Semantics for knowledge and change of awareness

Hans van Ditmarsch and Tim French

Abstract We examine various logics that combine knowledge, awareness, and change of awareness. An agent can become aware of propositional propositions but also of other agents or of herself. The dual operation to becoming aware, forgetting, can also be modelled. Our proposals are based on a novel notion of structural similarity that we call *awareness bisimulation*, the obvious notion of modal similarity for structures encoding knowledge and awareness.

1 Introduction

Modal logic has long been used to reason about knowledge and belief in multi-agent systems. While modal logics allow us to model uncertainty by varying the value of propositions between states (the so-called possible worlds), they still require that the agents are aware of all propositions and agents in the model. Thus reasoning in these models is undertaken under the closed world assumption. To every proposition in every state, every agent assigns some kind of value (true, false, unknown, or a value on some scale). In this paper we consider semantic structures and logics that differentiate between agents being uncertain (of the valuation) of a proposition, and agents being unaware (of the relevance) of a proposition.

Agents may be uncertain about the valuations of propositions, they may be unaware of these propositions, but they may also *become aware* of propositions, and be aware or become aware of other agents. We find that there are many subtleties and intricacies involved in defining the semantics for such dynamics. In this paper we will discuss these intricacies and in doing so make the following contributions:

Hans van Ditmarsch
University of Sevilla, Spain, e-mail: hvd@us.es

Tim French
University of Western Australia, Perth, Australia e-mail: tim@csse.uwa.edu.au

1. We will introduce a new form of model equivalence modulo the agents' awareness and uncertainty, called *awareness bisimulation*.
2. We will define a new type of knowledge, tentatively referred to as *speculative knowledge*. Speculative knowledge is essential to express the dynamic interactions between awareness and knowledge.
3. As well as considering awareness of atomic propositions, we also consider *awareness of agents*. This allows us to model such concepts as self-awareness and speculation about an unknown adversary.
4. We will present operators for awareness change—both *becoming aware* and *forgetting*—and give semantics for these operators that are consistent with our intuitions of awareness and knowledge.

Our work is rooted in: the tradition of epistemic logic [18] and in particular multi-agent epistemic logic [21, 11]; in various research since the 1980s on the interaction of awareness and knowledge [10, 22, 23, 16] — including a relation to recent works like [17, 14, 15]; and in modal logical research in propositional quantification, starting in the 1970s with [12] and followed up by work on bisimulation quantifiers [27, 19, 13].

Works treating awareness either follow a more *semantically* flavoured approach, where awareness concerns propositional variables in the valuation [10, 23, 16], or a more *syntactically* flavoured approach. In the latter, awareness concerns all formulas of the language in a given set, in order to model 'limited rationality' of agents. It is (also) pursued in [10] and in recent work like [14, 26]. We are straight into the semantic corner: within the limits of their awareness, agents are fully rational. For the static part of our logic we follow the semantics of [10].

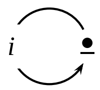
Our proposal extends the work in [7, 6, 8]; [7] presents the dynamics of awareness as a side issue, and shows that it combines well with dynamics of knowledge, as in a truthful informative announcements where a novel issue is addressed; [6] focusses on a special case of awareness, namely *public global awareness* when all agents are aware of the same propositions in all states; [8] gives an axiomatization and decidability for the logic where an agent may only become aware of a propositions in all states, another special case. The notion of awareness bisimulation is already presented in [8], but in a version for propositional variables only. In our current proposal we present the widest range of modelling alternatives.

2 Motivating example

Hans is going to a conference. He wakes up in the morning. His mind is empty. He is utterly unaware of his surroundings and even of himself. This situation is best characterized as:

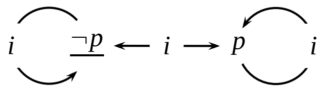
In words: a singleton domain Kripke model for an empty set of agents and an empty set of propositional variables, and therefore with an empty accessibility relation, and no valuation of atoms. Still, some formulas are true in this state: $\top$, $\neg\bot$, $\top \vee \neg\bot$, and so on.

A sprinkling of self-awareness starts to invade Hans' mind. He is able to reflect on his judgements again, yet another day! At this stage, the only thing he is able to avoid is contradictions, but there are no issues to reason about yet. This situation can be visualized as:



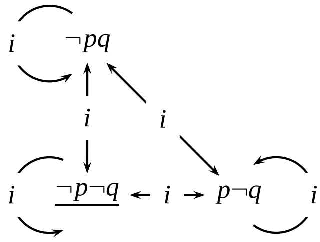
Hans is the agent i, who has full access to the singleton domain. Already, there is lots more to reason about. The following formulas are true in this state: $K_i\top$, $K_i\neg K_i\bot$, $\top \vee K_iK_i\neg\bot$, and so on. Hans knows that he can avoid contradictions, and that he does not know falsehoods, etc. With self-awareness, higher-order cognition has arrived!

Being aware of himself, and still rubbing stardust from his sleepy eyes, Hans realizes what is lacking: coffee. He starts to wonder if coffee would already be served in the restaurant below. It is still fairly early. Hans has reached the stage that he can reason about something, namely an actual proposition p (coffee is served) that he is uncertain about. This situation can be visualized as:



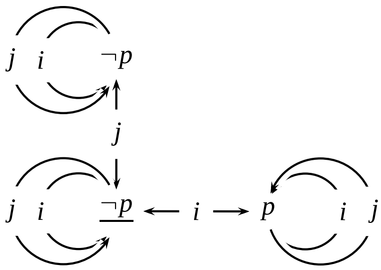
The things that are now true are beginning to look more and more like modal logic: $K_i(\neg K_i p \wedge \neg K_i \neg p)$ (Hans knows that he does not know whether coffee is served), $\neg K_i p$ (Hans does not know that coffee is served), etc. In fact, no coffee is served. The actual state in the figure, where p is false, is therefore underlined.

Now, while on his way to the lower floor, where the restaurant is located, someone in the elevator mentions that you can't have both coffee and orange juice for breakfast. This makes Hans aware that orange juice is an issue. After this, Hans does not know whether coffee is served and also does not know whether orange juice is served. But he knows that coffee and orange juice are not both served. We now get to the following situation. Unfortunately, Hans has still not found out that the breakfast area is closed. So actually (underlined), there is no coffee and no orange juice.



It has now become true that $K_i\neg(p \wedge q)$, and so on.

Starting with mere uncertainty about coffee, and forgetting the orange complications right above: there is yet another possible scenario, one wherein Tim also enters the picture. Hans recalls (i.e., he *becomes aware*) that Tim arrived a day earlier at the conference. Tim would certainly know if the restaurant is open or closed. And Hans assumes he and Tim would have been informed if the restaurant had been open, and that Tim may have thought Hans had already found out in another way if it had been closed. The resulting picture then becomes:



From now on, the full range of multi-agent epistemic logic is at our (and Hans', and Tim's) disposition. In the underlined state it is true that Tim does not know whether Hans knows p: $\neg K_j(K_i p \vee K_i\neg p)$, that Hans does not consider it possible that Tim knows that p: $\neg\neg K_i\neg K_j p$, and so on.

Going in the opposite direction, we can model forgetting, or rather *making an abstraction* from a given information state. Let us not bother the reader with more examples but get down to logical business. Consider the transition from the situation wherein i (only) is uncertain about p to the situation where i is also uncertain about q but knows that p and q (coffee and orange juice are served) cannot both be true. Figure 1 shows three distinct ways in which we can view this transition. Let us call the models *Le*, *Mi*, *Ri*, and *Bo* for the one on the left, middle (top), right, and bottom, respectively.

The transition from *Le* to *Bo* is how we would really like to view it: no mention of q whatsoever in the initial uncertainty. From a structural point of view, that may still seem more or less fine, but from a logical point of view, there certainly is a

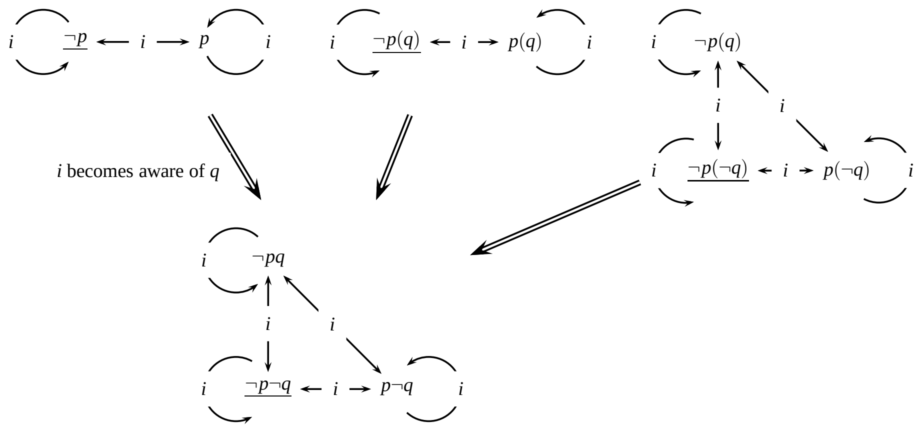


Fig. 1 Three different ways to view agent *i* becoming aware of proposition *q*. The depicted models are named *Le*, *Mi*, *Ri*, and *Bo* for the one on the left, middle (top), right, and bottom, respectively.

problem. Parameter sets of propositional variables and agents are typically given at the outset, and a logical language, and a class of structures to interpret the language in, are then given using such parameters. So, variable q should have some initial presence. Epistemic models are defined with a valuation of atomic propositions over states — we therefore have to choose a valuation for q in both states of the model on the left. The middle and right models *Mi* and *Ri* are two incompatible solutions for this. In *Mi* we have given q the value $\top$ in both states, but we wish to treat that as a don't care value. It is put in parentheses to indicate that i is unaware of p. In *Ri* we copied the resulting structure *Bo*, but with the difference that i is still unaware of q. Again, it is put in parentheses in all three states.

The transition from *Ri* to *Bo* can be seen as the agent becoming aware of structure that is already present in the model. It is common to see this as implicit knowledge becoming explicit, e.g., implicit knowledge that p and q are not both true.

The transition from *Mi* to *Bo* can be seen as a minimal way to expand the model *Mi* such that the agent to become aware of q in a specific way, while retaining all her explicit knowledge in *Mi*, namely ignorance of p. It is then essential to point out that there are many other ways in which agent i can become aware of q, e.g.:

- She can become aware of q such that she considers all four valuations of p and q possible;
- She can become aware of q such that she knows q (achieved by removing the parentheses from (q) in both states in *Mi*);
- She can become aware of q such that she knows $\neg q$;
- And so on...

Agents speculating about all the different ways in which one could become aware of a fresh variable is exactly what our proposed notion of *speculative knowledge* allows for, see Section 6. This view of becoming aware as building a model also allows for specific instructions on how the agent becomes aware of something, to select one of various possible ways to become aware; a matter that we will address in Section 9.

What the four structures in Figure 1 have in common is that if you abstract from q, they are are *the same*, in the technical sense that they are all p-bisimilar. (And also, they are i-bisimilar.) Once you take away the value of q, the left two states in *Ri* and *Bo* are indistinguishable: the same fact p is true there, and from either we can access a p-state as well as a $\neg p$-state and nothing else. The notion of restricted bisimilarity is the foundation of our approach to study the interactions of awareness and knowledge. We propose a notion of *awareness bisimilarity* in Section 4 and our main contribution is on the various logical uses of this semantic primitive.

3 Language

Given are a countably infinite set of propositional variables (propositions) P and a countably infinite set of agents N. The sets P and N are disjoint. Propositional variables are named p,q,r, possibly indexed or quoted, and agent variables are named i,j,k, possibly indexed or quoted. For all sets X and Y and all x, write $X+x$ for $X \cup \{x\}$, write $X-x$ for $X \setminus \{x\}$, write $\overline{Y}$ for $X \setminus Y$, and similarly $\overline{x}$ for $X-x$.

We augment multi-agent epistemic logic several new operators: $A_i\varphi$, to mean that agent i *is* aware of al the propositions and agents in φ; $A_i^+ p\varphi$ for agent i *becoming aware* of proposition (variable) p, after which φ is true (and similarly, for becoming aware of agents); and $A_i^- p\varphi$ for agent i *forgetting* proposition p, after which φ is true (and similarly, $A_i^- j\varphi$ stand for agent i forgetting agent j, after which φ is true). The construct $K_i^I\varphi$ stands for "agent i implicitly knows φ". We also have a primitive $\forall p\varphi$ in the language. This is a technical device that stands for (awareness) bisimulation quantification over p, after which φ is true (and a similar quantification operator $\forall i\varphi$ over agent variables). Its meaning is very much linked to the semantic primitives that we still have to introduce, and it will later be combined with other primitive operators to define more familiar notions.

Definition 1 (Language). Given are a countably infinite set of propositional variables (propositions) P, and a (disjoint) countably infinite set of agents N. The language $\mathscr{L}$ of knowledge and awareness is defined as

$$\varphi ::= \top \mid p \mid \varphi \wedge \varphi \mid \neg\varphi \mid K_i^I\varphi \mid \forall p\varphi \mid \forall i\varphi \mid A_i\varphi \mid A_i^+ p\varphi \mid A_i^+ i\varphi \mid A_i^- p\varphi \mid A_i^- i\varphi$$

where $i \in N$ and $p \in P$. Implication $\rightarrow$, disjunction $\vee$, and equivalence $\leftrightarrow$ are defined by abbreviation. For $\neg K_i^I\neg\varphi$ we write $L_i^I\varphi$, and for $\neg\forall p\neg\varphi$ we write $\exists p\varphi$. *Explicit knowledge* is defined as $K_i^E\varphi := K_i^I\varphi \wedge A_i\varphi$. Given a finite set $X = x_1,\ldots,x_n$ of variables (propositional or agent variables, or both), we write $\forall X \varphi$ for $\forall x_1 \ldots \forall x_n \varphi$, and where $\forall \emptyset \varphi = \varphi$.

We explicitly include the formula $\top$ in the language, as the usual abbreviation $p \vee \neg p$ complicates cases where not all agents are aware of p: an agent unaware of p (as in the introductory example, at the moment of creation) would then not explicitly know truth. Unlike usual in multi-agent epistemic logic, we assume an infinite set of agents. This is obligatory, because in any given system the agents can always become aware of yet another agent. No finite number, however large, will therefore do. The un(upper)labeled K_i operator in the introductory section denotes explicit knowledge K_i^E.

The semantics of the awareness operator A_i will be purely syntax-based, namely using the *free variables* of a formula. These are defined as follows. Note that $v(\varphi) \subseteq P \cup N$.

Definition 2 (Free variables). The free variables of a formula are defined by induction on formula structure: $v(\top) = \emptyset$, $v(p) = \{p\}$, $v(\varphi \wedge \psi) = v(\varphi) \cup v(\psi)$, $v(\neg \varphi) = v(\varphi)$, $v(K_i \varphi) = v(\varphi) + i$, $v(\forall p \varphi) = v(\varphi) - p$, $v(\forall i \varphi) = v(\varphi) - i$, $v(A_i \varphi) = v(\varphi) + i$, $v(A_i^+ p \varphi) = v(\varphi) + i + p$, $v(A_i^+ j \varphi) = v(\varphi) + i + j$, $v(A_i^- p \varphi) = v(\varphi) + i + p$ and $v(A_i^- j \varphi) = v(\varphi) + i + j$.

Example 1. Consider Hans again, who is uncertain about coffee and is becoming aware of his uncertainty about orange juice. Hans is unaware of oranges: $\neg A_i q$; Hans does not (explicitly) know whether there is coffee: $\neg(K_i^E p \vee K_i^E \neg p)$; after becoming aware of oranges, Hans explicitly knows that coffee and oranges are not both served: $A_i^+ q K_i^E \neg(p \wedge q)$; and this entails that he is aware of oranges: $A_i^+ q A_i q$; Hans considers it (explicitly!) possible that after becoming aware of oranges, he knows that orange juice and coffee are both available: $L_i^E \exists q A_j^+ K_i^E (p \wedge q)$. Note that $v(L_i^E \exists q A_j^+ K_i^E (p \wedge q)) = \{p, i\}$, but that $v(A_j^+ K_i^E (p \wedge q)) = \{p, q, i\}$.

4 Structures

We augment standard epistemic (Kripke) models with a parameter for awareness, and subsequently introduce a proper notion of bisimilarity for these structures.

Definition 3 (Epistemic awareness model). An *epistemic awareness model* for N and P is a tuple $M = (S, R, \mathscr{A}, V)$ that consists of a *domain* S of (propositionual) *states* (or 'worlds'), an *accessibility function* $R : N \to \mathscr{P}(S \times S)$, an *awareness function* $\mathscr{A} : N \to S \to \mathscr{P}(P \cup N)$ and a *valuation function* $V : P \to \mathscr{P}(S)$. For $R(i)$ we write R_i and for $\mathscr{A}(i)$ we write $\mathscr{A}_i$; accessibility function R can be seen as a set of *accessibility relations* R_i, and V as a set of *valuations* $V(p)$. A pointed epistemic awareness model (M, s) is an *epistemic awareness state*.

The awareness function $\mathscr{A}$ may be varied to reflect different logics. The logic of *public global awareness* results if the value of $\mathscr{A}$ is the same for all agents and for all states. The logic of *individual global awareness* results if the awareness function is the same in all states, but may vary among agents. These logics are discussed in [6]. If there are no constraints placed on the awareness function $\mathscr{A}$, we call the logic that of *individual local awareness* [8]. For the sake of generality we will also assume no restrictions on the accessibility function R. However, for practical reasoning purposes we will often require that the relation satisfies some simple properties (such as reflexivity, transitivity etc.), where the typical model classes are $KD45$ (transitive, euclidean, and serial) and $S5$ (transitive, euclidean, and reflexive). Further, given an arbitrary model M we will refer to the elements of the tuple as $(S^M, R^M, \mathscr{A}^M, V^M)$.

The property of *awareness introspection* [16] holds if all agents know when they are aware of a proposition or of another agent: "If $(s,t),(s,u) \in R_i$, then $\mathscr{A}_i(t) = \mathscr{A}_i(u)$." We make no commitment to awareness introspection. The image of the awareness function is a set of propositions *and* agents. Awareness of agents is not commonly considered. The property of *self-awareness* holds if for all agents and states, $i \in \mathscr{A}_i(s)$.

We now introduce the notion of awareness bisimulation. Consider the following scenario (also depicted in Figure 3).

Example 2. In state s agent i is aware of agent j and of proposition p, state u is accessible for agent i from state s, and in state u agent j is aware of proposition p and also of proposition q. That agent j is also aware of q in u should leave agent i indifferent, as she is not aware of q in s!

This sort of similarity is captured in the following notion, named *awareness bisimulation*. Informally, given a model and a set (of propositions and agents) $A \subseteq P \cup N$, another model is an A awareness simulation if it cannot be distinguished from the first by formulas built only from proposition and agent variables in A, and only in the scope of agents who are aware of those propositions and agents.

Definition 4 (Awareness bisimulation). Let epistemic awareness models $M = (S, R, \mathscr{A}, V)$ and $M' = (S', R', \mathscr{A}', V')$ be given. For all $A \subseteq P \cup N$ we define the relation $\mathfrak{R}[A]$ by $(s,s') \in \mathfrak{R}[A]$ iff:

- **[atoms]** for all $p \in A$, $s \in V(p)$ iff $s' \in V'(p)$;
- **[aware]** for all $i \in A$, $\mathscr{A}_i(s) \cap A = \mathscr{A}'_i(s') \cap A$;
- **[forth]** for all $i \in A$, if $t \in S$ and $R_i(s,t)$ then there is a $t' \in S'$ such that $R'_i(s',t')$ and $(t,t') \in \mathfrak{R}[A \cap \mathscr{A}_i(s)]$;
- **[back]** for all $i \in A$, if $t' \in S'$ and $R'_i(s',t')$ then there is a $t \in S$ such that $R_i(s,t)$ and $(t,t') \in \mathfrak{R}[A \cap \mathscr{A}'_i(s')]$.

Epistemic awareness state (M',s') is an *A-awareness-bisimulation* of epistemic awareness state (M,s) iff $(s,s') \in \mathfrak{R}[A]$. We write $(M',s') \underline{\leftrightarrow}_A (M,s)$.

The 'aware' clause can be considered as an additional 'atoms' requirement, due to the nature of our models where states have more structure than merely propositionual truth. If we were to replace $\mathfrak{R}[A \cap \mathscr{A}_i(s)]$ in the **back** and **forth** clauses with $\mathfrak{R}[A]$, we would have the definition of a standard (restricted) bisimulation over labelled transition structures [25]. (Restricted to P' and N', for $P' \cup N' = A$.) Thus every bisimulation is an awareness bisimulation. Vice versa, if all agents are aware of all propositional variables and agents, the awareness bisimulation is a standard bisimulation (for the relations R_i). This is what we desire: we then revert to the standard multi-agent epistemic situation, where awareness plays no role.

Proposition 1. *The relation* $\leftrightarroweq_A$ *is an equivalence relation.*

Proof. This can be seen by examining Definition 4. If $(s,s') \in \mathfrak{R}[A]$, then the 'level of awareness' in s and s' is the same, as required by **aware**: for all $i \in A$, $\mathscr{A}_i(s) \cap A = \mathscr{A}_i'(s') \cap A$. Reflexivity, symmetry, and transitivity follow directly.

Definition 4 is more complex than the definition of standard bisimulation, however its motivation is very simple. Two states are A-awareness-bisimilar if, for any observer aware only of the propositions and agents in A, the states appear identical. It gives us the "A-perspective" of a an epistemic awareness model. We also call it *observational equivalence*. Let that observer be agent i in state s, then her perspective is that of $\mathscr{A}_i(s)$-awareness-bisimilarity. We might also say that her view of the model is that of its $\mathfrak{R}[\mathscr{A}_i(s)]$ equivalence class. Clearly, if this holds for all agents, noone can make an explicit distinction.

Definition 5 (Observational equivalence). Two epistemic awareness states (M,s) and (M',s') are *observationally equivalent* iff they are $\mathscr{A}_i(s)$-awareness bisimilar for all $i \in N$, i.e., iff

$$(M,s) \leftrightarroweq_{\bigcup_{i\in N} \mathscr{A}_i(s)} (M',s') \ .$$

The crucial part of the definition of awareness bisimulation is that in 'forth', in the requirement "$(t,t') \in \mathfrak{R}[A \cap \mathscr{A}_i(s)]$", the bisimulation for **state *t*** is (further) restricted to the propositions and agents visible for agent i in **state *s***, the i-predecessor of t. (And similarly for 'back'.) An honoured principle (also in economics, and in artificial intelligence) is that incompleteness precedes uncertainty. The awareness function of an agent in a given state (incompleteness) determines what the agent can 'see' in all accessible states (uncertainty), *and so on*. This chaining of awareness is expressed with awareness bisimulation. This chaining requirement was present in epistemic awareness structures since its inception in [10]. We have merely employed it to the full and in the one and only way, for structural similarity.

Example 3. In Figure 2 agent i is aware of p but unaware of q in state s. In the figure, names of states are followed, between parentheses, by values of propositions (as in t) or by the propositions the agent is aware of (as in s). (These conventions for visualization are different from those in Figure 1, but they generalize better to the multi-agent situation. We will keep the conventions throughout the contribution.) The three depicted epistemic states, wherein i (from left to right) implicitly knows q, knows $\neg q$, or does not know whether q, are observationally indistinguishable for the agent: they are $\{p,i\}$-awareness bisimilar. A $\{p,i\}$-awareness bisimulation between (e.g.) the left and the right picture is $\mathfrak{R} = \{(s,s''),(t,t''),(t,t''')\}$.

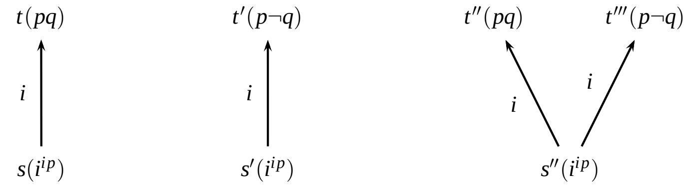


Fig. 2 Agent i is aware of p but unaware of q in state s.

Example 4. In Figure 3, agents i, j are aware of themselves and of each other in all state. Also, in state s agent i is aware of proposition p, and agent j is aware of propositions p and q. In state u agent j is also aware of q. In s, agent i explicitly knows p and agent j explicitly knows p and q. That agent j is not aware of q in u leaves agent i indifferent, as she is not aware of q in s. The model rooted in state s is observationally equivalent ($\{p,q,i,j\}$-awareness bisimilar) to (e.g.) the model with root s', where agent j is (obligatory) aware of q in s' but *not* aware of q in u', and where q is still false in u'. But it is also observationally equivalent to the model with root s'' where the awareness for both agents remains the same but the value of q is true in u''. This freedom does not exist for values of q in states t, t', t'': as j is aware of q in s (and s', s''), he knows explicitly that q and therefore, making q false in one of the states t, t', t'' would make an observational difference. The crucial aspect of awareness bisimulation is that in state s, j is aware that q is true in state t, but not from the perspective of i, whose access to state t is limited by what she considers possible for j within the limits of her awareness: e.g., we say that agent i considers it 'speculatively' possible that, counterfactually, j considers it possible that q is false in t. (It is counterfactual because j does not really consider that possible.) Also, i considers it speculatively possible that q is false in u, and that j is aware of q in u, and that j is unaware of q in u.

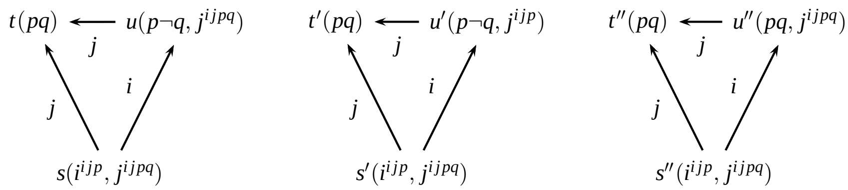


Fig. 3 Agent j's awareness of q in state t is inconsequential to agent i in state s, as i is not aware of q.

Examples with a $KD45$ or $S5$ setting as in Section 2 are given in Section 7.

5 Semantics

Having defined the language and the structures, we now define the semantics.

Definition 6 (Semantics). Let $M = (S, R, \mathscr{A}, V)$ be given.

$$
\begin{array}{ll}
(M,s) \models \top & \\
(M,s) \models p & \text{iff } s \in V(p) \\
(M,s) \models \varphi \wedge \psi & \text{iff } (M,s) \models \varphi \text{ and } (M,s) \models \psi \\
(M,s) \models \neg\varphi & \text{iff } (M,s) \not\models \varphi \\
(M,s) \models K_i^I \varphi & \text{iff } \forall t \in sR_i, (M,t) \models \varphi \\
(M,s) \models \forall p \varphi & \text{iff } \forall (M',s') \underline{\leftrightarrow}_{\overline{p}} (M,s), (M',s') \models \varphi \\
(M,s) \models \forall i \varphi & \text{iff } \forall (M',s') \underline{\leftrightarrow}_{\overline{i}} (M,s), (M',s') \models \varphi \\
(M,s) \models A_i \varphi & \text{iff } v(\varphi) \subseteq \mathscr{A}_i(s) \\
(M,s) \models A_i^+ p \varphi & \text{iff } (M^{s \mapsto_i p}, s) \models \varphi \\
(M,s) \models A_i^+ j \varphi & \text{iff } (M^{s \mapsto_i j}, s) \models \varphi \\
(M,s) \models A_i^- p \varphi & \text{iff } (M^{s \not\mapsto_i p}, s) \models \varphi \\
(M,s) \models A_i^- j \varphi & \text{iff } (M^{s \not\mapsto_i j}, s) \models \varphi
\end{array}
$$

where

$$
\begin{array}{ll}
M^{s \mapsto_i p} & = (S, R, \mathscr{A} \cup \{(i,(s,p))\}, V) \\
M^{s \mapsto_i j} & = (S, R, \mathscr{A} \cup \{(i,(s,j))\}, V) \\
M^{s \not\mapsto_i p} & = (S, R, \mathscr{A} \setminus \{(i,(s,p))\}, V) \\
M^{s \not\mapsto_i j} & = (S, R, \mathscr{A} \setminus \{(i,(s,j))\}, V)
\end{array}
$$

The set of validities (and the logic) is called *DLILA* (Dynamic Logic of Individual Local Awareness).

The semantics for $A_i^+ p$ (and similarly, for becoming aware of agents, and for forgetting) is only one of several options, to be discussed in Section 9.2. Using the semantics for the knowledge operator and the bisimulation quantifier, we can also define a novel notion called speculative knowledge, to which we devote Section 6. This notion will justify the use of bisimulation quantifiers. Even for the semantics of the quantifier, we have various options, depending on the model class over which quantification is considered. This is discussed in Section 7. So in fact, we do not just have a logic *DLILA* here, but a range of logics for knowledge and awareness, depending on setting various parameters.

Before the examples, some results. First, a negative result: *The semantics of DLILA are* **not** *invariant to bisimulation.* Consider a singleton model M with state s, accessible for a single agent i, where p is true and where the agent is unaware of p. This is clearly bisimilar to the two-state model N consisting of states t and t' such that $R_i = \{(t,t'),(t,t')\}$ and where p is still true in both states and where i is unaware of p in both. Clearly $(N,t) \underline{\leftrightarrow}_{\{p,i\}} (M,s)$, but $(N^{t \mapsto p}, t) \not\underline{\leftrightarrow}_{\{p,i\}} (M^{s \mapsto p}, s)$: the states t and t' can now be distinguished because i is aware of p in the former but not in the latter. In Section 9.1 we present some bisimulation invariant alternatives.

Proposition 2. *Awareness bisimilar states satisfy the same explicit knowledge: If* $(M,s) \models A_i \varphi \wedge K_i^I \varphi$ *and* $(M,s) \underline{\leftrightarrow}_{\mathscr{A}_i(s)} (M',s')$*, then* $(M',s') \models A_i \varphi \wedge K_i^I \varphi$.

Proof. Note that $A_i\varphi$ means $v(\varphi) \subseteq \mathscr{A}_i(s)$. In the language restricted to $\mathscr{A}_i(s)$ the epistemic awareness states (M,s) and (M',s') are therefore bisimilar in the standard sense, from which follows logical equivalence, thus equivalence of $A_i\varphi \wedge K_i^I\varphi$ in both states.

Proposition 3. *Duality of becoming aware and forgetting:* $\models A_i p \rightarrow (\varphi \leftrightarrow A_i^- pA_i^+ p\varphi)$ *and* $\models \neg A_i p \rightarrow (\varphi \leftrightarrow A_i^+ pA_i^- p\varphi)$. *(And similarly for agent variables.)*

Proof. Proposition 3 immediately follows from the semantics.

Example 5. Consider again Figure 2, and the roots of the models. In all three cases agent i explicitly knows that p. But she does not explicitly know in state s that q, because accessible state t is p-awareness bisimilar to (e.g.) t' wherein q is false. After becoming aware of q in state s, she explicitly knows q: then, any state $\{p,q\}$-awareness bisimilar to t *must* satisfy q. So $A_i^+ qK_i^E q$ is true.

Example 6. In Figure 3, in state s, it is true that both agents know that p, and that agent j knows that q but agent i is unaware of q. Therefore i considers it possible that agent j is not aware of q: $(s,u) \in R_i$, $(u,u') \in \mathfrak{R}[p]$, and $q \notin \mathscr{A}_j(u')$. Still, if i becomes aware of q, she knows that j is aware of q, and then she (mistakenly) believes that he incorrectly believes that q is false.

One way to pursue this logic further is to investigate its theoretical properties. In [8] decidability is shown for the logic $DLILA_K$ where all epistemic awareness states are trees, (and where only awareness of propositions is modelled, not of agents,) via an embedding into poly-modal logic. Poly-modal logic enjoys uniform interpolation with respect to both propositional atoms and actions [5]. Also in [8], a complete axiomatization is given for that logic, for a version of becoming aware where the agent becomes aware in all states, not merely in the actual state. The axioms in that logic that do not involve dynamics are comparable to those found in [10]. But in the underlying paper we keep focussing on the semantics, on intriguing concepts that we can express in the semantics with the available primitive operators, and on a number of semantic variations.

6 Implicit, speculative, and explicit knowledge

The presence of bisimulation quantifiers in the logic allows us to define a novel epistemic operator, called speculative knowledge.[1] This treatment of knowledge is a main innovation in these semantics.

Definition 7 (Speculative knowledge). Let an epistemic awareness model (M,s) be given.

$$(M,s) \models K_i^S\varphi \text{ iff } \forall t \in sR_i,\ \forall (M',t') \underline{\leftrightarrow}_{\mathscr{A}_i(s)} (M,t), (M',t') \models \varphi$$

[1] In [8], K_i^S was called introspective knowledge and written as K_i.

If the set of proposional and agent variables were finite this definition is equivalent to

$$(M,s) \models K_i^S \varphi \ \text{ iff } (M,s) \models K_i^I \forall \overline{\mathscr{A}_i(s)} \varphi$$

But we did not require the sets of propositional and agent variables to be finite. There still is a more convenient direct way to define speculative knowledge with a quantifier, that always works. Let $X = \overline{\mathscr{A}_i(s)} \cap v(\varphi)$. Then

$$(M,s) \models K_i^S \varphi \ \text{ iff } (M,s) \models K_i^I \forall X \varphi$$

We note that the value of φ in any model only depends on the (finite set of) variables occurring in φ. (Any further bisimulation variation that we take into account, does not matter.)

An agent speculatively knows φ only if in all accessible states φ remains true for *every* possible interpretation of all propositions and agents that she is unaware of. We achieve this by extending the agent's accessibility relation by composing it with bisimulation modulo those propositions and agents of which the agent is unaware. This speculative knowledge is neither explicit nor implicit knowledge.

In the definition of speculative knowledge, the order of epistemic access and speculation is important. Given a state s, in the semantics for $K_i^S \varphi$ *first* we access a possible world t and *only then* we compare observably indistinguishable states, but from the agent's perspective in the source state s, not in that target state t. This mirrors the corresponding crucial aspect of awareness bisimulation. The meaning of the quantifier only makes sense in the context of modal accessibility.

Example 7. The semantics of becoming aware is simple, but should really be seen in the context of speculative knowledge, where the quantification comes into play. Suppose that the agent is unaware of p and that p is true in all accessible states. We then have that $A_i^+ p K_i^E p$ is true: after the agent becomes aware of p, p is true. But although the agent considers that as a possibility, she does not know that, and she also considers it possible that after becoming aware of p, she knows that p is false, or that she is uncertain about p: all true are $L_i^S A_i^+ p K_i^E p$, $L_i^S A_i^+ p K_i^E \neg p$, and $L_i^S A_i^+ p \neg (K_i^E p \vee K_i^E \neg p)$.

Example 8. Proposition 3 states that if an agent i first forgets a proposition p and then recalls the proposition p, then nothing changes: $\models A_i p \rightarrow (\varphi \leftrightarrow A_i^- p A_i^+ p \varphi)$. **However**, if an agent i forgets a proposition p, she does not *know* that recalling the proposition p will return the world to its original state. Not *explicitly*, which is obvious, but also not *speculatively*: she can speculate over many different ways in which this were not to come to pass: $\varphi \rightarrow A_i^- p K_i^S A_i^+ p \varphi$ is invalid! Although ostensibly the dynamic operators are $A_i^+ p$ and $A_i^- p$, to grasp the dynamic expressivity we need the context of the agent's knowledge in combination with quantification.

Let us list once more the three notions of knowledge. Given a model M with a state $s \in M$.

implicit knowledge	$K_i^I\varphi$	φ is true in all i-accessible states,
speculative knowledge	$K_i^S\varphi$	$K_i^I\forall(\overline{\mathscr{A}_i(s)}\cap v(\varphi))\varphi$
explicit knowledge	$K_i^E\varphi$	$K_i^I\varphi\wedge A_i\varphi$

Past literature on knowledge and awareness has focused on the difference between implicit knowledge ("knowing" something without being fully aware of that thing) and explicit knowledge ("knowing" something as well as being fully aware of that thing). Speculative knowledge is strictly weaker than explicit knowledge and strictly stronger than implicit knowledge. It allows us to reason about the process of becoming aware, and that is our reason to complicate the existing picture. For explicit and implicit knowledge we have retained the usual definitions. Interaction between the three kinds of knowledge includes:

Proposition 4. $\models K_i^E\varphi\rightarrow K_i^S\varphi$ *and* $\models K_i^S\varphi\rightarrow K_i^I\varphi$.

On the other hand, $\not\models K_i^I\varphi\rightarrow K_i^S\varphi$ (you can implicitly know that p but, as you are unaware of p, you do not speculatively know that p), and $\not\models K_i^IK_i^I\varphi\rightarrow K_i^SK_i^I\varphi$ (as you may be unaware of yourself).

Explicit knowledge is the truest form of knowledge. It is the knowledge an agent is totally aware of. Implicit knowledge is the idealized form of knowledge. If proposition φ is true in every state an agent considers possible, then the agent implicitly knows φ even though she may be unaware of some of the propositions and agents in φ. As φ is not in the agent's language of discourse, the agent cannot access this information. If we were to speculate about the agent's interpretation of φ, we must assume that any states observationally indistinguishable to the agent for the propositions and agents of which that agent is aware, are just as possible in the agent's perspective. This allows us to model how agents can perceive other agents becoming aware in the most general way.

There is another reason to consider this novel notion of speculative knowledge: *Maybe K^S is the notion that should really be called implicit knowledge.* That should delegate K^I to the level of a mere technical primitive, or something called 'latent' knowledge, say. Works predating [10] such as [20] support this view. In Levesque's presentation, variables in the unaware set can be speculated over and that is exactly what in his writings is called implicit knowledge. We have more to say about this, but in this contribution we focus on awareness dynamics.

If the agent is aware of every proposition and agent in the formula φ, then the agent speculatively knows φ, if and only if the agent explicitly knows φ, if and only if the agent implicitly knows φ. The agent speculatively knows φ if for every world indistinguishable from the current world, for every epistemic awareness model bisimilar to that world with respect to all propositions and agents in φ, we have that φ is true. If two worlds are bisimilar with respect to all propositions and agents in φ, then the interpretation of φ in those two worlds is identical (see Proposition 2). So in the presence of complete awareness the interpretation of all three forms of knowledge is as for epistemic logic.

Example 9. John knows the firewall is working. He is not aware of any new worm that may attack the system. However, he speculatively knows that in any possible world where such a worm existed, the firewall would be working. This is captured by the validity

$$K^S_{john} f \wedge \neg A_{john} w \to K^S_{john}(w \to f)$$

where f is *the firewall is working* and w is *the worm is attacking the system*.

Example 10. The animal knows, of course. But it certainly does not know that it knows. *(Teilhard de Chardin)*

Suppose that we have a simple animal or agent that is not self-aware. That is while the agent may have the capacity to reason about a number of propositions it knows, it is not able to reason about its own *knowing* of those propositions. This is precisely captured by the semantics of speculative knowledge. If this animal, a, is aware of a nocturnal threat, t, and aware of the proposition that it is night, n, then the agent may know that in every possible world that it is day, there is no threat. In terms of the semantics, in every world s that the agent considers possible, in every world $\{n,t\}$-awareness-bisimilar to s (for the propositions and agents n and t are the only aspects of this possible world that the animal may note), we have $t \to n$ being true. However, the agent does not know that it knows. Speculatively, it considers it feasible that some agent a may know such a proposition, but without any self-awareness it makes no connection between such an agent and itself.

7 KD45 and S5

Apart from the logic *DLILA* we also consider the logics $DLILA_{\mathrm{L}}$, where every modal operator K^I_i satisfies the axioms of the logic L. In [8] we investigate $DLILA_K$ where all epistemic awareness states are trees. Other typical choices of L are S5 and KD45. One should be careful to note that these are *not* simple cases of restriction. Restricting the underlying logic to L (for example KD45) means that in interpreting the formula $\forall p\varphi$ (or $\forall i\varphi$), we may only consider pointed models (M',t') that satisfy the constraints of L (so transitive, serial and euclidean for KD45). The validities of $DLILA_{\mathrm{L}}$ therefore do not necessarily extend those of *DLILA*. Indeed, each axiomatization poses new problems. Specific logics will also require us to vary the semantic interpretation of the operators $A^+_i p$ and $A^-_i p$. If awareness introspection holds (indistinguishable states have the same level of awareness, so that you know of what you are aware), then apart from becoming aware of some variable p in the actual state, you should also become aware of that variable in all accessible states; or else, awareness introspection is not preserved. In this section we assume that this is indeed the case. (In Section 9.1 we will more systematically investigate such versions of becoming aware.)

Example 11. Consider the case of $DLILA_{KD45}$, where every agent's accessibility relation is transitive, serial and euclidean. Crucially, in $KD45$, strong beliefs may be mistaken, but you do not consider that possible: to yourself, your beliefs appear knowledge. So $L^E_i(\neg p \wedge K^E_i p)$ is inconsistent. However, in $DLILA_{KD45}$ it is valid that an agent i considers it possible that she becomes aware of a propositional variable p that is false and that she believes to be true. That is nothing but speculating about becoming aware of false information that you had reason to accept! A validity of the language is

$$\neg A_i p \rightarrow L_i^S A_i^+ p(\neg p \wedge K_i^E p)$$

The interpretation of this formula is shown in Figure 4. The crucial aspect is that the pair $(s,t) \in \mathfrak{R}[\emptyset]$ (the dashed line): agent i cannot a priori distinguish the reality of p being true in the believed state from the speculative option that p is false there but believed true. However, after becoming aware of p (in both s and t) this option is out of reach, as $(s,t) \notin \mathfrak{R}[p]$.

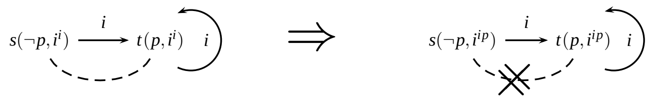


Fig. 4 You can become aware of a false belief.

More than any other example, we find speculation over false beliefs a very strong vindication for our novel notion of speculative knowledge.

Example 12. Now consider the case of $DLILA_{S5}$, where every agent's accessibility relation is an equivalence. This is the case for the coffee and orange juice examples in Section 2. Consider Figure 1 again on page 5, reprinted here as Figure 5, and the logical descriptions in Example 1 on page 7. Let s be the actual state in Ri, where p and q are both false.

- After becoming aware of oranges, Hans explicitly knows that coffee and oranges are not both served: $Ri, s \models A_i^+ q K_i^E \neg(p \wedge q)$
- Hans considers it speculatively possible that after becoming aware of oranges, he knows that orange juice and coffee are both available: $Ri, s \models L_i^S A_j^+ q K_i^E (p \wedge q)$.

Another formalization of the second is $L_i^S K_i^E (p \wedge q)$, i.e., $L_i^I \exists q K_i^E (p \wedge q)$. This is because $\exists q A_j^+ q \varphi$ implies $\exists q \exists q \varphi$ and therefore, trivially, $\exists q \varphi$ for any φ (but not the other way round): the bisimulation quantification over q also allows to vary the awareness level of q, not just the valuation. In other words, the epistemic awareness state wherein you have become aware of a variable q is $\overline{q}$-bisimilar to the epistemic awareness state wherein you were not yet aware of that variable.

8 Building models while becoming aware

There is a completely different way to model becoming aware, but for which we already have all the necessary tools. Figure 5 depicts two different ways to become aware of q in a model wherein q already has a value.

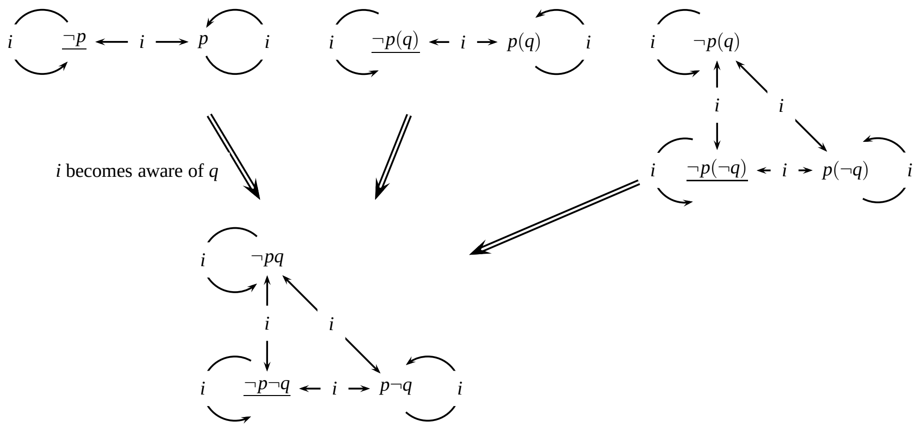


Fig. 5 Same as Figure 1 on page 5. The models are named *Le*, *Mi*, *Ri*, and *Bo*.

So far, we have modelled becoming aware as in the transition from *Ri* to *Bo*. In all the states of *Ri*, we move q to the set of propositions of which i is aware. Technically, this is the $A_i^{S+}q$ operation in the $S5$-version where the agent becomes aware of the variable in all accessible states.

From the model *Mi* we can also arrive at *Bo*, namely by increasing the epistemic complexity, by changing the value of q, and by moving q to the set of aware propositions. Technically, this is a combination of $\{i, p\}$-restricted bisimulation, i.e., a quantification $\exists q$, followed by the same $A_i^+ q$. Having atom q change its value during the $\exists q$ operation is a price we pay: the price is that $K_i^I \varphi$ no longer meaning implicit knowledge in the usual sense of 'knowledge that the agent can become aware of'. Before the transition, $K_i^I q$ is true. But in the transition this gets lost: afterwards, she is explicitly ignorant of q. Therefore, the process of becoming aware can now not be seen as explicitizing implicit knowledge.

This alternative formalization of 'becoming aware of φ' (also presented in [6]), in the sense of 'there is a way in which the agent can become aware of q, after which φ', comes with the following definition.

Definition 8 (Becoming aware while building the world).

$$
\begin{array}{lll}
(M,s) \models \mathbf{A}_i^+ p\varphi \text{ iff } & (M,s) \models \exists p A_i^+ p\varphi & \text{if } p \notin \bigcup_{j\in N} \mathscr{A}_i(s) \\
& (M,s) \models A_i^+ p\varphi & \text{otherwise}
\end{array}
$$

Becoming aware of agents, and forgetting, is defined similarly.

The case distinction in the definition is necessary. Let i be unaware and become aware of p, and consider an agent $j \neq i$ who is already aware of p. If p changes value in the $\exists p$-operations, than j would *observe* that change and be puzzled. It would be like Tim already sipping coffee in the dining room while Hans' becomes aware of coffee. Hans speculates over coffee maybe not being available there. As a result of that Tim sees to his consternation both the cup and the saucer slowly fade, until they have disappeared from his hands. This, we do not want.

With this view of becoming aware, we can think of agents gradually becoming aware of themselves, of other agents, and of propositions, and in this way gradually expanding a structure representing their uncertainties. This view is not incompatible with the other modelling of becoming aware—wherein the real and always present structure of the world is gradually revealed, so to speak—but then we have to think of the smaller structures as equivalence classes instead. For example, the model *Le* in Figure 1 is the $\underline{\leftrightarrow}_{\{i,p\}}$-equivalence class.

It seems a bit strange that forgetting is also model building. Should it not instead be some deconstruction, and make models simpler again? If you forget variable p, you *may* have this variable keep the same value in the resulting model in all states. But, as the value is now don't care, there is nothing against giving that variable the same value on the whole model. And that means *loss* of structure. Another way of seeing this: view the resulting model as a $\overline{p}$-equivalence class *but* keep a value for p: give p the same value in all states of that equivalence class. Then, perform a bisimulation contraction. Doing that, the model resulting from forgetting will typically be smaller.

We tend to view the novel becoming aware operator as a demonstration of the flexibility our structures allow in defining logics and in customizing operators. With the new operator come new questions too: given that $\boldsymbol{A}_i^+ p$ models a new way of becoming aware, and using the modelling principle that unaware variables in structures may have no-care values, can we also be precise about *which* way to become aware we want in some descriptive operator? Yes, we can. The mind reset one needs for an elegant solution is to view change of awareness and change of information simultaneously, and then to regard change of awareness as a special case. This is the focus of the next section.

9 Dynamics of awareness and dynamics of knowledge

In this section we elaborate on different ways in which we can further develop the framework for knowledge and change of awareness, towards a direction that combines change of knowledge with change of awareness. A number of strands also developed in other works meet here, mainly [7], [14], [26], and [9]. Concerning Section 9.2: [7] and [14] also address the matter of public announcements and awareness, although quite differently. Chapter 3 of Velazquez' PhD thesis [26] presents action models in a version involving awareness, a explicit motivation for and precursor of our proposal in Section 9.3. In [9] van Eijck *et al.* combine structures for disjoint or overlapping vocabularies. Although awareness is not the subject of their investigations, the modal product construction applied in epistemic awareness actions in Section 9.3 is clearly reminiscent of and inspired by their work.

9.1 Different ways of becoming aware

We recall once more the semantics of becoming aware—we take the case of becoming aware of a proposition, but the entire story unwinding here equally applies to becoming aware of an agent; and equally, it applies to forgetting of proposition or agents.

$$(M,s) \models A_i^+ p\varphi \text{ iff } (M^{s\mapsto_i p}, s) \models \varphi$$

where

$$M^{s\mapsto_i p} = (S, R, \mathscr{A} \cup \{(i,(s,p))\}, V)$$

This is only one way to model becoming aware. First, we recall an observation from Section 5: if the model is not a bisimulation contraction, point-wise becoming aware does not preserve bisimilarity. A logically inelegant way to correct this is to require it to be performed on bisimulation contractions only. But it has other drawbacks too, that we already encountered in Section 7: point-wise becoming aware does not preserve awareness introspection (same level of awareness in epistemically indistinguishable states). In $KD45$ and $S5$ structures awareness introspection guarantees that agents know that they are aware of propositions. Consider again the case of $S5$ structures, and an agent with a non-singleton equivalence class, wherein she is unaware of p throughout the class. After becoming of aware of p in the actual state only, she has lost awareness introspection. This is undesirable for many applications. In Section 7 we proposed an alternative semantics for becoming aware, wherein the agent becomes aware in all accessible states. Now, let us look at this more systematically.

Consider the following three alternative semantics for an individual agent becoming aware—including the point-based one we already know. All three consist of making an unaware variable into an aware variable, i.e., changing the set $\mathscr{A}$ in a model but leaving all other parameters the same. Given state s, one can make agent i aware of the propositional variable p:

- in the actual state (only):
  $M^{s\mapsto_i p} = (S, R, \mathscr{A} \cup \{(i,(s,p))\}, V)$;
- in the actual state and all states accessible for agent i:
  $M^{sR_i\mapsto_i p} = (S, R, \mathscr{A} \cup \{(i,(s,p))\} \cup \{(i,(t,p)) \mid t \in S \text{ and } R_i(s,t)\}, V)$;
- in all states of the model:
  $M^{S\mapsto_i p} = (S, R, \mathscr{A} \cup \{(i,(t,p)) \mid t \in S\}, V)$.

To distinguish them in the language we name them $A_i^{s+}p$, $A_i^{R+}p$, and $A_i^{S+}p$. To this triple we add another one, and for which we also introduce a novel operator in the logical language:

$$(M,s) \models A^+ p\varphi \text{ iff } (M^{S\mapsto p}, s) \models \varphi$$

where

$$M^{S\mapsto p} = (S, R, \mathscr{A} \cup \{(i,(t,p)) \mid t \in S, i \in N\}, V)$$

In other words: all agents become aware of p. It is not definable in terms of the previous three (or any other single-agent based version of becoming aware), because there is an infinite number of agents. Except for the first, point-based becoming aware, the three variations of becoming aware are all bisimulation invariant.

One road down the stamp-collecting track is to expand this with further variations and epicycles. But there is a way to unify these alternatives in a single definition, at negligable cost. Instead of specifying directly where in the structure the agent has to become aware, we can also require this more properly in a purely linguistic way, namely by way of a precondition for becoming aware in the logical language. Consider the following definition for *conditionally becoming aware*—we write the becoming aware operator between square brackets '[' and ']' in order to distinguish it clearly from the *being* aware operator $A_i\varphi$ and the becoming aware operators like A_i^+ considered so far.

Definition 9 (Conditionally becoming aware).

$$(M,s) \models [A_i^\varphi p]\varphi \text{ iff } (M',s) \models \varphi$$

where M' is as M except that

$$p \in \mathscr{A}_i'(s) \text{ iff } (M,s) \models \varphi$$

For one thing, this immediately solves our worries about bisimulation preservation. The language with this operator as in inductive construct instead of all the A_i^+ and A_i^- operators is clearly bisimulation invariant. But, subject to some restrictions, we also have that all of the former are now definable. Let (M,s) be a finite model and a bisimulation contraction. Then:

$$\begin{array}{l}(M,s) \models A_i^{s+} p\varphi \text{ iff } (M,s) \models [A_i^{A_i p \vee \delta_s} p]\varphi \\ (M,s) \models A_i^{s-} p\varphi \text{ iff } (M,s) \models [A_i^{A_i p \wedge \neg\delta_s} p]\varphi \\ (M,s) \models A_i^{S+} p\varphi \text{ iff } (M,s) \models [A_i^{\top} p]\varphi\end{array}$$

where δ_s is a *distinguishing formula* [4, 2] for s in M (these always exist for finite models). A similar take can be used for all agents becoming aware. On finite $S5$ models, we also have

$$(M,s) \models A_i^{R+} p\varphi \text{ iff } (M,s) \models [A_i^{A_i p \vee L_i^I \delta_s} p]\varphi$$

If M is not a bisimulation restriction, the actual state s may be similar to another state in M, and then the distinguishing formula cannot capture the actual state alone. So then, $A_i^+ p$ is not definable in this way. That is the price we pay. The reason that we consider this neglibable cost is that it is rather a prize we carry, why on earth would one *want* to separate the actual state from a bisimilar one?

Finally, note that

$$[A_i^\varphi p]\varphi \rightarrow \exists p\varphi$$

is valid, and therefore also, e.g., $A_i^+ p\varphi \rightarrow \exists p\varphi$, as already noted in Example 12. This is because $(M,s) \leftrightarrows_{\overline{p}} (M',s)$, where M' is the model with the changed awareness function in Definition 9. We also have that $(M,s) \leftrightarrows_{\overline{p}} (M^{s \mapsto_i p},s)$, etc.

9.2 Public announcements and awareness

Consider an extension of the language $\mathscr{L}$ with operators $[!\varphi]\psi$ for 'after truthful public announcement of φ, ψ holds' [24]. We now can model the notion of 'addressing the issue', i.e., of announcing something of which the listeners were as yet unaware. Let us write

Definition 10 (Becoming aware in a public announcement).

$$[!_A\varphi]\psi := [!\varphi]A^+v(\varphi)\psi$$

where $A^+v(\varphi)$, defined in the previous Section 9.1, means that in all states of the model all agents become aware of all the unaware (agent and propositional) variables in the announced formula φ. It is imperative that $A^+v(\varphi)$ is after the announcement and not before. After all, a truthful announcement to all could be

"You are not aware of the (fact that there is a) revolution in Egypt!"

This is true at the moment of the announcement, but after that no longer: all have now become aware. We have just discovered the *unsuccessful awareness update*, an announcement to i of the archetypical form $!_A(p \wedge \neg A_i p)$.

Another way to model addressing an issue is along the lines of [7] and the model building treatment in Section 8:

Definition 11 (Becoming aware in a public announcement—variation).

$$[!_{\mathbf{A}}\varphi]\psi := \exists X[!\varphi]A^+v(\varphi)\psi$$

where, as before, X is the set of variables (agents or propositions) in φ of which no agent is aware, i.e., $X = \{x \in v(\varphi) \mid x \notin \bigcup_{i\in N} \mathscr{A}_i(s)\}$.

Now, we increase the complexity of the model to the extent necessary to incorporate the new information, while all the time ensuring consistency with prior explicit knowledge.

9.3 Epistemic awareness actions

We now combine the conditionally becoming aware operators and the obvious generalization of the announcements that address an issue, in order to be able to model any combination of awareness change and information change, including private actions that are more complex than public information change wherein some but not all become aware, and so on. We propose *epistemic awareness actions*.

Definition 12 (Epistemic awareness action). An *epistemic awareness action model* for N and P is a tuple $\mathsf{M} = (\mathsf{S}, R, \mathscr{A}, \mathsf{pre}, \mathsf{post})$ that consists of a *domain* S of *actions*, an *accessibility function* $R : N \to \mathscr{P}(\mathsf{S} \times \mathsf{S})$, an *awareness function* $\mathscr{A} : N \to \mathsf{S} \to$

$\mathscr{P}(P \cup N)$ (which determines the new awareness level of each agent after execution of s), a *precondition function* $\mathsf{pre} : \mathsf{S} \to \mathscr{L}$ that to each action assigns a precondition for its execution, and a *postcondition function* $\mathsf{post} : \mathsf{S} \to P \to \mathscr{L}$ assigns in each action to each propositional variable a formula (which determines the new value of that variable after that action). A pointed epistemic awareness action model (M, s) is an *epistemic awareness action*.

The execution of an epistemic awareness action in an epistemic awareness state is the usual suspect, the restricted modal product $(M \otimes \mathsf{M})$ [1] of both, wherein only pairs (s, s) survive such that $s \models \mathsf{pre}(\mathsf{s})$, where the value of a proposition p in a pair (s, s) is determined by the postcondition function, namely such that $(s, \mathsf{s}) \in V(p)$ iff $M, s \models \mathsf{post}(s)(p)$, and with the novelty that the awareness function of the epistemic awareness action overwrites the previous level of awareness: agent i is aware of proposition p in pair (s, s) of the modal product, i.e., $(M \otimes \mathsf{M}), (s, \mathsf{s}) \models A_i p$, iff $p \in \mathscr{A}_i(\mathsf{s})$. (And similarly for awareness of agent variables.) In other words, prior to execution, $p \in \mathscr{A}_i(s)$; but post-execution, $p \in \mathscr{A}_i(\mathsf{s})$: in (s, s), awareness in state s is replaced by awareness in s. (There are more efficient ways to handle this, but they are less general.)

We can now define various of the previously defined operators as epistemic awareness action: those for becoming aware, for combinations of becoming aware and becoming informed (addressing an issue), or, most poignantly, those involving bisimulation quantifiers in the 'model building' version of becoming aware.

Example 13 (Becoming aware). $M, s \models [A_i^{\varphi} p]\varphi$ iff $M \otimes \mathsf{M}', (s, \mathsf{s}') \models \varphi$, where M' consists of two actions s' and t', indistinguishable from one another for all agents, with $\mathsf{pre}(\mathsf{s}') = \varphi$ and $\mathsf{pre}(\mathsf{t}') = \neg\varphi$, with post' the trivial assignment for both actions ($\mathsf{post}'(\mathsf{s})(q) = q$ for all $q \in P$, i.e., including p; for $\mathsf{s} = \mathsf{s}', \mathsf{t}'$), and the awareness function in M' adds p to the awareness set for agent i in the φ-states of M, i.e.: for all (s, s) in the domain of the product (for all s and s such that $M, s \models \mathsf{pre}(\mathsf{s})$), $\mathscr{A}'$ should be defined such that for agent i, $q \in \mathscr{A}_i'(\mathsf{s})$ iff $q \in \mathscr{A}_i(s)$ or $q = p$; and for all other agents j, $\mathscr{A}_j' = \mathscr{A}_j$.

With the additional proviso that models are finite bisimulation contractions, the operators $A_i^+ p$ and $A_i^- p$ are special cases of the above (see Section 9.1).

Example 14 (Addressing a novel issue in an announcement). $M, s \models [!_A \varphi]\psi$ iff $M \otimes \mathsf{M}'', (s, \mathsf{s}'') \models \varphi$, where M'' consists of the single action s'', accessible to all agents in N, with precondition φ, with trivial postcondition, and such that $p \in \mathscr{A}_i''(\mathsf{s})$ iff $p \in \mathscr{A}_i(s)$ or $p \in v(\varphi)$.

Example 15. Consider the transition of model *Mi* to *Bo* in Figure 5. This is only one way for i to become aware of q. Now, we can be precise about *which* exact way. In the model *Mi*, let the left state where p is false and q is true be called s and the right state, where p and q are both true, be called t.

Consider the epistemic awareness action model M consisting of two actions s and t, with universal access for i on this two-action domain, and such that

- $\mathsf{pre}(\mathsf{s}) = \top$, $\mathsf{post}(\mathsf{s})(q) = \bot$, $\mathsf{post}(\mathsf{s})(p) = p$ and $\mathscr{A}_i(\mathsf{s}) = \{p, q\}$;

- $\mathsf{pre}(\mathsf{t}) = \neg p$, $\mathsf{post}(\mathsf{t})(q) = \top$, $\mathsf{post}(\mathsf{t})(p) = p$ and $\mathscr{A}_i(\mathsf{s}) = \{p, q\}$.

An execution of epistemic awareness action (M, s) in (Mi, s) is depicted in Figure 6.

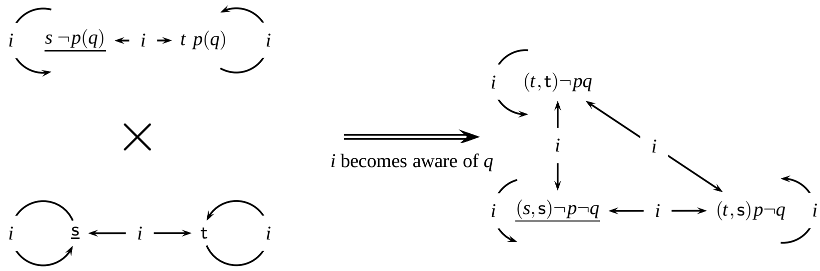


Fig. 6 Agent *i* becoming aware of *q* as a modal product construction. Compare to Figure 1.

This is not the end of the story, but the beginning of another story. For one thing, the epistemic awareness actions are unsuitable as primitives in the language in this form, because infinitely many variables may be assigned a new value according to the postcondition function and also because there are infinitely many agents. There are standard ways to deal with this, e.g., [3] allows only for a *finite* number of propositional variables to be assigned a new value (such that the infinite remainder does not change value). This makes it possible to apply an inductive language definition again, with $[\mathsf{M}, \mathsf{s}]\varphi$ as an additional inductive construct (for 'after execution of epistemic awareness action (M, s), φ'). But the infinity of agents causes additional (inductive definition) language trouble: we cannot enumerate the action models anymore, we will never arrive at the ones for an infinite number of agents... We think to have tricks around this, for example, we are only interested in models with *initially* only a finite number of agents that are aware of a finite number of other agents and a finite number variables (so, e.g., public announcement means public announcement to *those* agents only); and instead of an awareness function in an action model overwriting the old awareness function, we can envisage an awareness function merely *changing* a finite part of the awareness. We are confidently expecting to resolve these issues in future research, and also to find an axiomatization.

In epistemic awareness actions, information change and awareness change are combined.

To model pure awareness change, we need additional constraints. That is easy: each agent should not be allowed to learn any non-trivial formula in every equivalence class, i.e., for all $i \in N$, and for all $\mathsf{s} \in \mathsf{S}$ (of some given awareness epistemic action)

$$\bigvee_{\mathsf{t} \in \mathsf{s}R_i} \mathsf{pre}(\mathsf{t}) \leftrightarrow \top$$

If the disjuntion of all preconditions in any equivalence class is true, then the agent will not gain hard information, but will only (possibly) change his level of awareness and his view on the change of others' levels of awareness.

To model pure information change, the requirement is that the awareness does not change anywhere, i.e., the awareness function $\mathscr{A}$ in an action model M should be such that for all agents i and pairs (s, s) such that $s \models \mathsf{pre}(\mathsf{s})$:

$$\mathscr{A}_i(\mathsf{s}) = \mathscr{A}_i(s)$$

Thinking a bit ahead in the more remote future, we expect that (on finite models) each execution of an epistemic awareness action model in a given epistemic awareness state results in an *awareness simulation* of that epistemic awareness state, and that, vice versa, each awareness simulation can be modelled as the execution of an epistemic awareness action; where an *awareness simulation* is like an awareness bisimulation but without the **forth** requirement. This would prove that epistemic awareness actions are a suitable and adequate formalization for *any conceivable change of awareness or information*, similar to the result for epistemic actions and simulations in [7].

10 Comparison

Our structures for knowledge and awareness are *almost* like those in [10]. We consider multi-K structures, i.e., any accessibility relation, whereas [10] assume $KD45$ epistemic relations and, to mention another similar approach, [16] assume (implicitly) multi-$S5$. Our multi-K choice facilitates some theoretical results, such as the proof of decidability in [8]. For propositional variables our awareness function is the most general (semantic) choice, as in [10]. As far as we know, awareness of agents is not considered in the literature on knowledge and awareness, prior to our own work [6].

Our approach is in some respects simpler and more constrained than [16]. From the epistemic awareness structure we are able to implicitly derive a complete lattice of spaces via awareness bisimulation, whereas in [16] this structure is given explicitly. In other words, we have a succinct, technical tool to derive that result. For example, consider Figure 2 in [16, p.86], reprinted here as Figure 7. The two possibilities in row jp, uq and columns p, q and $p, \neg q$ in $S_{\{p,q\}}$ are both linked to the single possibility in row jp and column p in $S_{\{p\}}$. We would say that the latter represents the $\mathfrak{R}[p]$-equivalence class that contains the two former: if the agent is unaware of q, it is compatible with this level of awareness that q is true but also that q is false. Similarly, in this figure, the single possibility $\emptyset$ represents the $\mathfrak{R}[\emptyset]$-equivalence class of the four possibilities in the bottom row of $S_{\{p,q\}}$, and, if there were only one atom in the language, the $\mathfrak{R}[\emptyset]$-equivalence class of the two possibilities in the bottom row of $S_{\{p\}}$, etc.

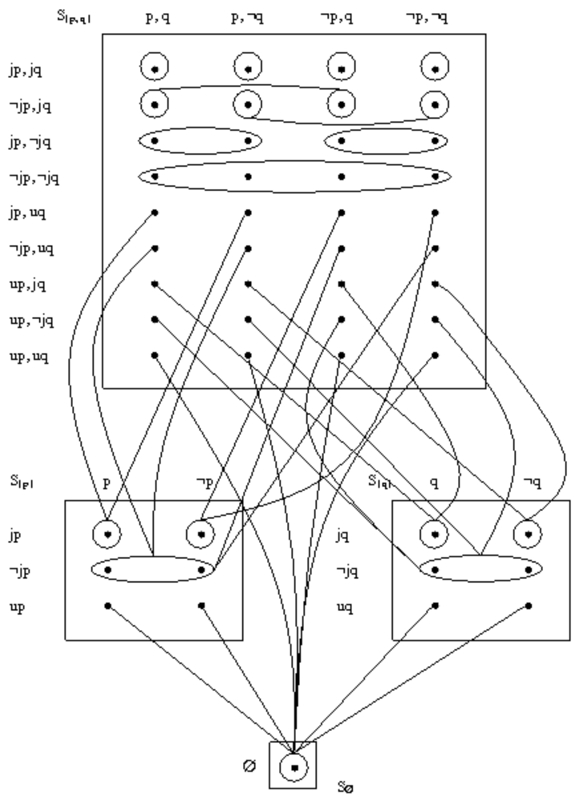


Fig. 7 An agent being aware and not being aware of propositions *p* and *q*. Reprinted from [16, p.86].

Now compare Figure 7 to Figure 1. Instead of arrows connecting states, draw ovals around indistinguishable states. What results? They are almost the same! This is not accidental. On purpose we have taken the same variables, and a single agent, exactly as in [16]. The state *Le* is found in Figure 7 as the lowermost left oval. The transition 'agent *i* becomes aware of proposition *q*' in Figure 1 is found in Figure 7 as a relation between the lower left square and the upper square. (Or, if you wish, as a relation between the lower left oval and the isolated points in the upper square that four curved lines are connected to.)

There is a difference, though. Our figure *Bo* represents uncertainty over three of four valuations. Figure 7 from [16] (intentionally) only visualizes part of all possible uncertainty and unawareness about two variables: an oval connecting three states is missing there.

Propositional quantification is integrated with awareness and knowledge in [15] (and in various precursors). This concerns quantification over the set of formulas of which an agent is aware. They interestingly mention that “Using semantic valuations [for quantification] does not work in the presence of awareness” [15, p.506]; although of course correct, we are wondering if our work may make the authors reconsider the suggested scope of that remark.

Dynamics of (propositional) awareness is also presented in [17] and in [14]. In the former, becoming aware means (initially) becoming ignorant about that proposition. It uses an algebraic approach. In the latter, the approach in Section 3 is similarly dynamic modal as ours, and it provides an interestingly integrated combination of syntactic and semantic awareness. This approach is taken much further in Velazquez’ recent PhD thesis [26].

11 Conclusions

We presented various logics that combine knowledge, awareness, and change of awareness. An agent can become aware of propositional propositions but also of other agents or of herself, and the dual operation of forgetting is also incorporated in the framework. Crucial to our results is the novel notion of *awareness bisimulation*, the obvious notion of modal similarity for structures encoding knowledge and awareness. Some tentative results link our quantificational logics with knowledge and awareness to extensions of dynamic epistemic logic with awareness. Axiomatizations for some of our proposals are future research, and also further elaboration of the notion epistemic awareness action.

Acknowledgements

Discussion with participants of ICLA Delhi were helpful to pursue this research. We wish to thank Thomas Ågotnes for encouraging us. Hans van Ditmarsch is also affiliated with the Institute of Mathematical Sciences Chennai (IMSC), India, as associated researcher.

References

1. Baltag, A., Moss, L., Solecki, S.: The logic of public announcements, common knowledge, and private suspicions. In: I. Gilboa (ed.) Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 98), pp. 43–56 (1998)
2. van Benthem, J.: Dynamic odds and ends. Tech. rep., University of Amsterdam (1998). ILLC Research Report ML-1998-08
3. van Benthem, J., van Eijck, J., Kooi, B.: Logics of communication and change. Information and Computation **204(11)**, 1620–1662 (2006)

4. Browne, M., Clarke, E., Grümberg, O.: Characterizing Kripke structures in temporal logic. In: H. Ehrig, R. Kowalski, G. Levi, U. Montanari (eds.) TAPSOFT '87, pp. 256–270. Springer (1987). LNCS 249
5. d'Agostino, G., Hollenberg, M.: Logical questions concerning the μ-calculus: Interpolation, Lyndon and Los-Tarski. Journal of Symbolic Logic **65**(1), 310–332 (2000)
6. van Ditmarsch, H., French, T.: Awareness and forgetting of facts and agents. In: Proceedings of WI-IAT Workshops 2009, pp. 478–483. IEEE Press (2009)
7. van Ditmarsch, H., French, T.: Simulation and information. In: J. Broersen, J.J. Meyer (eds.) Knowledge Representation for Agents and Multi-Agent Systems, pp. 51–65. Springer (2009). LNAI 5605. Presented at LOFT 2008 and KRAMAS 2008
8. van Ditmarsch, H., French, T.: Becoming aware of propositional variables. In: M. Banerjee, A. Seth (eds.) Logic and Its Applications - 4th Indian Conference, ICLA 2011. Proceedings, pp. 204–218. Springer (2011). LNCS 6521
9. van Eijck, J., Sietsma, F.: Message-generated Kripke semantics. In: Proceedings of the Tenth International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS 2011) (2011). To appear
10. Fagin, R., Halpern, J.: Belief, awareness, and limited reasoning. Artificial Intelligence **34**(1), 39–76 (1988)
11. Fagin, R., Halpern, J., Moses, Y., Vardi, M.: Reasoning about Knowledge. MIT Press, Cambridge MA (1995)
12. Fine, K.: Propositional quantifiers in modal logic. Theoria **36(3)**, 336–346 (1970)
13. French, T.: Bisimulation quantifiers for modal logic. Ph.D. thesis, University of Western Australia (2006)
14. Grossi, D., Velázquez-Quesada, F.: *Twelve Angry Men*: A study on the fine-grain of announcements. In: X. He, J. Horty, E. Pacuit (eds.) Logic, Rationality, and Interaction. Proceedings of LORI 2009, pp. 147–160. Springer (2009). LNCS 5834
15. Halpern, J., Rego, L.: Reasoning about knowledge of unawareness. Games and Economic Behavior **67(2)**, 503–525 (2009)
16. Heifetz, A., Meier, M., Schipper, B.: Interactive unawareness. Journal of Economic Theory **130**, 78–94 (2006)
17. Hill, B.: Awareness dynamics. Journal of Philosophical Logic (2009). To appear
18. Hintikka, J.: Knowledge and Belief. Cornell University Press, Ithaca, NY (1962)
19. Hollenberg, M.: Logic and bisimulation. Ph.D. thesis, University of Utrecht (1998)
20. Levesque, H.: A logic of implicit and explicit belief. In: R. Brachman (ed.) Proceedings of AAAI, pp. 198–202. AAAI Press (1984)
21. Meyer, J.J., van der Hoek, W.: Epistemic Logic for AI and Computer Science. *Cambridge Tracts in Theoretical Computer Science 41*. Cambridge University Press, Cambridge (1995)
22. Modica, S., Rustichini, A.: Awareness and partitional information structures. Theory and Decision **37**, 107–124 (1994)
23. Modica, S., Rustichini, A.: Unawareness and partitional information structures. Games and Economic Behavior **27**, 265–298 (1999)
24. Plaza, J.: Logics of public communications. In: M. Emrich, M. Pfeifer, M. Hadzikadic, Z. Ras (eds.) Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems: Poster Session Program, pp. 201–216. Oak Ridge National Laboratory (1989)
25. Stirling, C.: The joys of bisimulation (1998). LNCS 1450
26. Velazquez-Quesada, F.: Small steps in dynamics of information. Ph.D. thesis, University of Amsterdam (2011). ILLC Dissertation Series DS-2011-02
27. Visser, A.: Bisimulations, model descriptions and propositional quantifiers (1996). Logic Group Preprint Series 161, Department of Philosophy, Utrecht University