

## The dynamics of awareness

Johan van Benthem ·  
Fernando R. Velázquez-Quesada

Received: 9 September 2009 / Accepted: 21 July 2010 / Published online: 6 August 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Classical epistemic logic describes implicit knowledge of agents about facts and knowledge of other agents based on semantic information. The latter is produced by acts of observation or communication that are described well by dynamic epistemic logics. What these logics do not describe, however, is how significant information is also produced by acts of inference—and key axioms of the system merely postulate “deductive closure”. In this paper, we take the view that all information is produced by acts, and hence we also need a dynamic logic of inference steps showing what effort on the part of the agent makes a conclusion explicit knowledge. Strong omniscience properties of agents should be seen not as static idealizations, but as the result of dynamic processes that agents engage in. This raises two questions: (a) how to define suitable information states of agents and matching notions of explicit knowledge, (b) how to define natural processes over these states that generate new explicit knowledge. To this end, we use a static base from the existing awareness literature, extending it into a dynamic system that includes traditional acts of observation, but also adding and dropping formulas from the current ‘awareness’ set. We give a completeness theorem, and we show how this dynamics updates explicit knowledge. Then we extend our approach to multi-agent scenarios where awareness changes may happen privately. Finally, we mention further directions and related approaches. Our contribution can be seen as a ‘dynamification’ of existing awareness logics.

---

J. van Benthem (✉) · F. R. Velázquez-Quesada  
Institute for Logic, Language and Computation, Universiteit van Amsterdam, Amsterdam,  
The Netherlands  
e-mail: j.vanbenthem@uva.nl

F. R. Velázquez-Quesada  
e-mail: f.r.velazquezquesada@uva.nl

J. van Benthem  
Department of Philosophy, Stanford University, Palo, Alto, USA

**Keywords** Epistemic logic · Considering dropping · Observation · Awareness · Dynamic epistemic logic

## 1 The problem of omniscience: what is ‘missing in action’

The usual discussions of the problem of omniscience in epistemic logic revolve around the distribution axiom  $K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$ . Is knowledge closed under drawing logical inferences? If it is, so the story goes, then we have idealized our knowing agents too much.

But this story is misleading on two accounts. First, with the usual semantics of epistemic logic, the  $K$  operator really just describes implicit semantic information of the agent, which definitely has the preceding closure property. The point is rather that closure need not hold for a related, but different intuitive notion, viz. explicit “aware-that” knowledge  $\text{Ex } \varphi$ , in some suitable sense to be defined. So, what we really need is not “epistemic logic bashing”, but a richer account of agents’ attitudes. Our first question, then, is how to define explicit knowledge.

But there is more. The interesting issue is not whether explicit knowledge has deductive closure. It is rather: “*what do agents have to do to make their implicit knowledge explicit?*” Consider the premises  $\text{Ex } (\varphi \rightarrow \psi)$ ,  $\text{Ex } \varphi$  of the distribution axiom, saying that the agent explicitly knows both  $\varphi \rightarrow \psi$  and  $\varphi$ . As explicit knowledge should be also implicit knowledge, these should imply  $K\psi$ , that is, the agent knows  $\psi$  implicitly. But crucially, in order to make this information explicit, the agent has to do some *work*, namely, perform an act of “awareness raising” that leads to  $\text{Ex } \psi$ . Stated more syntactically, the usual implication  $\text{Ex } (\varphi \rightarrow \psi) \rightarrow (\text{Ex } \varphi \rightarrow \text{Ex } \psi)$  contains a *gap* [ ]:

$$\text{Ex } (\varphi \rightarrow \psi) \rightarrow (\text{Ex } \varphi \rightarrow [ ] \text{Ex } \psi)$$

and in that gap, we should place an action. Note that, then, the agent is no longer omniscient, but she is not defective either: with the right repertoire of actions available, she can do awareness raising as needed.

This paper explores these ideas. We first recall some simple epistemic awareness models and their known axiomatization. Then, we explore some proposals for defining explicit knowledge, picking one that we will work with. Next, we define basic dynamic actions that modify our models, and provide examples of their behaviour, alone and in combination. Representing the actions in the language yields a sound and complete logic that clarifies our initial issues. We also develop some formal properties and raise some open problems. Then we move to the multi-agent case, developing tools for private and even unconscious versions of our actions, which were public so far. Finally, we relate our proposal to earlier ones, and end with conclusions and further directions, in particular, toward agents with beliefs that are modified by default inferences.

## 2 A static system for different agent attitudes

In the long-standing discussion about the omniscience problem (Sim 1997; Moreno 1998) there is general agreement that there should be a difference between *explicit*

knowledge of an agent, and the *implicit knowledge* that she only has in principle (Konolige 1984; Levesque 1984; Lakemeyer 1986; Vardi 1986). This section introduces the semantic models that will allow us to differentiate the two notions.

We assume that the reader is familiar with classical epistemic logic (*EL*; Hintikka 1962). We have already stated our motivation for working with this: even though this system fails for its intended interpretation of ‘full-blooded knowledge’, like so many logical systems, it has turned out quite adequate for other, perhaps originally non-intended interpretations. In particular, it deals well with implicit knowledge of the semantic range kind (cf. van Benthem and Martínez 2008).

In a nutshell, the language of epistemic logic extends that of propositional logic with an operator  $K$  creating new formulas  $K\varphi$  that are read as “the agent knows  $\varphi$ ”. Semantically, we consider a set of possible worlds carrying valuations, with an accessibility relation indicating which world or situations the agent considers possible. Here boolean connectives are interpreted as usual, and  $K\varphi$  is true at a world  $w$  iff  $\varphi$  is true in all the worlds that the agent considers possible from  $w$ .

Awareness logics (due to Fagin and Halpern (1988) and taken up in an extensive further literature) extend the base language of *EL* with an operator  $A\varphi$  saying that the agent “is aware of  $\varphi$ ” (in less psychological terms, she “entertains  $\varphi$ ”). Notice that this is a matter of attention, and does not imply any attitude pro or con: the agent may believe  $\varphi$ , but also reject it. Stated in other, but related terms, “awareness of” does not imply “awareness that”.

**Definition 2.1** (*Language  $\mathcal{L}$* ) Let  $\mathbb{P}$  be a set of atomic propositions. Formulas  $\varphi$  of the *epistemic awareness language  $\mathcal{L}$*  are given by

$$\varphi ::= p \mid A\varphi \mid \neg\varphi \mid \varphi \wedge \psi \mid K\varphi$$

with  $p \in \mathbb{P}$ . Other Boolean connectives  $\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$ , as well the existential modal operator ( $\hat{K}$ ) are defined as usual.

Formulas  $A\varphi$  are read as “the agent is aware of  $\varphi$ ”, and formulas  $K\varphi$  as “the agent knows  $\varphi$  implicitly”. The language is interpreted in epistemic models assigning to the agent a set of formulas in each world, representing the information she is aware of (Fagin and Halpern 1988).

**Definition 2.2** (*Semantic model*) An *epistemic awareness model* is a tuple  $M = \langle W, R, \mathbf{A}, V \rangle$  where

- $\langle W, R, V \rangle$  is a standard epistemic model: a set of worlds  $W$ , an accessibility relation  $R \subseteq (W \times W)$ , and a valuation  $V : \mathbb{P} \rightarrow \wp(W)$ .
- $\mathbf{A} : W \rightarrow \wp(\mathcal{L})$  is the “awareness” function giving the formulas that the agent ‘has in mind’.  $\mathbf{A}(w)$  is the *awareness set* at  $w$ .

As usual, a *pointed model*  $(M, w)$  also has a distinguished world  $w$ .

The semantic interpretation of formulas in  $\mathcal{L}$  is entirely as expected:

**Definition 2.3** Let  $(M, w)$  be a pointed semantic model with  $M = \langle W, R, \mathbf{A}, V \rangle$ . Atomic propositions and boolean connectives are interpreted as usual; for  $A\varphi$  and  $K\varphi$  we have:

$$\begin{aligned}(M, w) \models A\varphi & \text{ iff } \varphi \in \mathbf{A}(w) \\ (M, w) \models K\varphi & \text{ iff for all } u \in W, Rwu \text{ implies } (M, u) \models \varphi.\end{aligned}$$

On these models we can impose standard epistemic assumptions about the accessibility relation, such as reflexivity, transitivity, and symmetry. Moreover, further conditions can be imposed on the  $\mathbf{A}$ -sets, like closure under commutation for conjunction and disjunction, or being generated by some subset of atomic propositions.<sup>1</sup> Nevertheless, these requirements are orthogonal to our main concerns in this paper, and we will not assume any of them.

It is easy to visualize how our mixed models work:

### Example 1



In this one-world model, the agent knows implicitly that  $p$  and also that  $q$ . But while she is explicitly aware that  $p$  holds, she is not aware that  $q$  holds, so her explicit knowledge about  $p, q$  differs.

In models with more than one world, a genuine issue of interplay arises: is awareness introspective? Here is an appealing property: being ‘aware of’ is implicitly known:

*Weak introspection.* The formulas that the agent is aware of are preserved under epistemic accessibility: that is, if  $\varphi \in \mathbf{A}(w)$  and  $Rwu$ , then  $\varphi \in \mathbf{A}(u)$ .<sup>2</sup>

Nevertheless, as mentioned before, we will not require any property in our intended class of models.

The calculus of reasoning with these notions is standard. The valid formulas of  $\mathcal{L}$  are those provable in the minimal modal logic  $K$  for implicit knowledge. In the case of the ‘*weak introspection*’ property, we just need the following axiom:

$$A\varphi \rightarrow KA\varphi$$

Soundness and completeness are proved by standard techniques.

### 3 Defining explicit knowledge

*Explicit knowledge as a defined notion.* Combining *implicit* and *awareness* information produces several kinds of *explicit* information. Options start with the  $K\varphi \wedge A\varphi$

<sup>1</sup> Such conditions are studied in depth in [Fagin and Halpern \(1988\)](#) and, more recently, in [Halpern \(2001\)](#).

<sup>2</sup> Note the effect of this property in combination with properties of  $R$ . With preorders,  $\mathbf{A}$  becomes *persistent*; with equivalence relations,  $\mathbf{A}$  becomes a function from equivalence classes to sets of formulas.

of [Fagin and Halpern \(1988\)](#), which says that  $\varphi$  is implicitly known and explicitly considered. But there are others. Personally, we favour:

$$\text{Ex } \varphi := K(\varphi \wedge A\varphi)$$

which is equivalent to  $K\varphi \wedge KA\varphi$ .<sup>3</sup> Both definitions satisfy the desideratum that explicit knowledge should imply implicit knowledge. Moreover, when the accessibility relation is transitive and euclidean, we get the following forms of positive and negative implicit introspection:

$$\text{Ex } \varphi \rightarrow K\text{Ex } \varphi \quad \neg\text{Ex } \varphi \rightarrow K\neg\text{Ex } \varphi$$

In what follows, our main focus will be on the dynamics of awareness, and nothing very much hinges on the precise choice of the notion of explicit knowledge. Henceforth, we will work with the definition  $\text{Ex } \varphi := K(\varphi \wedge A\varphi)$  for explicit knowledge (“awareness that”).<sup>4</sup>

*Explicit knowledge as a primitive notion.* It may be useful at this stage to explain how we are deviating here from our earlier work ([van Benthem 2008](#); [Velázquez-Quesada 2009](#)). There we assumed a *primitive notion* of explicit knowledge, treated as a map assigning a set of formulas to each world, just like our  $A$  function here. For a proper representation of explicit knowledge, we then needed to assume that all formulas in such sets were implicitly known, and hence in particular, that they were true. But this property is not preserved by standard update operations, since an epistemic formula true in a given pointed model may not be true after a public announcement: say, an ignorance statement may now have turned to knowledge. Other natural actions may change truth values of formulas involving  $A$ . Based on these observations, we found a need for dropping all persistence of explicit knowledge under update, except for purely factual assertions without modal operators. Our current approach to defining explicit knowledge will turn out to circumvent this difficulty, since the definition automatically ‘recomputes’ what is explicit knowledge following a dynamic update with epistemic side-effects.

For now, we move to our second main issue, viz. how agents can ‘improve’ their current brand of knowledge about a proposition. This is not a matter of static implications between brands of implicit and explicit knowledge. As we have said before, the correct question to ask here is a dynamic one: *what does an agent have to do to upgrade her implicit knowledge?* To some readers, introducing the explicit actions that lead to more omniscient states may demystify them. In Conan Doyle’s detective stories, the explanation offered at the end turns Holmes’ ‘magical powers’ into a sequence of

<sup>3</sup> In general, this definition is not equivalent to that of Fagin and Halpern, but it is if we assume weak introspection plus reflexivity of accessibility. Then we get  $K(\varphi \wedge A\varphi) \leftrightarrow (K\varphi \wedge A\varphi)$ .

<sup>4</sup> Another candidate that appeals to us is  $K\varphi \wedge AK\varphi$ , where the implicit knowledge itself has the agent’s explicit attention. We will not pursue this option—but once again, our framework can describe the dynamics of many proposals.

simple observations and deductive acts, making the procedure “elementary, my dear Watson”. While this is true, it also underscores the power of successive small steps.

#### 4 Operations on epistemic awareness models

Our epistemic awareness models suggest a natural and simple dynamics. Though the agent is not logically omniscient anymore, she can get new information by various acts, including observation and inference. But we want to dig deeper. In line with our definition for explicit knowledge, it also makes sense to look for simple actions transforming models that can be put together to analyze more complex informational acts. We will see later on how these transform explicit knowledge.

*Defining the basic actions.* Our models have two separate components for representing information: the accessibility relation and the awareness sets. The following operations modify these components in a simple way, allowing us to define complex epistemic actions later on.

The *consider* operation represents an “awareness raising” action:

**Definition 4.1** (The *consider* operation) Let  $M = \langle W, R, A, V \rangle$  be a model and  $\chi$  any formula in  $\mathcal{L}$ . The model  $M_{+\chi} = \langle W, R, A', V \rangle$  is  $M$  with its awareness sets extended with  $\chi$ , that is,

$$A'(w) := A(w) \cup \{\chi\} \quad \text{for every } w \in W$$

‘Considering’ extends the propositions that an agent is aware of, but we can also define a *drop* operation with the opposite effect: reducing awareness sets. This fits with the operational idea that agents can expand and shrink the set of issues having their current attention.

**Definition 4.2** (The *drop* operation) Let  $M = \langle W, R, A, V \rangle$  be a model and  $\chi$  a formula in  $\mathcal{L}$ . The model  $M_{-\chi} = \langle W, R, A', V \rangle$  reduces  $M$ ’s awareness sets by removing  $\chi$ , that is,

$$A'(w) := A(w) \setminus \{\chi\} \quad \text{for every } w \in W$$

This operation can be seen as a form of ‘forgetting’, an action usually disregarded in dynamic-epistemic logic (but see [van Ditmarsch et al. \(2009\)](#) and [van Ditmarsch and French \(2009\)](#) for proposals).

The preceding actions affect what an agent is aware of. The next, known from dynamic-epistemic logic, modifies her implicit knowledge by discarding those worlds where some *observed* formula  $\chi$  fails:

**Definition 4.3** (*Implicit observation*) Let  $M = \langle W, R, A, V \rangle$  be a model and  $\chi$  a formula in  $\mathcal{L}$ . The model  $M!_{\chi} = \langle W', R', A', V' \rangle$  has

- $W' := \{w \in W \mid (M, w) \models \chi\}$       –  $R' := R \cap (W' \times W')$
- $A'(w) := A(w)$  for every  $w \in W'$       –  $V'(w) := V(w)$  for every  $w \in W'$

The observation is implicit because, although it removes worlds, it does not affect what the agent is aware of in the preserved ones.

*Building complex actions.* Complex actions can now be built by combining basic ones. As an example, it seems natural to think that a public observation of some fact is in fact done consciously, generating awareness. The corresponding operation of “explicit seeing” can be defined in our framework.

**Fact** The *explicit seeing* operation, analogous in its effect to a *Public Announcement* in *EL* (Plaza 1989; Gerbrandy 1999), can be defined by means of an implicit observation followed by an act of consideration:

$$M_{ES(\chi)} := (M \upharpoonright_{\chi})_{+\chi}$$

The definition also works if we interchange the order of the operations—the reason being that we are transforming two independent components of our models.<sup>5</sup>

*Preserving static constraints.* Though we have not imposed strong constraints on our static models, it is interesting to note that some reasonable requirements, like weak introspection and equivalence relations for accessibility, are preserved by our operations.

**Proposition 1** *The consider operation preserves weak introspection and equivalence relations.*

*Proof* The equivalence property of  $R$  is obviously preserved, since  $R$  is not modified. For weak introspection, take a world  $w$  in  $M_{+\chi}$  and any  $\varphi \in A'(w)$ . Suppose  $Rwu$ . If  $\varphi$  is already in  $A(w)$ , then  $\varphi \in A(u)$  because  $M$  satisfies the principle, and then  $\varphi \in A'(u)$  by the definition of  $A'$ . If  $\varphi$  is not in  $A(w)$ , then it should be  $\chi$  itself, which by definition is also in  $A'(u)$ .  $\square$

By a similar argument, the *drop* operation, too, preserves the two properties.

**Proposition 2** *Dropping preserves equivalence relations and weak introspection.*

Finally, our actions of implicit observation have the same effect:

**Proposition 3** *Implicit observation preserves weak introspection and equivalence relations.*

*Proof* Equivalence relations are preserved automatically since we go to a sub-model. Next, for weak introspection, use the fact that the sub-model  $M \upharpoonright_{\chi}$  has the same

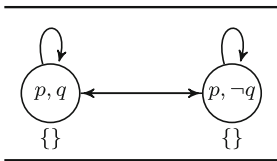
<sup>5</sup> Still, one might argue that implicit observation and considering take place *simultaneously*. While this makes sense, we will not pursue it here.

awareness sets at its worlds as  $M$ , while its epistemic accessibility is a subrelation of that for  $M$ . □

It is also worthwhile to notice how closure properties one might impose on the  $A$ -sets (cf. the already mentioned closure under commutation for conjunction and disjunction, or being generated by some subset of atomic propositions, as in cf. [Fagin and Halpern \(1988\)](#)) are not preserved by the operations *consider* and *drop*. Indeed, we think these properties over-idealize the content over agents' awareness.

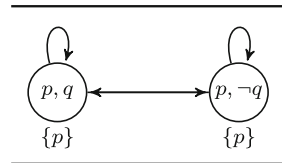
### 5 The actions in action

Consider the following model:

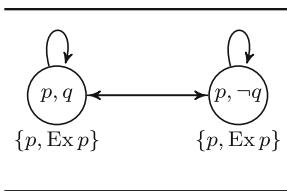


In the leftmost world the agent does not even know implicitly that  $q$ . But she knows implicitly that  $p$ , though not explicitly.

After the agent *considers*  $p$ , we get the model on the right: in both worlds, the agent now knows explicitly that  $p$ .

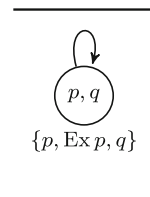


Lifting a restriction in [van Benthem \(2008\)](#), [Velázquez-Quesada \(2009\)](#), our agent can also get explicit knowledge about her own awareness, or implicit and explicit information. Here is how this can happen:



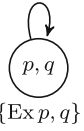
When she *considers*  $Ex\ p$ , we get the model on the left. The agent knows explicitly that she has explicit knowledge of  $p$ . By acting, she has achieved positive introspection.<sup>6</sup>

Next, consider the above *explicit seeing* of  $q$ : an *implicit observation* followed by *consideration* of  $q$ . This yields the model on the right where  $q$  is now explicitly known by the agent ( $Ex\ q$ ).



<sup>6</sup> We are cheating a little here, since  $Ex\ p$  is really a defined notion in the language.





Finally, *dropping*  $p$  makes the agent lose earlier explicit knowledge about it (in our case, we get  $\neg Ex p$ ). Moreover, by our definition of explicit knowledge, she no longer has explicit knowledge that  $Ex p$ , since the latter formula is no longer true, and therefore, it is no longer implicitly known.<sup>7</sup>

There are many further scenarios with complex many-world patterns, but the above will suffice to show the interest of our setting.

### 6 A complete dynamic logic

In order to express how our dynamic operations affect implicit knowledge, awareness, and explicit knowledge, we extend the static epistemic awareness language with modalities representing each basic operation. If  $\chi$  and  $\varphi$  are formulas in the resulting extended language (still called  $\mathcal{L}$  in this section), then so are

- $[+\chi]\varphi$     *after the agent considers  $\chi$ ,  $\varphi$  is the case.*
- $[-\chi]\varphi$     *after the agent drops  $\chi$ ,  $\varphi$  is the case.*
- $[!\chi]\varphi$     *after the agent implicit observes  $\chi$ ,  $\varphi$  is the case.*

**Definition 6.1** Let  $(M, w)$  be a pointed semantic model with  $M = \langle W, R, A, V \rangle$ , and let  $\chi, \varphi$  be formulas in the extended language  $\mathcal{L}$ :

$$\begin{aligned}
 (M, w) \models [+\chi]\varphi & \text{ iff } (M_{+\chi}, w) \models \varphi \\
 (M, w) \models [-\chi]\varphi & \text{ iff } (M_{-\chi}, w) \models \varphi \\
 (M, w) \models [!\chi]\varphi & \text{ iff } (M, w) \models \chi \text{ implies } (M_{!\chi}, w) \models \varphi
 \end{aligned}$$

The main difference among the new modalities is the precondition. The agent can consider or drop a formula  $\chi$  without any further requirement, but for her to implicitly observe  $\chi$ ,  $\chi$  needs to be *true*.

#### 6.1 Dynamic completeness theorem

We now formulate a sound and complete logic for the semantic validities in the extended language  $\mathcal{L}$ :

**Theorem 1** *The valid formulas of the dynamic-epistemic awareness language  $\mathcal{L}$  (in epistemic awareness models) are just those provable by the axioms and rules for the static base language (see Sect. 2) plus the reduction axioms and modal inference rules listed in Table 1.*

<sup>7</sup> This may seem strange since the formula  $Ex p$  still occurs in the world, but this only means that the agent is aware of it, not that she still endorses it.

*Proof* We use standard techniques from dynamic-epistemic logic (cf. [van Benthem et al. 2006](#)). Our axioms run through all cases needed to reduce a innermost occurrence of a dynamic operator. Iterating this produces an equivalent formula in the static base language.  $\square$

These axioms express the syntactic basics of the considering and dropping operations, merged with the axioms of public announcement logic *PAL* ([Plaza 1989](#); [Gerbrandy 1999](#)). For instance, how do the propositions that the agent is aware of change when the agent *considers*  $\chi$ ? Our axioms show the two possibilities. After considering  $\chi$ , the agent is aware of a  $\varphi \neq \chi$  iff she was aware of  $\varphi$  before; but also, considering  $\chi$  always makes the agent aware of  $\chi$ . The *drop* operation has an analogous effect in the opposite direction. The rest of the axioms are simple commutation clauses, indicating the independence of modifying the domain of worlds and the awareness sets.

### 6.2 How the logic describes our major issues

Our logic states how each basic operator of the language is affected by our three actions. By combining effects and unfolding definitions, the logic also explains how derived notions of *explicit knowledge* fare under these actions. We discuss a few cases, using our earlier definition  $K(\varphi \wedge A\varphi)$ , and suppressing detailed calculations:

*Explicit knowledge.* For the action of *considering*  $\chi$  and explicit knowledge about a different formula  $\varphi$ , an application of the reduction axioms gives us the following valid principle

$$[+\chi]Ex \varphi \leftrightarrow K([+\chi]\varphi \wedge A\varphi) \quad (\text{for } \varphi \neq \chi).$$

We leave it to the reader to put this analysis into words. One might have expected a simpler direct reduction principle  $[+\chi]Ex \varphi \leftrightarrow Ex \varphi$ , but this formula is not valid, since the *consider* action may have changed truth values for sub-formulas of  $\varphi$ .

**Table 1** Sound and complete logic for dynamic epistemic awareness logic

|   |   |   |   |
|---|---|---|---|
| $\vdash [+\chi]p$                                   | $\leftrightarrow p$                                   | $\vdash [-\chi]p$                                   | $\leftrightarrow p$                                 |
| $\vdash [+\chi]A\varphi$                            | $\leftrightarrow A\varphi$ for $\varphi \neq \chi$    | $\vdash [-\chi]A\varphi$                            | $\leftrightarrow A\varphi$ for $\varphi \neq \chi$  |
| $\vdash [+\chi]A\chi$                               | $\leftrightarrow \top$                                | $\vdash [-\chi]A\chi$                               | $\leftrightarrow \perp$                             |
| $\vdash [+\chi]\neg\varphi$                         | $\leftrightarrow \neg[+\chi]\varphi$                  | $\vdash [-\chi]\neg\varphi$                         | $\leftrightarrow \neg[-\chi]\varphi$                |
| $\vdash [+\chi](\varphi \wedge \psi)$               | $\leftrightarrow [+\chi]\varphi \wedge [+\chi]\psi$   | $\vdash [-\chi](\varphi \wedge \psi)$               | $\leftrightarrow [-\chi]\varphi \wedge [-\chi]\psi$ |
| $\vdash [+\chi]K\varphi$                            | $\leftrightarrow K[+\chi]\varphi$                     | $\vdash [-\chi]K\varphi$                            | $\leftrightarrow K[-\chi]\varphi$                   |
| From $\vdash \varphi$ infer $\vdash [+\chi]\varphi$ |   | From $\vdash \varphi$ infer $\vdash [-\chi]\varphi$ |   |
| $\vdash [!\chi]p$                                   | $\leftrightarrow \chi \rightarrow p$                  |   |   |
| $\vdash [!\chi]A\varphi$                            | $\leftrightarrow \chi \rightarrow A\varphi$           |   |   |
| $\vdash [!\chi]\neg\varphi$                         | $\leftrightarrow \chi \rightarrow \neg[!\chi]\varphi$ |   |   |
| $\vdash [!\chi](\varphi \wedge \psi)$               | $\leftrightarrow [!\chi]\varphi \wedge [!\chi]\psi$   |   |   |
| $\vdash [!\chi]K\varphi$                            | $\leftrightarrow \chi \rightarrow K[!\chi]\varphi$    |   |   |
| From $\vdash \varphi$ infer $\vdash [!\chi]\varphi$ |   |   |   |

In the particular case of explicit knowledge about  $\chi$  itself, however, we get the following.

**Fact** The formula  $[+\chi]\text{Ex } \chi \leftrightarrow K \chi$  is valid.

*Proof* Using our reduction axioms, we get

$$\begin{aligned} [+\chi]\text{Ex } \chi &\leftrightarrow [+\chi]K(\chi \wedge A\chi) \\ &\leftrightarrow K[+\chi]\chi \wedge K[+\chi]A\chi \\ &\leftrightarrow K[+\chi]\chi \end{aligned}$$

But the latter is equivalent to  $K \chi$ , since we have quite generally that

$$\text{the formula } \chi \leftrightarrow [+\chi]\chi \text{ is valid.}$$

The reason is that, given our semantics, an act of considering  $\chi$  can only change truth values for  $A\chi$  and formulas containing it. But the formula  $\chi$  itself is too short to be affected by this. □

This shows how a *consider* action makes implicit knowledge explicit. Next, going back to the initial discussion about the distribution axiom, the formula

$$\text{Ex } (\varphi \rightarrow \psi) \rightarrow (\text{Ex } \varphi \rightarrow K \psi)$$

is valid, since the antecedent implies the premises of the distribution law for  $K$ . Then, *considering* is the action that ‘fills the gap’:

$$\text{Ex } (\varphi \rightarrow \psi) \rightarrow (\text{Ex } \varphi \rightarrow [+ \psi]\text{Ex } \psi) \text{ is valid}$$

One might think that the real act here is a richer one of *drawing the inference*, but in our analysis it is the explicit consideration of the conclusion that ‘gives the last little push’ toward explicit knowledge.

Thus, our current proposal realizes the intuitive expectations in our Introduction. But it can describe more, including the behaviour of explicit knowledge under the *drop* operation. Here is what happens with formulas  $\varphi$  that differ from the dropped  $\chi$ :

$$[-\chi]\text{Ex } \varphi \leftrightarrow K([-\chi]\varphi \wedge A\varphi) \quad (\text{for } \varphi \neq \chi)$$

For explicit knowledge about  $\chi$  itself, we get the following.

**Fact** The formula  $[-\chi]\text{Ex } \chi \leftrightarrow K \perp$  is valid.

*Proof* Using our reduction axioms as above,

$$\begin{aligned} [-\chi]\text{Ex } \chi &\leftrightarrow [-\chi]K(\chi \wedge A\chi) \\ &\leftrightarrow K([-\chi]\chi \wedge [-\chi]A\chi) \\ &\leftrightarrow K([-\chi]\chi \wedge \perp) \\ &\leftrightarrow K \perp \end{aligned}$$

□

Under the natural requirement of consistency for knowledge (technically, seriality for the accessibility relation), this validity becomes

$$\neg[-\chi]Ex \chi$$

read as “one never has explicit knowledge about  $\chi$  after dropping it”.

Still, even after dropping it, the agent does keep  $\chi$  as implicit knowledge, witness the following valid law:

**Fact** The formula  $Ex \chi \rightarrow [-\chi]K \chi$  is valid.

*Proof*

$$\begin{aligned} Ex \chi &\rightarrow K \chi \wedge KA \chi \\ &\rightarrow K \chi \\ &\rightarrow K[-\chi]\chi \\ &\rightarrow [-\chi]K \chi \end{aligned}$$

Our proof uses the validity of

$$\chi \leftrightarrow [-\chi]\chi$$

whose justification is analogous to the one for  $\chi \leftrightarrow [+ \chi]\chi$  □

Finally, we analyze the effect of an implicit observation over explicit knowledge. For any  $\varphi$  and  $\chi$ , unfolding the definition of explicit knowledge via our axioms (we suppress intermediate steps here) gives

$$[!\chi]Ex \varphi \leftrightarrow (\chi \rightarrow K([!\chi]\varphi \wedge (\chi \rightarrow A\varphi)))$$

Again, we leave it to the reader to state this fact informally. This outcome is our solution to the earlier-mentioned problem of update making explicit knowledge ‘out of synch’ with implicit knowledge. (Recall that this was the reason for the restriction to purely factual assertions in [van Benthem \(2008\)](#), [Velázquez-Quesada \(2009\)](#)). Explicit knowledge is now a defined notion, so it automatically re-adjusts to whatever happens to the modalities  $K$  and  $A$ , and our logic tells us precisely how.

We have extracted the effect of our basic epistemic actions over explicit knowledge defined as  $K(\varphi \wedge A\varphi)$ . Thus, we replace discussion whether agents *have* epistemic closure by a much richer picture of what they can *do* to update and “upgrade” their knowledge.

Moreover, this style of analysis works not only for the stated notion of explicit knowledge; it can also provide us with validities expressing the way other possible notions of explicit knowledge are affected by dynamic actions, like the already mentioned  $K\varphi \wedge AK\varphi$ .

### 6.3 Schematic validities and algebra of actions

While all this seems straightforward dynamic epistemic technique, there is a catch. In deriving the principles of the previous section, we have used more than the reduction axioms of our logic per se. Several important ‘schematic’ principles did not follow from our reduction axioms. In particular, we have used the two principles

$$[+\chi]\chi \leftrightarrow \chi \quad \text{and} \quad [-\chi]\chi \leftrightarrow \chi$$

whose validity involved additional considerations. Of course, each specific instance of such a formula can be derived, given our completeness theorem. But that does not mean there is any illuminating *uniform derivation* of an “algebraic” sort. Indeed, an explicit characterization of the schematic validities in dynamic-epistemic logics (valid for all substitutions of formulas for proposition letters) is a well-known open problem (cf. [van Benthem 2010](#)), even in the case of public announcement logic. Given the clear interest of such general principles here, that problem becomes even more urgent.

*Algebra of actions.* We end with one particular source of schematic validities. As important as it is to understand how actions affect our information, their general algebraic structure is of interest, too. We briefly discuss some validities, to show that “*algebra of actions*” raises some interesting issues:

- In general, *drop* does not neutralize *consider*:  $[+\chi][-\chi]\varphi \leftrightarrow \varphi$  is not valid. If the agent is initially aware of  $\chi$ , *consider* makes no change, but *drop* does, yielding a model where  $\chi$  is not in the awareness set. The actual validity is the qualified

$$\neg A\chi \rightarrow ([+\chi][-\chi]\varphi \leftrightarrow \varphi)$$

- The dual case behaves in the same way: *consider* does not neutralize *drop* in general—but we do have:

$$A\chi \rightarrow ([-\chi][+\chi]\varphi \leftrightarrow \varphi)$$

As for unqualified algebraic laws, we do have idempotence:

- A sequence of *consider* actions for the same formula has the same effect as a single one, and the *drop* operation behaves similarly:

$$[+\chi]\varphi \leftrightarrow [+\chi][+\chi]\varphi \quad \text{and} \quad [-\chi]\varphi \leftrightarrow [-\chi][-\chi]\varphi$$

Next, given the dynamics of the system, we do not expect strong commutation laws between considering and dropping. Nevertheless, we do expect commutation of these operations with implicit observation, since the latter modifies an independent component of our models. For example, the following formulas,

$$[!\chi][+\chi]\varphi \leftrightarrow [+ \chi][!\chi]\varphi \quad \text{and} \quad [!\chi][-\chi]\varphi \leftrightarrow [-\chi][!\chi]\varphi,$$

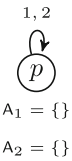
are valid even for formulas  $\chi$  using the modality  $A$ .<sup>8</sup>

We will not pursue this action algebra here (there are even more validities when we restrict attention to just factual assertions), but it clearly involves uniform schematic validities once more that are not immediately obvious from our earlier completeness theorem.<sup>9</sup>

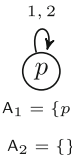
### 7 From single to multi-agent scenarios

So far, we have considered activities of single agents, including their observations, but also their acts of inference. Now the latter are typically private, and hence it makes sense to look at scenarios with privacy. But a bit paradoxically, privacy only becomes visible in a multi-agent setting.<sup>10</sup> Here is a first simple illustration, with two agents:

**Example 2** Consider the following model  $M$ , generalizing the single-agent framework to a multi-agent setting in a straightforward way:

|  |  |
|--|--|
| <div style="text-align: center;"> <p>1, 2</p>  <p><math>A_1 = \{\}</math></p> <p><math>A_2 = \{\}</math></p> </div> | <p>In the only world of the model, each agent knows implicitly that <math>p</math>, but no agent is aware of <math>p</math> (<math>\neg A_1 p \wedge \neg A_2 p</math>). Moreover, agents have implicit higher-order knowledge about each other. E.g., agent 2 knows implicitly that agent 1 is not aware of <math>p</math> (<math>K_2 \neg A_1 p</math>).</p> |
|--|--|

Now let an event take place: agent 1 considers  $p$ :  $M_{(1,+p)}$  is given by

|   |   |
|---|---|
| <div style="text-align: center;"> <p>1, 2</p>  <p><math>A_1 = \{p\}</math></p> <p><math>A_2 = \{\}</math></p> </div> | <p>In the new situation, agent 1 is aware of <math>p</math> (<math>A_1 p</math>), and now has explicit knowledge about it. But there is more: agent 2 now knows implicitly that agent 1 is aware of <math>p</math> (<math>K_2 A_1 p</math>), but without knowing this explicitly.</p> |
|---|---|

Is this a realistic scenario? Independently of the modelling, it seems strange that an internal action that takes place only in agent 1’s mind can affect the information of agent 2. To get clearer on this, we need a more detailed analysis of how epistemic awareness models should change in a setting allowing privacy.

<sup>8</sup> The reason is, once again, that  $[+\chi]$  and  $[-\chi]$  can only change the truth value of  $A\chi$ , and hence that of  $\chi$  cannot be affected.

<sup>9</sup> Here is one more analogy with public announcement logic *PAL*. The latter system, too, has an algebra of actions, viz. successive announcements—but it tends to go unnoticed, as two successive announcements can be compressed into a single one. This compression disappears when the operation changes the accessibility relation, as it is done in dynamic-epistemic logics for changes in belief or preferences: two successive *upgrades* cannot be compressed into a single one (van Benthem and Liu 2007; van Benthem 2007).

<sup>10</sup> With a single agent, the private is the public realm.

### 7.1 Multi-agent static framework

The extension of the static epistemic awareness framework to a setting with many agents in a group  $\mathbb{A}\mathcal{G}$  is straightforward. In the language of multi-agent  $\mathcal{L}$ , we just add agent indexes to the  $A$  and the  $K$  modalities ( $A_i$  and  $K_i$ , respectively). In the semantic models,  $R$  becomes a *function* from  $\mathbb{A}\mathcal{G}$  to  $\wp(W \times W)$  returning an *accessibility relation*  $R_i$  for each agent  $i \in \mathbb{A}\mathcal{G}$ , and  $\mathbf{A}$  becomes a function from  $\mathbb{A}\mathcal{G} \times W$  to  $\wp(\mathcal{L})$  returning the *awareness set*  $\mathbf{A}_i(w)$  of each agent  $i$  at each possible world  $w$ . The semantic interpretation of formulas is then as before, using  $\mathbf{A}_i$  and  $R_i$  to interpret formulas of the form  $A_i\varphi$  and  $K_i\varphi$ , respectively.

Again, in this multi-agent case we will not impose special semantic constraints. But note how, in a setting with private actions, just as in *DEL* in general, we may want to change our emphasis from *knowledge* to one on *belief*, since some actions may make agents’ information out of synch with reality.

### 7.2 Multi-agent actions: the general case

To make our actions *private*, we need a mechanism that lets actions affect agents in different ways. The *action models* of Baltag et al. (1999) will do this, provided we extend them in essentially the manner of van Benthem et al. (2006). That is, events can now really change the world, coming not just with *preconditions* on their executability, but also with *postconditions* describing what changes they bring about:

**Definition 7.1** (Multi-agent action model) With  $\mathbb{P}$  the set of atomic propositions and  $\mathbb{A}\mathcal{G}$  the finite set of agents, a *multi-agent action model* is a tuple  $E = \langle S, T, \text{Pre}, \text{Pos} \rangle$  where:

- $\langle S, T, \text{Pre} \rangle$  is an action model (Baltag et al. 1999) with  $S$  a finite non-empty set of *events*,  $T : \mathbb{A}\mathcal{G} \rightarrow \wp(W \times W)$  a function returning an *accessibility relation*  $T_i$  for each agent  $i \in \mathbb{A}\mathcal{G}$  and  $\text{Pre} : S \rightarrow \mathcal{L}$  the *precondition* function indicating where each event can be executed.
- $\text{Pos} : (\mathbb{A}\mathcal{G} \times S \times \wp(\mathcal{L})) \rightarrow \wp(\mathcal{L})$  is the *postcondition* function, assigning a new set of formulas in  $\mathcal{L}$  to every tuple of an agent, event, and (old) set of formulas in  $\mathcal{L}$ .

A *pointed action model*  $(E, s)$  has a distinguished event  $s$ .

Recall that we want to model private versions of our operations that *modify* awareness sets. This is exactly the role of the  $\text{Pos}$  function, a generalization of the *substitution function* in van Benthem et al. (2006) for representing factual change. The following update rule describes how our action models modify epistemic awareness models.

**Definition 7.2** (Product update) Let  $M = \langle W, R, \mathbf{A}, V \rangle$  be a multi-agent semantic model and let  $E = \langle S, T, \text{Pre}, \text{Pos} \rangle$  be a multi-agent action model. The *product model*  $M \otimes E = \langle W', R', \mathbf{A}', V' \rangle$  is given by

- $W' := \{ (w, s) \mid (M, w) \models \text{Pre}(s) \}$
- $V'(w, s) := V(w)$
- $R'_i(w, s)(w', s')$  iff  $R_i w w' \& T_i s s'$
- $\mathbf{A}'_i(w, s) := \text{Pos}_i(s, \mathbf{A}_i(w))$

Note how the Pos function works: for each agent  $i$  and each event  $s$ , Pos takes agent  $i$ 's awareness set at  $w$  in  $M$ , and returns her awareness set at  $(w, s)$  in  $M \otimes E$ . Later on, we will look at restrictions on the syntactic format of definition for the postcondition function.<sup>11</sup>

In order to express how product updates affect the agents' information, the *extended* multi-agent language  $\mathcal{L}$  has extra modalities:

if  $(E, s)$  is a pointed action model and  $\varphi$  is a formula in the extended multi-agent  $\mathcal{L}$ , then so is  $[(E, s)]\varphi$ .

The semantic interpretation of these new formulas is as follows:

**Definition 7.3** Let  $(M, w)$  be a pointed multi-agent semantic model and let  $(E, s)$  be a pointed action model with  $E = \langle S, T, \text{Pre}, \text{Pos} \rangle$ .

$$(M, w) \models [(E, s)]\varphi \text{ iff } (M, w) \models \text{Pre}(s) \text{ implies } (M \otimes E, (w, s)) \models \varphi$$

Now it is time to look at concrete cases illustrating the mechanism.

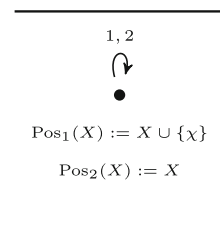
### 7.3 Public *considering* and *dropping*

For a start, our multi-agent setting generalizes the single agent case, since we can define action models for our earlier (now public) *consider* and *drop* operations.

**Definition 7.4** (Public *consider* action) Let  $\chi$  be a formula in multi-agent  $\mathcal{L}$ . The action of *agent  $j$  publicly considering  $\chi$*  is the pointed action model  $(\text{Pub}_{+\chi}^j, \bullet)$  with  $\text{Pub}_{+\chi}^j = \langle S, T, \text{Pre}, \text{Pos} \rangle$  given by

- $S := \{\bullet\}$
- $T_i := \{(\bullet, \bullet)\}$  for every agent  $i$
- $\text{Pre}(\bullet) := \top$
- $\text{Pos}_j(\bullet, X) := X \cup \{\chi\}$ ,
- $\text{Pos}_i(\bullet, X) := X$  for  $i \neq j$

The diagram on the right shows the action model  $\text{Pub}_{+\chi}^1$  in the 2-agent case (with the preconditions omitted).



<sup>11</sup> Substitution functions are highly constrained and simply definable, our Pos function is much less constrained, though in cases of interest, it, too, usually has a simple syntactic definition.



**Definition 7.5** (Public drop action) Let  $\chi$  be any formula in multi-agent  $\mathcal{L}$ . The action of agent  $j$  publicly dropping  $\chi$ , given by the pointed action model  $(\text{Pub}_{-\chi}^j, \bullet)$ , differs from public considering only in its postcondition function for agent  $j$  in  $\bullet$ :

$$\text{Pos}_j(\bullet, X) := X \setminus \{\chi\}$$

The public versions of the actions have just one event, and their accessibility relations  $T_i$  indicate that all involved agents recognize this. Moreover, the precondition in the lonely world is simply  $\top$ . Then, the application of  $(\text{Pub}_{+\chi}^j, \bullet)$  ( $(\text{Pub}_{-\chi}^j, \bullet)$ , resp.) on a multi-agent static model  $M$  yields a copy of  $M$  in which  $\chi$  has been added to (respectively, removed from) the awareness set of agent  $j$  in all worlds.

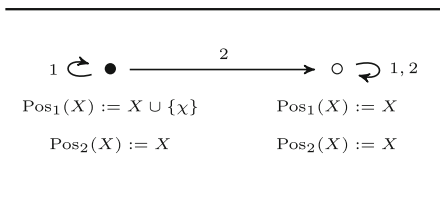
7.4 Private considering and dropping

But our mechanism can also define private actions. Here are simple versions of the earlier consider and drop. As usual, these encode what takes place, but also how different agents ‘view’ this:

**Definition 7.6** (Private consider action) Let  $\chi$  be any formula in multi-agent  $\mathcal{L}$ . The action of agent  $j$  privately considering  $\chi$  yields the pointed action model  $(\text{Pri}_{+\chi}^j, \bullet)$  with  $\text{Pri}_{+\chi}^j = \langle S, T, \text{Pre}, \text{Pos} \rangle$  as

$$\begin{aligned} -S &:= \{\bullet, \circ\} & -T_i &:= \begin{cases} \{(\bullet, \bullet), (\circ, \circ)\} & \text{if } i = j \\ \{(\bullet, \circ), (\circ, \circ)\} & \text{otherwise} \end{cases} & -\text{Pre}(\bullet) &= \text{Pre}(\circ) := \top \\ & & -\text{Pos}_j(\bullet, X) &:= X \cup \{\chi\}, & \text{Pos}_j(\circ, X) &:= X \\ & & -\text{Pos}_i(\bullet, X) &:= X, & \text{Pos}_i(\circ, X) &:= X \text{ for } i \neq j \end{aligned}$$

The diagram on the right shows the model  $\text{Pri}_{+\chi}^1$  for 2 agents (preconditions again omitted).



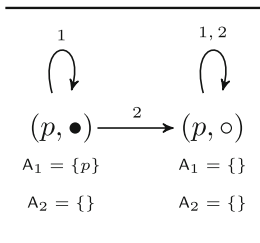
**Definition 7.7** (Private drop action) Let  $\chi$  be any formula in multi-agent  $\mathcal{L}$ . The action of agent  $j$  privately dropping  $\chi$ , given by the pointed action model  $(\text{Pri}_{-\chi}^j, \bullet)$ , differs from private considering only in its postcondition function for agent  $j$  in  $\bullet$ :

$$\text{Pos}_j(\bullet, X) := X \setminus \{\chi\}$$

The difference between the public and the private version of the actions is that the private actions involve two events: one in which  $\chi$  is added to (removed from) agent

$j$ 's awareness set (the event  $\bullet$ ), and another in which there is no change (the event  $\circ$ ). Moreover, the accessibility relations  $T_i$  indicate that, while  $j$  recognizes which event is the real one (our  $\bullet$ ), the other agents do not consider that event possible, sticking to the 'no change' option. Then, the application of  $(\text{Pri}_{+\chi}^j, \bullet)$  ( $(\text{Pri}_{-\chi}^j, \bullet)$ , resp.) on a multi-agent static model  $M$  yields a model containing two copies of  $M$ : one, recognized as the real one only by  $j$ , in which  $j$ 's awareness set changed, and another, viewed by the other agents as the only possibility, in which nothing happened.

**Example 3** Recall the model  $M$  from Example 2. After agent 1 considers  $p$  privately (i.e., after applying  $(\text{Pri}_{+p}^1, \bullet)$ ), we get a better version of our initial example that started the thread of this section:



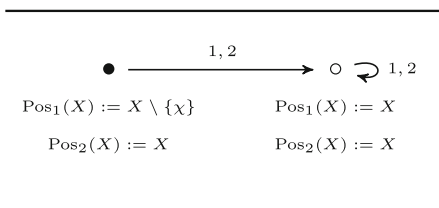
In the evaluation point, here the leftmost world, agent 1 is aware of  $p$  ( $A_1 p$ ), just like she does after publicly considering  $p$ . But this time, agent 2's implicit belief does not change: she continues to believe implicitly that agent 1 is not aware of  $p$  ( $K_2 \neg A_1 p$ ).

### 7.5 Unconscious versions

The flexibility of the postcondition mechanism is great. We can represent many further scenarios, even *unconscious* actions, hidden from all agents, including the one that 'performs' it! We just give an illustration:

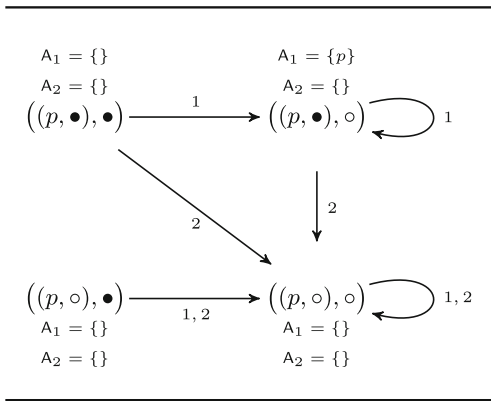
**Definition 7.8** (Unconscious *drop* action) Let  $\chi$  be a formula in the multi-agent  $\mathcal{L}$ . The action of *agent  $j$  unconsciously dropping  $\chi$*  yields the pointed action model  $(\text{Unc}_{-\chi}^j, \bullet)$ , differs from its *private* counterpart only in the definition of the accessibility relation:

$$T_i := \{(\bullet, \circ), (\circ, \circ)\} \quad \text{for every agent } i$$



The diagram on the left depicts  $\text{Unc}_{-\chi}^1$  in a 2-agent scenario.

**Example 4** Consider the model  $(M \otimes \text{Pri}_{+p}^1, (p, \bullet))$  of Example 3. If agent 1 unconsciously drops  $p$ , we get the following updated model:



In  $((p, \bullet), \bullet)$ , the agent is not aware of  $p$  ( $\neg A_1 p$ ), but she implicitly believes that she is aware of it ( $K_1 A_1 p$ ).

The updated model contains two copies of the original one. Here, agent 1 considers only the rightmost world of the upper copy, and agent 2 only the rightmost world of the lower one.

Much more can be said about this scenario, and we feel that we have a promising take here on unconscious actions such as *forgetting*. But our purpose here was just to demonstrate the flexibility of the framework.

### 7.6 Completeness of the multi-agent system

In principle, there is a complete dynamic logic for our product-update mechanism for awareness, and it looks like our earlier single-agent logic, with indices attached. Its principles for atomic formulas, Boolean operations, and implicit knowledge are the usual ones from the dynamic-epistemic logic *DEL*. As an illustration, we have the valid equivalence

$$\vdash [(E, s)]K_i \varphi \leftrightarrow \text{Pre}(s) \rightarrow \bigwedge_{Tsr} K_i [(E, r)]\varphi$$

But to formulate a precise result, the crucial issue is stating the right reduction axiom for awareness given the *postconditions*. Consider the earlier axioms that we gave for our two basic syntactic operations of *consider* and *drop*. These described the postconditions (the effect of the operations on the  $\mathbf{A}$ -sets) inside the language, exploiting the simple format of their effects. For instance, we have  $\varphi$  in our  $\mathbf{A}$ -set after an act  $+\chi$  if we had  $\varphi$  before, or  $\varphi$  is actually the just added formula  $\chi$ . This case distinction in the reduction axiom reflects directly the simple disjunctive *definition of the postcondition* for the action  $+\chi$ :  $\mathbf{A}'(w) := \mathbf{A}(w) \cup \{\chi\}$ .

The same is true in our more general setting with action models: *simple uniform definitions of postconditions in our action models will allow us to provide matching reduction axioms*. We will not pursue the technicalities of this mechanism here, but just say this. To get a proper completeness theorem, there should be a simple definition scheme  $\delta(\phi)$  within our language for postconditions of awareness-set changing

events, stating when a formula  $\phi$  belongs to the new set. The syntax of this scheme can then be used to derive reduction axioms.<sup>12</sup>

We have merely made some proposals, and explored a simple syntactic version of *DEL* product update. But even at this preliminary stage, our private awareness analysis seems an interesting extensions of the scope of standard dynamic epistemic logic.

## 8 Other approaches

Many approaches deal with the problem of logical omniscience. Here we mention some of them, and their relation to our work.

*Static approaches.* A crucial idea in the logical omniscient discussion is to differentiate between what the agent actually knows and what follows from this explicit knowledge. Then, explicit knowledge is weakened somehow, obtaining in this way an agent that does not know *explicitly* everything that follows from her explicit knowledge. In (Konolige 1984), the author studied agents with arbitrary rules of inference; Levesque (Levesque 1984) introduced *partial* and *incoherent* situations. Our approach has followed the already mentioned influential Fagin and Halpern (1988) in using a syntactic notion of awareness (an idea that, as indicated by an anonymous referee, goes back to syntactic notions of knowledge (Eberle 1974; Moore and Hendrix 1979)).

All these approaches deal with agents with limited information, but what seems new in our approach is the way in which we study explicit actions of awareness change.

*Temporal approaches.* When dealing with dynamic systems that evolve, another approach is to look at them from a temporal perspective. Indeed, Fagin and Halpern (1988), Elgot-Drapkin and Perlis (1990) and Duc (1997) have explored how the information (knowledge, or beliefs) of an agent change over time. Nevertheless, what we add here is an explicit constructive account of the local dynamics that creates these temporal histories.

*Other dynamic approaches.* In the work of Jago (2006) and later Jago (2009) application of inference rules are indicated explicitly in the logical language. These are then studied in an abstract *dynamic logic* style, while we provide *DEL*-style constructive model transformations.

*Earlier DEL approaches.* Unlike recent approaches to inferential dynamics (van Benthem 2008; Velázquez-Quesada 2009), our agents' explicit knowledge is not only about factual formulas. It also includes 'higher' information about their own information. While this may seem a small 'in-house' difference, it is a big step forward in the *DEL*-style treatment of awareness dynamics. In this connection, Grossi and Velázquez-Quesada (2009) also deals with higher-order information. The present

<sup>12</sup> The system for epistemic and factual change in (van Benthem et al. 2006) is an illustration, exploiting the fact that postconditions of its events were definable as simple syntactic substitutions.

proposal extends this to a real multi-agent treatment, allowing us to handle public, private and even unconscious versions of the *consider* and *drop* actions.

## 9 Conclusions and further directions

Our paper shows how a significant informational dynamics can take place over existing epistemic awareness models, generalizing acts of observation and inference. We also show how this leads to perspicuous useful technical systems and results about these, in the spirit of dynamic epistemic logic. Thus, we have shown that the ‘reductionist approach’ to explicit knowledge in terms of implicit semantic knowledge and syntactic awareness is feasible and interesting in its own right.

Of course, we also found lots of open problems. Our multi-agent setting can describe many more agent activities than what we have shown, but we have only scratched the surface. Also, many technical issues remain unresolved: in particular, the issue of schematic validities and action algebra, and the precise logic of postconditions, dependent on their format of definition.

But beyond this, we also see drawbacks to the approach taken here. In particular, we still feel that the reductionist approach, which solves the long-standing problem of dealing with complex formulas, also has its drawbacks. In an intuitive sense, explicit knowledge may be *sui generis*, and not reducible to implicit knowledge plus awareness in any of the existing senses in the literature. I might be thinking of  $\varphi$ , and also know it implicitly, and still fail to see directly that it is true. Think of a conclusion that I am pondering, and that in fact follows from some premises whose truth I am aware of. I could still fail to see how it follows explicitly. In fact, this shows that we have not analyzed real ‘acts of inference’, since for us, merely becoming aware of  $\varphi$  was enough to upgrade knowledge from implicit to explicit.<sup>13</sup> Thus, eventually we want a much richer account of inference, allowing us to deal with important issues like *preferences* over possible inferences and an explicit treatment of their *justifications* (cf. Artemov and Nogina (2005) and related papers on ‘justification logics’).

A most urgent desideratum in our view concerns the interpretation shift in our multi-agent section. It was *beliefs* of agents that made more sense there than *knowledge*. But once we take beliefs seriously, we should redo our analysis in the setting of dynamic logics for acts of *belief revision* (cf. Baltag and Smets 2008; van Benthem 2007) that work over epistemic plausibility models. Indeed, it makes much sense to relate belief revision, not just to new observations, but also to new inferences. This inferential dynamics should then also include special mechanisms that affect belief, rather than knowledge, that is: *default rules*. It is a significant issue how all this should be done.<sup>14</sup> We refer to the dissertation Velázquez-Quesada (2011) for details and further results on awareness dynamics, default reasoning, and belief revision.

<sup>13</sup> Grossi and Velázquez-Quesada (2009) explore an explicit knowledge definition that involves a further requirement: the formula should be acknowledged as true via a rule-based inference.

<sup>14</sup> van Benthem (2009) is a first exploration, including some major changes in the plausibility relation, that now does not just order worlds, but worlds plus partial syntactic descriptions.

But our main point is not technical results, or concrete directions. It is rather the general picture of agency arising from our analysis. We replaced agents with “supernatural” abilities like omniscience by human ones that must, and can work to improve their information. The resulting mathematical model is rich, and much more attractive than the usual ones. We mentioned Sherlock Holmes at some point in our story, famous for combining observation and deduction. Dynamic logics are about what makes this tick. As visitors to Reichenbach Falls in Switzerland can see, our hero died a (fictional) death at the hands of the evil mathematician Professor Moriarty. But though these two minds were enemies, their fields were not. We hope to have shown how delicate philosophical issues can profit from mathematical modeling.

**Acknowledgements** The authors would like to thank three anonymous referees. Their comments and suggestions have greatly improved this paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Artemov, S. N., & Nogina, E. (2005). Introducing justification to epistemic logic. *Journal of Logic and Computation*, 15(6), 1059–1073.
- Baltag, A., Moss, L. S., & Solecki, S. (1999). *The logic of public announcements, common knowledge and private suspicious*. Technical Report SEN-R9922. Amsterdam: CWI.
- Baltag, A., & Smets, S. (2008). A qualitative theory of dynamic interactive belief revision. In G. Bonanno, W. van der Hoek, & Wooldridge, M. (Eds.), *Logic and the foundations of game and decision theory (LOFT7), texts in logic and games* (Vol. 3, pp. 13–60). Amsterdam: University Press.
- Duc, H. N. (1997). Reasoning about rational, but not logically omniscient, agents. *Journal of Logic and Computation*, 7(5), 633–648.
- Eberle, R. A. (1997). A logic of believing, knowing, and inferring. *Synthese*, 26(3–4), 356–382.
- Elgot-Drapkin, J. J., & Perlis, D. (1990). Reasoning situation in time I: Basic concepts. *Journal of Experimental and Theoretical Artificial Intelligence*, 2(1), 75–98.
- Fagin, R., & Halpern, J. Y. (1988). Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34(1), 39–76.
- Gerbrandy, J. (1999). *Bisimulations on planet Kripke*. Ph.D. thesis, Institute for Logic, Language and Computation (ILLC), Universiteit van Amsterdam (UvA), Amsterdam, The Netherlands. ILLC Dissertation Series DS-1999-01.
- Grossi, D., & Velázquez-Quesada, F. R. (2009). Twelve angry men: A study on the fine-grain of announcements. In X. He, J. F. Horty, & Pacuit, E. (Eds.), *LORI, lecture notes in computer science* (Vol. 5834, pp. 147–160). New York: Springer.
- Halpern, J. Y. (2001). Alternative semantics for unawareness. *Games and Economic Behavior*, 37, 321–339.
- Hintikka, J. (1962). *Knowledge and belief: An introduction to the logic of the two notions*. Ithaca, NY: Cornell University Press.
- Jago, M. (2006). Rule-based and resource-bounded: A new look at epistemic logic. In T. Ågotnes, & Alechina, N. (Eds.), *Proceedings of the workshop on logics for resource-bounded agents, organized as part of the 18th European summer school on logic, language and information (ESSLLI)* (pp. 63–77). Malaga, Spain.
- Jago, M. (2009). Epistemic logic for rule-based agents. *Journal of Logic, Language and Information*, 18(1), 131–158.
- Konolige, K. (1984). *Belief and incompleteness*. Technical Report 319. Menlo Park, California: SRI International.

- Lakemeyer, G. (1986). Steps Towards a first-order logic of explicit and implicit belief. In Halpern, J. Y. (Ed.), *Proceedings of 1986 conference on theoretical aspects of reasoning about knowledge* (pp. 325–340). San Francisco, CA: Morgan Kaufmann Publishers Inc.
- Levesque, H. J. (1984). A logic of implicit and explicit belief. In *Proceedings of AAAI-84* (pp. 198–202). Austin, TX
- Moore, R. C., & Hendrix, G. (1979). *Computational models of beliefs and the semantics of belief sentences*. Technical Report 187. Menlo Park, California: SRI International.
- Moreno, A. (1998). Avoiding logical omniscience and perfect reasoning: A survey. *AI Communications*, 11(2), 101–122.
- Plaza, J. A. (1989). Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, & Ras, Z. W. (Eds.), *Proceedings of the 4th international symposium on methodologies for intelligent systems* (pp. 201–216). Tennessee, USA: Oak Ridge National Laboratory, ORNL/DSRD-24.
- Sim, K. M. (1997). Epistemic logic and logical omniscience. A survey. *International Journal of Intelligent Systems*, 12(1), 57–81.
- van Benthem, J. (2007). Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17(2), 129–155.
- van Benthem, J. (2008). Merging observation and access in dynamic logic. *Journal of Logic Studies*, 1(1), 1–17.
- van Benthem, J. (2009). Logic, mathematics, and general agency. In P. Bour, M. Rebuschi, & Rollet, L. (Eds.), *Festschrift for Gerhard Heinzmann*. Nancy: Laboratoire d'histoire des sciences et de la philosophie.
- van Benthem, J. (2010). *Logical dynamics of information and interaction*. Cambridge: Cambridge University Press (to appear).
- van Benthem, J., & Liu, F. (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logics*, 17(2), 157–182.
- van Benthem, J., & Martínez, M. (2008). The stories of logic and information. In P. Adriaans & van Benthem J. (Eds.), *Philosophy Of information, handbook of the philosophy of science* (pp. 217–280). Amsterdam: North-Holland.
- van Benthem, J., van Eijck, J., & Kooi, B. (2006). Logics of communication and change. *Information and Computation*, 204(11), 1620–1662.
- van Ditmarsch, H., & French, T. (2009). Awareness and forgetting of facts and agents. In *Proceedings of the 2009 IEEE/WIC/ACM international joint conference on web intelligence and intelligent agent technologies (WI-IAT 2009)*. Milan.
- van Ditmarsch, H., Herzig, A., Lang, J., & Marquis, P. (2009). Introspective forgetting. *Synthese (Knowledge, Rationality and Action)*, 169(2), 405–423.
- Vardi, M. Y. (1986). On epistemic logic and logical omniscience. In Halpern, J. (Ed.), *Proceedings of the 1986 conference* (pp. 293–305). Los Altos, CA: Morgan Kaufmann Publishers Inc.
- Velázquez-Quesada, F. R. (2009). Inference and update. *Synthese (Knowledge, Rationality and Action)*, 169(2), 283–300.
- Velázquez-Quesada, F. R. (2011). *Small steps in dynamics of information*. Ph.D. thesis, Institute for Logic, Language and Computation (ILLC), Universiteit van Amsterdam (UvA), Amsterdam, The Netherlands (to appear).