}, \{\approx_a : a \in G\}, \rightarrow_{Dk})$ , satisfying the following conditions: all relations are reflexive;  $\rightarrow_a \approx_a \subseteq \rightarrow_a$ ; and  $\rightarrow_{Dk} \subseteq \rightarrow_a$ . A *pseudo-model for LGK<sup>+</sup>* is a pseudo-model for *LGK* in which all the relations are transitive. A *pseudo-model for LGK<sup>±</sup>* is a pseudo-model for *LGK* in which all the relations are equivalence relations.

Given a pseudo-model  $M$ , the semantics of our basic language is given by interpreting the modalities  $K_a, Q_a$  and  $Dk$  as Kripke modalities for the corresponding relations, and defining the modalities  $Ck$  and  $Dk$  in the same way as on interrogative epistemic models.

For the language *LaGK*, we also define the modality  $K_a^G$  in the same way as on interrogative epistemic models (using the appropriate definitions for each of the logics *LaGK*, *LaGK<sup>+</sup>*, *LaGK<sup>±</sup>*).

It is easy to see that *each of these logics is sound with respect to the (corresponding) class of pseudo-models*.

Let now  $M$  be the canonical structure constructed in Step 1, let  $\varphi$  be a consistent formula, and let  $s_\varphi$  be a maximally consistent theory that contains  $\varphi$ . The *Fisher-Ladner closure of  $\varphi$*  is the smallest set of formulas  $\Sigma_\varphi$  satisfying the following conditions:  $\varphi \in \Sigma_\varphi$ ;  $\Sigma_\varphi$  is closed under subformulas;  $\Sigma_\varphi$  is closed under single negations (i.e. if  $\psi \in \Sigma_\varphi$  is *not* of the form  $\neg\theta$ , then  $\neg\psi \in \Sigma_\varphi$ ); if  $K_a\psi \in \Sigma_\varphi$  then  $Dk\psi \in \Sigma_\varphi$ ; if  $K_a\psi \in \Sigma_\varphi$  and  $\psi$  is *not* of the form  $Q_a\theta$ , then  $K_aQ_a\psi \in \Sigma_\varphi$ ; if  $Ck\psi \in \Sigma_\varphi$  then  $\bigwedge_{a \in G} K_aCk\psi \in \Sigma_\varphi$ ; finally, if  $Gk\psi \in \Sigma_\varphi$  then  $\bigwedge_{a \in G} DkQ_aGk\psi \in \Sigma_\varphi$ . For *LaGK*, we need to add another closure requirement:  $K_a^G\psi \in \Sigma_\varphi$  implies  $DkQ_aK_a^G\psi \in \Sigma_\varphi$ .

It is easy to see that *the Fisher-Ladner closure of  $\varphi$  is always finite*.

We define an equivalence relation on the set  $S$  of all maximally consistent theories in the canonical structure  $M$  above, by putting:

$$s \cong_\varphi t \text{ iff } s \cap \Sigma_\varphi = t \cap \Sigma_\varphi.$$

We can now define a *filtration*  $M^f = (S^f, \{\rightarrow_a^f : a \in G\}, \{\approx_a^f : a \in G\}, \rightarrow_{Dk}^f, \rightarrow_{Ck}^f, \rightarrow_{Gk}^f, \|\bullet\|^f)$  of the canonical structure  $M$  above, by putting:

- $S^f = \{[s] : s \in S\}$ , where  $[s]$  is the equivalence class of  $s$  modulo  $\cong_\varphi$ ;
- $[s] \rightarrow_a^f [t]$  if and only if  $\forall K_a\psi \in \Sigma_\varphi (K_a\psi \in s \Rightarrow \psi \in t)$ ; for *LGK<sup>+</sup>*, we use instead a transitive version:  $[s] \rightarrow_a^f [t]$  if and only if  $\forall K_a\psi \in \Sigma_\varphi (K_a\psi \in s \Rightarrow K_a\psi \in t)$ ; while for *LGK<sup>±</sup>*, we use a transitive and symmetric version:  $[s] \rightarrow_a^f [t]$  if and only if  $\forall K_a\psi \in \Sigma_\varphi (K_a\psi \in s \Leftrightarrow K_a\psi \in t)$ ;
- $[s] \rightarrow_{Dk}^f [t]$  if and only if  $\forall Dk\psi \in \Sigma_\varphi (Dk\psi \in s \Rightarrow \psi \in t)$ ; again, we use the corresponding transitive version for *LGK<sup>+</sup>*, and the transitive and symmetric version for *LGK<sup>±</sup>*;
- $[s] \approx_a^f [t]$  if and only if  $\forall Q_a\psi \in \Sigma_\varphi (Q_a\psi \in s \Leftrightarrow Q_a\psi \in t)$ ;
- $\rightarrow_{Ck}^f = (\bigcup_{a \in G} \rightarrow_a^f)^+$ ;
- $\rightarrow_{Gk}^f = (\bigcup_{a \in G} \rightarrow_{Dk}^f \approx_a^f)^+$ ;
- for *LaGK*, *LaGK<sup>+</sup>*, *LaGK<sup>±</sup>*, we add the appropriately filtrated version for  $\rightarrow_a^G$ ;
- $\|p\| = \{[s] : s \in \|p\|_M\} = \{[s] : p \in s\}$ , for  $p \in \Phi \cap \Sigma$ , and  $\|p\| = \emptyset$  for  $p \in \Phi \setminus \Sigma$ .

It is now easy to check that  $M^f$  is a finite pseudo-model, and that it is indeed a filtration<sup>12</sup> of the canonical Kripke structure  $M$ .

By the usual properties of filtration, it follows that, for all formulas  $\psi \in \Sigma_\varphi$ , we have:

$$[s] \models_{M^f} \psi \text{ iff } s \models_M \psi \text{ iff } \psi \in s.$$

Hence,  $\varphi$  is satisfied at state  $[s_\varphi]$  in  $M^f$ .

This gives us completeness with respect to finite pseudo-models, and thus decidability.

<sup>12</sup>See (Blackburn et al., 2002) for the definition and properties of filtrations.

### Step 3: Unraveling.

Let  $\varphi, s_\varphi, \Sigma_\varphi$  be as above, and  $M^f$  be the finite pseudo-model constructed in Step 2. A *history in  $M^f$*  (with origin  $[s_\varphi]$ ) is any finite sequence  $h := ([s_0], R_0, [s_1], \dots, R_{n-1}, [s_n])$  such that

- for all  $k \leq n : [s_k] \in S^f$ ;
- $[s_0] = [s_\varphi]$ ;
- for all  $k < n : R_k \in \{\rightarrow_a^f : a \in G\} \cup \{\approx_a^f : a \in G\} \cup \{\rightarrow_{Dk}^f\}$ ;
- for all  $k < n : [s_k]R_k[s_{k+1}]$ .

For any history  $h := ([s_0], R_0, [s_1], \dots, R_{n-1}, [s_n])$ , we write  $last(h) = [s_n]$ .

The *unraveling of  $M^f$  around  $[s_\varphi]$*  is a tree-structure  $\vec{M} = (\vec{S}, R_{\rightarrow_a}, R_{\approx_a}, R_{\rightarrow_{Dk}})$  such that

- $\vec{S} = \{h; h \text{ history in } M^f\}$ ;
- $hR_{\rightarrow_a}h'$  iff  $h' = (h, \rightarrow_a^f, [s'])$  for some  $s'$ ;
- $hR_{\approx_a}h'$  iff  $h' = (h, \approx_a^f, [s'])$  for some  $s'$ ;
- $hR_{\rightarrow_{Dk}}h'$  iff  $h' = (h, \rightarrow_{Dk}^f, [s'])$  for some  $s'$ .

We now convert  $\vec{M}$  into an *interrogative epistemic model*  $\vec{M} = (\vec{S}, \rightarrow, \approx, \|\bullet\|)$  as follows:

- we put  $\approx_a := (R_{\approx_a} \cup R_{\approx_a}^{-1})^*$ , i.e. the least equivalence relation that includes  $R_{\approx_a}$ ;
- for  $LGK$ , we take  $\rightarrow_a := \approx_a \cup (R_{\rightarrow_a} \approx_a) \cup (R_{\rightarrow_{Dk}} \approx_a)$ ; for  $LGK^+$ , we take  $\rightarrow_a := (\approx_a \cup R_{\rightarrow_a} \cup R_{\rightarrow_{Dk}})^*$ ; and for  $LGK^\pm$ , we put  $\rightarrow_a := (\approx_a \cup R_{\rightarrow_a} \cup R_{\rightarrow_a}^{-1} \cup R_{\rightarrow_{Dk}} \cup R_{\rightarrow_{Dk}}^{-1})^*$ , i.e. the least equivalence relation that includes  $R_{\approx_a} \cup R_{\rightarrow_a} \cup R_{\rightarrow_{Dk}}$ .
- $\|p\| = \{h \in \vec{S} : last(h) \in \|p\|_{M^f}\}$ .

It is easy now to check that *the map  $last : \vec{S} \rightarrow S^f$ , which maps each history  $h \in \vec{S}$  into its last element  $last(h) \in S^f$ , is a bounded morphism<sup>13</sup> between pseudo-models  $M^f$  and  $\vec{M}$* . Since bounded morphisms preserve the truth of modal formulas, it follows that our consistent formula  $\psi$  is satisfied at history  $([s_\varphi])$  in the interrogative epistemic model  $\vec{M}$ . This finishes our completeness proof for the static logics (Theorems 1 and 2).

### Step 4: Completeness for the dynamic logics

We now prove Theorem 3. It is easy to see that the *Reduction Axioms are sound*. For completeness, we show that there exists a *translation function  $t$*  from formulas of our dynamic languages to formulas in the corresponding static languages, defined recursively as follows:  $t(p) = t(!p) := p$ ;  $t(\neg\varphi) := \neg t(\varphi)$ ;  $t(\varphi \wedge \psi) := t(\varphi) \wedge t(\psi)$ ;  $t(\Delta\varphi) := \Delta t(\varphi)$  for all operators  $\Delta \in \{K_a, Q_a, Dk, Ck, Gk, K_a^G\}$ ;  $t(!\neg\varphi) := t(\neg!\varphi)$ ;  $t(!(\varphi \wedge \psi)) := t(!\varphi) \wedge t(!\psi)$ ;  $t(!K_a\varphi) = t(!K_a^G\varphi) := t(K_a^G!\varphi)$ ;  $t(!Dk\varphi) := t(Dk!\varphi)$ ;  $t(!Ck\varphi) = t(!Gk\varphi) := Gkt(!\varphi)$ .

Now we can show that, for every formula  $\varphi$  of our dynamic languages  $LGK!$  and  $LaGK!$ , we have:

- $t(\varphi)$  is a formula in the corresponding static language ( $LGK$  or  $LaGK$ );
- the formula  $\varphi \Leftrightarrow t(\varphi)$  is a theorem in the corresponding dynamic logic  $LGK!$  or  $LaGK!$ .

The proof is by induction, using the Reduction Axioms. This result immediately implies Theorem 3.

<sup>13</sup>See (Blackburn et al., 2002) for the definition and properties of bounded morphisms.