Epistemic logic and information update

A. Baltag, H.P. van Ditmarsch, and L.S. Moss

September 18, 2006

1 Introduction

Epistemic logic investigates what agents know or believe about certain factual descriptions of the world, and about each other. This makes precise what information is (statically) available in a given system, and what the general principles are for knowledge and belief. The information in a systems may also change due to certain events, or observations by the agents, or communication between the agents. This requires an *information update*. Such information updates have been investigated in computer science as interpreted systems, and in philosophy and in artificial intelligence as belief revision. A more recent development is called *dynamic epistemic logic*. Dynamic epistemic logic is an extension of epistemic logic with dynamic modal operators for belief change (i.e., information update). Dynamic epistemic logics are the focus of our contribution, but their relation to other ways to model dynamics will also be discussed in some detail.

Observations on how epistemic states change as a result of new information have been around since the setting of the field of epistemic logic by Hintikka [27]. This focussed initially on 'puzzling' phenomena involving higher-order belief change, for a single agent. A typical example from [27] is the so-called 'Moore'-problem about the inadequacy of information updates with 'fact p is true and I don't believe that'. As this example crops up in very different settings, and as it is so crucial for a proper understanding of dynamic epistemics, we discuss its origin in some detail, as a proper historical primer to the subject area. Hintikka's 'Knowledge and belief' [27, p.64] provides a list of excellent references on the topic of Moore-sentences. This also reveals an interesting development of the notion—we explain this using the notation used in this chapter. In [42, p.78] Moore writes that if I assert a proposition φ, I *express* or *imply* that I *think* or *know* φ, in other words I express $B\varphi$ or $K\varphi$. But φ cannot be said to *mean* $B\varphi$ [42, p.77] as this would cause, by substitution, an infinite sequence $BB\varphi$, $BBB\varphi$, ad infinitum. "But thus to believe that somebody believes, that somebody believes, that somebody believes ... quite indefinitely, without *ever* coming to anything which is what is believed, is to believe nothing at all" [42, p.77]. All this is in the context of a discussion on whether moral judgements are judgements about our feelings, or about our beliefs. Moore does not state in [42] (to our knowledge) that $\varphi \wedge \neg B\varphi$ cannot be believed. In Moore's "A reply to my critics", a chapter in the 'Library of Living Philosophers' volume dedidated to him, he writes " 'I went to the pictures last Tuesday, but I don't believe that I did' is a perfectly absurd thing to say, although *what* is asserted is something which is perfectly possibly logically" [43, p.543]. The absurdity follows from the implicature 'asserting φ implies $B\varphi$' pointed out in [42]. In other words, $B(p \wedge \neg Bp)$ is 'absurd' for the example of factual information p. As far as we know, this is the first full-blown occurrence of a Moore-sentence.

Then in [44, p.204] Moore writes " 'I believe he has gone out, but he has not' is absurd. This, though absurd, is not self-contradictory; for it may quite well be true." This is an example of $\neg p \wedge Bp$. Together with [43] it also sufficiently shows, we think, that Moore really had either of the general forms $\varphi \wedge \neg B\varphi$ or $\varphi \wedge B\neg\varphi$ in mind. Note that he does not claim that $B(\varphi \wedge \neg B\varphi)$ is inconsistent ('self-contradictory') as such, but 'only' that asserting $\varphi \wedge \neg B\varphi$ implies $B\varphi$, which contradicts $\neg B\varphi$ [43, pp.204–205]. The further development of this notion, addressed in our contribution, firstly puts Moore-sentences in a *multi-agent* perspective of announcements of the form 'I say to you that: p is true and that *you* don't believe that', and, secondly, puts Moore-sentences in a *dynamic* perspective of announcements that cannot be believed after being announced. Both perspectives appear to go beyond Moore.

The area of epistemic logic appears to thrive on such puzzling phenomena, others involve unfaithful wives, unfaithful husbands, or differently coloured hats (this puzzle possibly originating in the 1950s [71, 17] is by now generally known as the 'wisemen' or 'muddy children puzzle', for a logical treatment see [46, 47]), letters with mysterious contents opened under diverse conditions, and coins being thrown for heads or tails, or turned, also including various devices as one-way mirrors, or lying, cheating and bluffing about the results. Yet another productive area for explanation and exposition of dynamic epistemic phenomena are *card deals*. An introductory example illustrating dynamics for more than one agent is given in this setting. A similar setting will then often reappear during the later formal treatment of the logics. Of course this is a mere choice for reasons of a succinct exposition, we apologize to the reader who would have preferred more diverse examples.

Three players Anne, Bill, and Cath each hold one card from a stack of three cards clubs, hearts, and spades. They know their own card, but do not know which other card is held by which other player. Also, all of the previous is common knowledge. Assume that the actual deal is that Anne holds clubs, Bill holds hearts and Cath holds spades. Now Anne announces that she does not have hearts. What was known before this announcement, and how does this knowledge change as a result of that action? Before, Cath did not know that Anne holds clubs, but afterwards she knows that Anne holds clubs. This is because Cath can reason as follows: "I have spades, so Anne must have clubs or hearts. If she says that she does not have hearts, she must therefore have clubs." Bill knows that Cath now knows Anne's card, even though he does not know himself what Anne's card is. Both before and after, players know which card they hold in their hands. Note that the *only* change that appears to have taken place is epistemic change, and that no factual change has taken place, such as cards changing hands. How do we describe such an information update in an epistemic setting? We can imagine various other actions that affect the knowledge of the players, for example, the action where Anne *shows* her clubs card to Bill, in such a way that Cath sees that Anne is doing that, but without seeing the actual card. How does that affect the knowledge of the players about each other? After that action, Cath still does not know whether Anne holds clubs or hearts. But Cath now knows that Bill knows Anne's card.

Overview The knowledge of these card players and how this changes as a result of such information updates will modelled in dynamic epistemic logic. We start with a concise introduction to *epistemic logic*. We then give an overview of the *interpreted systems* way to model dynamics. After that we pay ample attention to *public announcement logic* (as in "Anne announces that she does not have hearts"), which models one particular form of dynamics. A well-known generalization of that for more complex dynamic events (such as 'Anne shows

Figure 1: A pointed Kripke model, also known as an epistemic state, that represents Anne's knowledge of the card deal where Anne holds clubs, and where hearts is on top of the two-card stack on the table. On the right a simpler visualisation of the model, used when the accessibility relation, as here, is an equivalence relation.

clubs to Bill' above) goes under the name of *action model logic*. Our closing observations are on links between (theory) *belief revision* [1] to dynamic epistemic logic. References to the literature are mainly given near the end of each section.

2 Epistemic Logic

One agent As in the inception of epistemic logic by Hintikka, we start by modelling the knowledge of a single agent.

> *Anne draws the clubs card from a stack of three different cards clubs, hearts, and spades. Is hearts or spades the top card of the two left on the table?*

We would like to be able to evaluate system descriptions such as "Anne knows that she holds clubs" (in fact equated with Anne *seeing* that she holds clubs, as she can look at her card), "Anne does not know that hearts is on top," and "Anne considers it possible that hearts is on top." Let propositional letter $Clubs_a$ stand for the factual/atomic proposition 'the clubs card is held by Anne', and similarly $Hearts_t$ for 'the hearts card is on top' (of the two-card stack with hearts and spades). Further, we use standard propositional connectives $\wedge$ for conjunction, $\vee$ for disjunction, $\neg$ for negation, $\rightarrow$ for implication, and $\leftrightarrow$ for equivalence. A formula of the form $K\varphi$ expresses that 'Anne knows that φ'—'K' is the common name in epistemic logic (Know) for the $\Box$-type modal operator—and a formula of the form $\hat{K}\varphi$ ($\hat{K}$ is the dual of K) expresses that 'Anne considers possible that φ'. The informal descriptions are then formalized as $K\,Clubs_a$, $\neg K\,Hearts_t$, and $\hat{K}\,Hearts_t$, respectively. There is no generally accepted notation for 'considers possible that'. The 'hat' in the notation $\hat{K}\varphi$—the notation we will keep using—is reminiscent of the diamond in $\Diamond\varphi$. Other notations for $\hat{K}\varphi$ are $M\varphi$, $\langle K\rangle$, and $k\varphi$.

We interpret such formulas on Kripke models. Kripke models consist of a domain of abstract states, an accessibility relation between those states, and a valuation of propositional letters in a given state. Also, one reasons from the perspective of an actual state in the model. In our example the two states correspond to two deals of cards, namely the deal $\clubsuit\heartsuit\spadesuit$ where Anne holds the clubs card, and the hearts card is on top of the spades card (both facedown) on the table, and the deal $\clubsuit\heartsuit\spadesuit$ where Anne also holds clubs but where hearts is on the bottom of the two-card stack. These 'names' for states are of course suggestive, as they express which facts are true in which state. The binary relation of accessibility between states expresses what the player knows about the facts. Assume that deal $\clubsuit\heartsuit\spadesuit$ is actually the case. Anne considers that possible, so that the pair $(\clubsuit\heartsuit\spadesuit, \clubsuit\heartsuit\spadesuit)$ is in the accessibility relation, but she also considers it possible that hearts is at the bottom of the stack, so that $(\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit)$ is also in the accessibility relation. Similarly, reasoning from the hypothetical case that $\clubsuit\spadesuit\heartsuit$ were the actual deal, $(\clubsuit\spadesuit\heartsuit, \clubsuit\spadesuit\heartsuit)$ and $(\clubsuit\spadesuit\heartsuit, \clubsuit\heartsuit\spadesuit)$ are in the accessibility relation. The resulting pointed Kripke model (or epistemic state, or information state) is depicted in Figure 1. It is formally an epistemic state $(D, \clubsuit\heartsuit\spadesuit)$ where the model $D = \langle S, R, V\rangle$ consists of a domain $S = \{\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit\}$, accessibility relation R with $R = \{(\clubsuit\heartsuit\spadesuit, \clubsuit\heartsuit\spadesuit), (\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit), \dots\}$, and valuation V such that $V(\mathit{Clubs}_a) = \{\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit\}$ and $V(\mathit{Hearts}_t) = \{\clubsuit\heartsuit\spadesuit\}$ (we identify a fact with the subset of the domain where it is true).

A proposition is known in an epistemic state if and only if it is true in all accessible states. For example, $D, \clubsuit\heartsuit\spadesuit \models K\mathit{Clubs}_a$ (Anne knows that she holds clubs in the actual state), because for all states s, if $(\clubsuit\heartsuit\spadesuit, s) \in R$ then $D, s \models \mathit{Clubs}_a$. This is true because the only states that are accessible from $\clubsuit\heartsuit\spadesuit$ are $\clubsuit\heartsuit\spadesuit$ itself and $\clubsuit\spadesuit\heartsuit$—we have $R(\clubsuit\heartsuit\spadesuit, \clubsuit\heartsuit\spadesuit)$ and $R(\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit)$—and both $D, \clubsuit\heartsuit\spadesuit \models \mathit{Clubs}_a$ and $D, \clubsuit\spadesuit\heartsuit \models \mathit{Clubs}_a$. Finally, $D, \clubsuit\heartsuit\spadesuit \models \mathit{Clubs}_a$ because $\clubsuit\heartsuit\spadesuit \in V(\mathit{Clubs}_a) = \{\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit\}$, and, similarly, $D, \clubsuit\spadesuit\heartsuit \models \mathit{Clubs}_a$ because $\clubsuit\spadesuit\heartsuit \in V(\mathit{Clubs}_a) = \{\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit\}$.

If one assumes certain properties of knowledge, the accessibility relation for the agent is an equivalence relation. The (often contested) properties are that 'what you know is true', which is formalized by the schema $K\varphi \to \varphi$; that 'you are aware of your knowledge', which is formalized by the schema $K\varphi \to KK\varphi$, and that 'you are aware of your ignorance', which is formalized by the schema $\neg K\varphi \to K\neg K\varphi$. Without $K\varphi \to \varphi$ but *with*, instead, the schema $K\varphi \to \hat{K}\varphi$ the operator models introspective belief instead of knowledge, in that case we write B instead of K. For equivalence relations, we will occassionally write $\sim$ instead of R (namely when discussing interpreted systems), and use a simpler visualization wherein we only visually link states that are in the same class, as in Figure 2 (which will be explained below).

More agents Many features of formal dynamics can be presented based on the single-agent situation. For example, the action of Anne picking up from the table the card that has been dealt to her, is a significantly complex epistemic action. But a proper and more interesting perspective is that of the multi-agent situation. This is because players may now have knowledge about each others' knowledge. Even for a single fact the Kripke models representing such knowledge can become arbitrarily complex. To distinguish different knowledge operators and corresponding accessibility relations for different agents, we label them.

Consider three players Anne, Bill, and Cath (a, b, c), each of who blindly draws a card from the stack of three cards clubs, hearts, and spades. Assume that the actual deal is that Anne draws clubs, Bill hearts, and Cath spades. This is represented by state $\clubsuit\heartsuit\spadesuit$. Propositional letters q_a stand for 'agent a holding card q.' We can now describe that "Bill considers it possible that Anne has spades but actually Anne has clubs" as $\hat{K}_b\mathit{Spades}_a \wedge \mathit{Clubs}_a$; and "Anne knows that Bill knows that Cath knows her own card" as $K_aK_b(K_c\mathit{Clubs}_c \vee K_c\mathit{Hearts}_c \vee K_c\mathit{Spades}_c)$. Figure 2 depicts the corresponding epistemic state.

The language, structures, and semantics of *multi-agent epistemic logic*, relative to a set of agents A and a set of atoms P as background parameters, are formally defined as follows.

Definition 1 (Language)

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi$$

where $p \in P$, and $a \in A$.

Other connectives are defined by notational abbreviation.

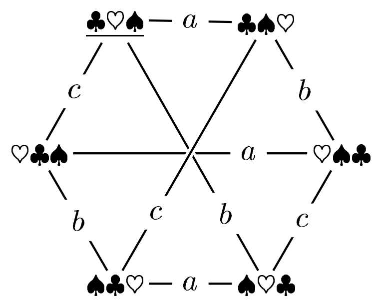

Figure 2: The epistemic state (*Hexa*, ♣♡♠) for the card deal where Anne holds clubs, Bill holds hearts, and Cath holds spades. For example, Anne cannot distinguish ♣♡♠ from ♣♠♡ as she holds clubs in both deals. But Bill cannot distinguish ♣♡♠ from ♠♡♣ as he holds hearts in both. As all accessibility relations are equivalence relations, we only have to link states in the same class by labelled (undirected) arcs.

Definition 2 (Structures) An *epistemic model* $M = \langle S, \sim, V \rangle$ consists of a *domain* S of (factual) *states* (or 'worlds'), *accessibility* (function) $R\ : A \rightarrow \mathcal{P}(S \times S)$, and a *valuation* (function) $V : P \rightarrow \mathcal{P}(S)$. For $s \in S$, (M, s) is an *epistemic state*. For valuation $V(p)$ we also write V_p and for accessibility relation $R(a)$ we also write R_a.

Definition 3 (Semantics)

$$\begin{array}{lll} M, s \models p & \text{iff} & s \in V_p \\ M, s \models \neg\varphi & \text{iff} & M, s \not\models \varphi \\ M, s \models \varphi \wedge \psi & \text{iff} & M, s \models \varphi \text{ and } M, s \models \psi \\ M, s \models K_a\varphi & \text{iff} & \text{for all } \ t \in S : (s,t) \in R_a \text{ implies } M, t \models \varphi \end{array}$$

Formula φ is valid on model M, notation $M \models \varphi$, if and only if for all states s in the domain of M: $M, s \models \varphi$. Formula φ is valid, notation $\models \varphi$, if and only if for all models M (of the class of models for the given parameters of A and P): $M \models \varphi$. These validity notions will be similarly defined for all expansions of the language, in the continuation.

Notes Hintikka is broadly acknowledged as the father of modern epistemic logic, and his 'Knowledge and Belief' [27] (republished in 2005 by King's College, London) given as the principal historical reference. Hintikka himself thinks that von Wright [73] deserves these credits. Modern epistemic logic started to flourish after modal logic (with its roots in Aristotle) was formalised and given a possible world semantics. It is hard to track down the exact origins of this semantics, but it is widely known as Kripke semantics, after Kripke, who devoted a number of early papers to the semantics of modal logic [32]. A contemporary and thorough reference for modal logic is the monograph [10]. From the late 1970s, epistemic logic became subject of study or applied in the areas of artificial intelligence (as in (R.C.) Moore's early work [45] on reasoning about actions and knowledge), philosophy (as in Hintikka's [28]), and game theory (e.g. Aumann [5]). In the 1980s, computer scientists became interested in epistemic logic. In fact, the field matured a lot by a large stream of publications around Fagin, Halpern, Moses and Vardi. Their important textbook '*Reasoning about Knowledge*' [15] which appeared in 1995, is in fact a survey of many papers co-authored by (subsets of) them over a period of more than ten years. We refer to [15] for more references. It counts as a standard reference in epistemic logic. Another standard reference is [41].

Common knowledge Well-known epistemic operators for a *group* of agents are 'general knowledge' and 'common knowledge'. We introduce the concepts by way of an example.

In the epistemic state $(Hexa, ♣♡♠)$ of Figure 2 both Anne and Bill know that the deal of cards is *not* ♠♣♡: both $K_a\neg(Spades_a \wedge Clubs_b \wedge Hearts_c)$ and $K_b\neg(Spades_a \wedge Clubs_b \wedge Hearts_c)$ are true. If a group of agents all know that φ, we say that φ is general knowledge. The modal operator for general knowledge of a group B is E_B. For an arbitrary subset $B \subseteq A$ of the set of agents A, we define $E_B\varphi := \bigwedge_{a\in B} K_a\varphi$. So in this case we have that $E_{ab}\neg(Spades_a \wedge Clubs_b \wedge Hearts_c)$—we abuse de language and write E_{ab} instead of $E_{\{a,b\}}$. Now even though φ may be generally known, that does not imply that agents know *about each other* that they know φ. For example, Bill, who has hearts, considers it possible that Anne has spades instead of clubs. In that case, Anne considers it possible that the card deal is ♠♣♡. So $\hat{K}_a\hat{K}_b(Spades_a \wedge Clubs_b \wedge Hearts_c)$ is true, and therefore $K_bK_a\neg(Spades_a \wedge Clubs_b \wedge Hearts_c)$ is false in $(Hexa, ♣♡♠)$.

One can construct formulas that are true to some extent $K_aK_bK_cK_aK_aK_b\varphi$ in some epistemic state but no longer if one adds one more operator at the start: $K_bK_aK_bK_cK_aK_aK_b\varphi$ is false—and where just the difference between the two is essential to model a given problem properly. A formula φ is *common knowledge* for a group B, notation $C_B\varphi$, if it holds for arbitrary long stacks of individual knowledge operators for individuals in that group. For example, if $B = \{a, b, c\}$, we get something (involving an enumeration of all finite stacks of knowledge operators) like $C_{abc}\varphi := \varphi \wedge K_a\varphi \wedge K_b\varphi \wedge K_c\varphi \wedge K_aK_a\varphi \wedge K_aK_b\varphi \wedge K_aK_c\varphi \wedge \ldots K_aK_aK_a\varphi \ldots$ Alternatively, we may see common knowledge as the conjunction of arbitrarily many applications of general knowledge: $C_B\varphi := \varphi \wedge E_B\varphi \wedge E_BE_B\varphi \wedge ....$ Such infinitary definitions are often undesirable for various reasons. Common knowledge C_B is typically added as a primitive operator to the language, whereas general knowledge is typically defined (for a finite set of agents) by notational abbreviation, as above. Common knowledge can be given its intended meaning in the semantics without changing the Kripke structures used to interpret individual knowledge, namely by an operation on the accessibility relations for the individual agents in the group. Something is common knowledge in the actual state, if it is true in all states in the transitive closure of the union of the accessibility relations for all agents in the group. The interaction between common and individual knowledge can then be specified in valid principles for the logic. Valid principles involving common knowledge are $C_B(\varphi \to \psi) \to (C_B\varphi \to C_B\psi)$ (distribution of common knowledge over implication), $C_B\varphi \to E_B(\varphi \wedge C_B\varphi)$ (use of common knowledge), and "from $\varphi \to E_B(\varphi \wedge \psi)$ infer $(\varphi \to C_B\psi)$ (induction rule).

Definition 4 (Language and semantics) Add an inductive clause $C_B\varphi$ to the definition of the language, where $B \subseteq A$. For the semantics, add clause:

$$M, s \models C_B\varphi \quad \text{iff} \quad \text{for all } \ t \in S : R_B^+(s,t) \text{ implies } M, t \models \varphi$$

where $R_B^+ \equiv (\bigcup_{a\in B} R_a)^+$.

Alternatively said, $C_B\varphi$ is true in an epistemic state (M, s) if φ is true in any state s_m that can be reached by a finite path of states $s_1, \ldots, s_m$ such that, for not necessarily different agents $a, b, c \in B$: $R_a(s_1, s_2)$, $R_b(s_2, s_3)$, and ..., and $R_c(s_{m-1}, s_m)$. 'Reachability by a finite path of non-zero length' is the same as 'being in the transitive closure'. There are two different but both wide-spread semantics for common knowledge. The one above is popular for logics of belief, and in philosophical circles. In computer science one often prefers to take the *reflexive* transitive closure $R_B^* \equiv (\bigcup_{a \in B} R_a)^*$ to interpret common knowledge (which means 'reachability by a finite path'). If all individual accessibility relations are equivalence relations, R_B^+ is also an equivalence relation, and in that case R_B^+ equals R_B^* anyway. Common knowledge for the entire group A of agents is called *public knowledge*.

In the model *Hexa*, access for any subgroup of two players, or for all three, is the entire model. For such groups B, $C_B\varphi$ is true in an epistemic state (Hexa, t) iff φ is valid on the model *Hexa*. For example, we have that "It is public knowledge that Anne knows her card", formally $\mathit{Hexa} \models C_{abc}(K_a \mathit{Clubs}_a \vee K_a \mathit{Hearts}_a \vee K_a \mathit{Spades}_a)$, and (possibly surprisingly) "Anne and Bill share the same knowledge as Bill and Cath," which is formally $\mathit{Hexa} \models C_{ab}\varphi \rightarrow C_{bc}\varphi$.

Notes In Lewis' 'Convention' [34] the notion of common knowledge was informally discussed. In the area of game theory Aumann's [5] gives one of the first formalisations of common knowledge. McCarthy formalises common knowledge in a rather off-hand way when solving a well-known epistemic riddle, the Sum and Product-riddle [40] (although at the time it was unknown to him that this riddle originated with the Dutch topologist Freudenthal [16]) as an abstract means towards solving this riddle. This work dates from the seventies but was only published later in a collection of McCarthy's work that appeared in 1990.

From two standard references [15, 41] to epistemic logic, Fagin et al. [15] defines common knowledge by transitive closure, whereas Meyer and Van der Hoek [41] define it by reflexive transitive closure. There is a recent resurging interest in variants of the notion, e.g. Artemov's evidence-based common knowledge, also known as justified common knowledge [2].

3 Interpreted Systems and Temporal Epistemic Logic

A general framework involving information change as a feature of *interpreted systems* was developed by Halpern and collaborators in the 1990s [15]. Central to this approach is the notion of the global state of a system. Given a number of agents or processors, each of which has a local state (such as 'holding clubs' for agent Anne), a global state is a list of all the local states of the agents involved in the system plus a state of the environment. The last represents actions, observations, and communications, possibly outside the sphere of influence of the agents. An example global state is $(\clubsuit\heartsuit\spadesuit, \emptyset)$ wherein Anne has local state $\clubsuit$, i.e., she holds clubs, Bill local state $\heartsuit$, and Cath local state $\spadesuit$, and where 'nothing happened so far in the environment,' represented by a value $\emptyset$. It is assumed that agents know their local state, in other words, that they can distinguish global states from one another wherein they have the same local state. This induces an equivalence relation among global states that the reader will obviously recognize as an accessibility relation. Another crucial concept in interpreted systems is that of a *run*: a run is a (typically infinite) sequence of global states. For example, when Anne says that she does not have hearts, this corresponds to a transition from global state $(\clubsuit\heartsuit\spadesuit, \emptyset)$ to global state $(\clubsuit\heartsuit\spadesuit, \mathsf{nohearts})$. Atomic propositions may also be introduced to describe facts. For example, not surprisingly, one may require an atom Hearts_a to be true in both global state $(\clubsuit\heartsuit\spadesuit, \emptyset)$ and in global state $(\clubsuit\heartsuit\spadesuit, \mathsf{nohearts})$.

Formally, an *interpreted system* $\mathcal{I}$ is a pair $(\mathcal{G}, \mathcal{R})$ consisting of a set of global states $\mathcal{G}$ and a set of runs $\mathcal{R}$ relating those states. A *global state* $g \in \mathcal{G}$ is a tuple consisting of local states g_a for each agent and a state g_ϵ of the environment. A *run* $r \in \mathcal{R}$ is a sequence of global

$$\begin{tikzcd}
\underline{(\clubsuit\heartsuit\spadesuit, u\heartsuit)} \arrow[r, dash] \arrow[d, "Hearts_t"] & (\clubsuit\spadesuit\heartsuit, u\heartsuit) \arrow[d, "\neg Hearts_t"] \\
(\clubsuit\heartsuit\spadesuit, y\heartsuit) \arrow[r, dash] & (\clubsuit\spadesuit\heartsuit, n\heartsuit)
\end{tikzcd}$$

Figure 3: Anne holds clubs, hearts is on top of spades on the two-card stack on the table, and Anne does not know (in the underlined, actual global state) if it is. The two visualized runs reveal which card is on top.

states. The m-th global state occurring in a run r is referred to as $r(m)$, and the local state for agent a in a global state $r(m)$ is written as $r_a(m)$.

A *point* (r, m) is a pair consisting of a run and a point in time m—this is the proper abstract domain object when defining epistemic models for interpreted systems. In an interpreted system, agents can distinguish global states from one another iff they have the same local state in both, which induces (for an indistinguishability relation that is an equivalence we choose to write $\sim$ instead of R)

$$(r, m) \sim_a (r', m') \text{ iff } r(m) \sim_a r'(m') \text{ iff } r_a(m) = r'_a(m')$$

With the obvious valuation for local and environmental state values, that defines an epistemic model. For convenience we keep writing $\mathcal{I}$ for that. Given an actual point (r', m'), we thus get an epistemic state $(\mathcal{I}, (r', m'))$. Epistemic and (LTL) temporal (next) operators have the interpretation

$$\begin{array}{lll}
\mathcal{I}, (r, m) \models X\varphi & \text{iff} & \mathcal{I}, (r, m+1) \models \varphi \\
\mathcal{I}, (r, m) \models K_a\varphi & \text{iff} & \text{for all } (r', m') : (r, m) \sim_a (r', m') \text{ implies } \mathcal{I}, (r', m') \models \varphi
\end{array}$$

It will be clear that subject to some proper translation (see e.g. [38]) interpreted systems correspond to some subclass of the $S5$ models: all relations are equivalence relations, but the interaction between agents is even more than that. The relation between Kripke models and interpreted systems is not entirely trivial, partly because worlds or states in Kripke models are abstract entities that may represent the same set of local states. The main difference between the treatment of dynamics in interpreted systems and that in dynamic epistemics is that in the former this is encoded in the state of the environment, whereas in the latter it emerges from the relation of a state (i.e., an abstract state in a Kripke model) to other states.

Example For a simple example, consider the single agent example in the previous section, and in Figure 1, wherein a single agent Anne holds clubs, and the hearts card is on top of the spades card (both facedown) on the table. She may now be informed about the card on top of the stack. This is represented by the interpreted system depicted in Figure 3. It consists of four global states. The card Anne holds represents her local state. The other cards are (in this case, unlike in the three-agent card deal) part of the environment. The state of the environment is represented by which of the two cards is on top, and by an ‘observation’ state variable *obs* that can have three values $u\heartsuit$, $y\heartsuit$, and $n\heartsuit$, corresponding to the state before the announcement which card is on top, the state resulting from the announcement that hearts is on top, and the other state resulting from the announcement that it is at the bottom.[1] The valuation V is now such that $V(Clubs_a) = \{(\clubsuit\heartsuit\spadesuit, u\heartsuit), (\clubsuit\spadesuit\heartsuit, u\heartsuit), (\clubsuit\heartsuit\spadesuit, y\heartsuit), (\clubsuit\spadesuit\heartsuit, n\heartsuit)\}$, and $V(Hearts_t) = \{(\clubsuit\heartsuit\spadesuit, u\heartsuit), (\clubsuit\heartsuit\spadesuit, y\heartsuit)\}$. The system consists of two runs, one from $(\clubsuit\heartsuit\spadesuit, u\heartsuit)$ to $(\clubsuit\heartsuit\spadesuit, y\heartsuit)$ (optionally extended with an infinite number of idle transitions), and the other run from $(\clubsuit\spadesuit\heartsuit, u\heartsuit)$ to $(\clubsuit\spadesuit\heartsuit, n\heartsuit)$. One can now compute that in the actual state $(\clubsuit\heartsuit\spadesuit, u\heartsuit)$ it is true that $\neg K_a Hearts_t$, but in state $(\clubsuit\heartsuit\spadesuit, y\heartsuit)$ she has learnt that hearts is on top: $K_a \neg Hearts_t$ is now true. Or, for another example, that in the actual state $XK_a Hearts_t$. How the treatment of announcements in interpreted systems relates to public announcement logic, will be made precise at the end of the following section.

Interpreted systems have been highly successful as an abstract architecture for multi-agent systems, where agents are either human operators or computer processors, and where the assumption that an agent 'knows its own state' is a realistic simplification. For that reason they can be said to model interaction between *ideal agents*. This assumption is also implicitly applied when modelling perfectly rational agents as in game theory and economics. Also, given that all the dynamics is *explicitly* specified in the runs through the system, it combines well with temporal epistemic logics (LTL, CTL) wherein dynamics is *implicitly* specified by referring to an underlying structure wherein such a change makes information sense. Temporal epistemic logics have been fairly successful. Their computational properties are well-known and proof tools have been developed. See, for example, [61, 13, 25]. The work of Fagin et al. [15] also generated lots of complexity results on knowledge and time, we also mention the work of van der Meyden in this respect, e.g. [60, 61].

Their are two rather pointed formal differences between the temporal epistemic approach and the dynamic epistemic approach.

Closed versus open systems First, the temporal epistemic description takes as models systems together with their whole (deterministic) history and future development, in the shape of 'runs'. As such, it can be easily applied to 'closed' systems, in which all the possible developments are fixed in advance, where there are no accidents, surprises or new interactions with the outside world, and thus the future is fully determined. Moreover, in practice the approach is more applicable to closed systems having a *small* number of possible moves: that's the only ones for which it is feasible to draw the transition graph of the full history.

In contrast, the dynamic epistemic approach can also be applied to 'open' systems. This is for example the case with epistemic protocols which can be modified or adapted at any future time according to new needs, or which can interact with an unpredictable environment. But it is also applicable to closed systems in which the number of possible different changes is large or indefinite.

There are two analogies here to be made. The first is with open-versus-closed-system paradigms in programming. People in concurrency are usually interested in open systems. The program might be run in many different contexts, in the presence of many other programs, etc. More recently (in the context of mobile computation), people have looked at approaches that allow programs to be changed at any time inside the same logical frame. The temporal logic approach is not fit for this, since it assumes the full current program to be fixed and given as 'the background model'. That is why people in this area have used totally different kinds of formalisms, mainly process algebraic, such as π-calculus. In contrast to that, dynamic

[1] An environmental local state variable with options 'unknown', 'true' and 'false' to model an announcement formula is used in work in progress by Ji Ruan.

epistemic logics are interesting in that, although based on a modal logic, which is not an algebraic kind of formalism, they are able to express changes in an open system through the semantic trick of changing the models themselves, via 'epistemic updates'.

The second analogy is with *game theory*. The temporal approach is like the description of a game through explicitly giving its full extensional form: the graph of all possible plays. For instance, chess (in this approach) is defined as the set of all possible chess plays. This set is finite but huge! But there is of course another (more commonly used) way to describe a game: by giving only the 'rules of the game' (which type of actions are allowed in which type of situations), together maybe with an 'initial state' (or set of states) and some 'winning rules'. This is a much more economic and insightful way to describe a game. Of course, once this is given, one could draw the game in extensional form as the set of all plays, if one is given enough computational power... If we neglect the winning rules, the dynamic epistemic approach can naturally describe epistemic games in precisely this way: one gives an epistemic Kripke model of 'initial states' and also an epistemic Kripke model or other semantically precise description of possible 'epistemic actions', including preconditions that tell us on which type of states a given action can be applied. Then one can play the game, by repeatedly updating the state model with the action model. A 'full play' or 'run' of the game is obtained when we reach a state (at the end of many updates) on which no action (in our given action model) can be applied.

Information change description The second difference between the interpreted systems and the dynamic epistemics approach simply concerns the ability to model and classify various 'types' or 'patterns' of information change, or information exchange, such as public announcements, private announcements, game announcements etc. The dynamic epistemic approach obviously has this in-built ability, while the temporal approach doesn't have it, at least not in a direct, usable manner. In the temporal approach, one can only say what is true 'before' and 'after' a given action, and thus only implicitly get some information about the type of the action itself, through its input-output behaviour. Moreover, this information is *not* enough to isolate the type of the action, since it only gives us the *local* input-output behaviour of a given action; and different actions may behave identically in one local context, but differ in general. For instance in the two players and two cards case, in an epistemic state in which the fact that the card deal is $\clubsuit\heartsuit\spadesuit$ is common knowledge, public announcement of that fact will have the same input-output description as a 'skip' action corresponding to 'nothing happens'. But in the epistemic state where the cards were dealt but not seen, or the subsequent one where all players only know their own card, this fact was not common knowledge and its public announcement will in that case induce an informative (i.e. non-skip) transition.

4 Dynamic Epistemic Logic

We now move on to a different form of dynamics of knowledge. Starting from the perspective of epistemic logic, knowledge and belief change can also be modelled by expanding the logic with dynamic modal operators to express such change. A simple form of dynamics is that caused by so-called public announcements. It is simple from the perspective of change, not particularly simple seen as an extension of epistemic logic. Public announcement logic is discussed in Subsection 4.1. The relation between interpreted systems and public announcement logic will be readdressed at the end of that section. More complex dynamics are treated in Subsection

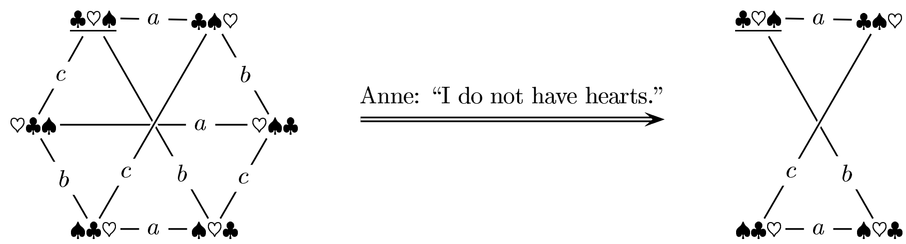

Figure 4: Given epistemic state $(\mathit{Hexa}, \clubsuit\heartsuit\spadesuit)$ where Anne hold clubs, Bill holds hearts, and Cath holds spades, the effect of Anne saying that she does not have hearts.

4.2.

4.1 Public Announcements

Assume our usual deal $\clubsuit\heartsuit\spadesuit$ of cards where Anne holds clubs, Bill hearts, and Cath spades. Anne now says ('announces') that she does not have the hearts card. Therefore she makes public to all three players that all deals where Hearts_a is true can be eliminated from consideration: everybody knows that everybody else eliminates those deals, etc. They can therefore be *publicly* eliminated. This results in a restriction of the model *Hexa* as depicted in Figure 4. The public announcement "I do not have hearts" can be seen as an 'epistemic program' with 'precondition' $\neg\mathit{Hearts}_a$, and it is interpreted as an 'epistemic state transformer' of the original epistemic state, exactly as a program in *dynamic* modal logic: given some program π, in dynamic logic $[\pi]\psi$ means that after every execution of π (state transformation induced by π), formula ψ holds. For announcements we want the form $[\varphi]\psi$ to mean that after (every) announcement of φ, formula ψ holds. The effect of such a public announcement of φ is the restriction of the epistemic state to all worlds where φ holds. So, 'announce φ' can indeed be seen as an epistemic state transformer, with a corresponding dynamic modal operator $[\varphi]$.

We *appear* to be moving away slightly from the standard paradigm of modal logic. So far, the accessibility relations were between states in a given model underlying an epistemic state. But all of a sudden, we are confronted with an accessibility relation between epistemic states as well. "I do not have hearts" induces a(n) (epistemic) state transition such that the pair of epistemic states in Figure 4 is in that relation. The epistemic states take the role of the points or worlds in a seemingly underspecified domain of 'all possible epistemic states'. By lifting accessibility between points in the original epistemic state to accessibility between epistemic states, we can get the dynamic and epistemic accessibility relations 'on the same level' again, and see this as an 'ordinary structure' on which to interpret a perfectly ordinary multimodal logic. (There is also a clear relation here with interpreted systems, which will be discussed in Subsection 4.1, later.) A crucial point is, that this 'higher-order structure' is induced by the initial epistemic state and the actions that can be executed there, and not the other way round. So it is standard modal logic after all.

Anne's announcement "I do not have hearts" is a simple epistemic action in various respects. *It is public.* A 'private' event would be when she learns that Bill has hearts without Bill or Cath noticing anything. This required a more complex action description. *It is truthful.* She could also have said "I do not have clubs." She would then be lying, but, e.g., may have reason to expect that Bill and Cath believe her. This would also require a more complex action description. *It is deterministic.* In other words, it is a state transformer. A non-deterministic action would be that Anne whispers into Bill's ear a card she does not hold, on Bill's request for that information. This action would have two different executions: "I do not have hearts", and "I do not have spades." Such more complex actions can be modelled in the action model logic presented in the next subsection.

Definition 5 (Language and semantics) Add an inductive clause $[\varphi]\varphi$ (BNF-format) to the definition of the language. For the semantics, add clause:

$$M, s \models [\varphi]\psi \quad \text{iff} \quad M, s \models \varphi \text{ implies } M|\varphi, s \models \psi$$

where $M|\varphi = \langle S', R', V' \rangle$ is defined as

$$\begin{array}{lcl} S' & \equiv & \{s' \in S \mid M, s' \models \varphi\} \\ R'_a & \equiv & R_a \cap (S' \times S') \\ V'_p & \equiv & V_p \cap S' \end{array}$$

In other words: the model $M|\varphi$ is the model M restricted to all the states where φ holds, including access between states (a submodel restriction in the standard meaning of that term). The interpretation of the dual $\langle\varphi\rangle$ of $[\varphi]$ will be obvious: $M, s \models \langle\varphi\rangle\psi$ if and only if $M, s \models \varphi$ and $M|\varphi, s \models \psi$. A proof system for this logic originates with and is proved sound and complete in [8], with precursors (namely completeness results for the logic *with* announcements but *without* common knowledge) in [49] and [21]. There are some alternative semantics for public announcements. Gerbrandy [19, 20] and Kooi [31] propose a different semantics for announcements in a setting possibly more suitable for 'belief'. The execution of such announcements is not conditional to the truth of the announced formula. Yet another semantics in a setting where introspective agents *remain* introspective after announcements has recently been proposed by Steiner [55].

Example After Anne's announcement that she does not have hearts, Cath knows that Anne has clubs (see Figure 4). We can verify this with a semantic computation as follows:

> In order to prove that $\mathit{Hexa}, \clubsuit\heartsuit\spadesuit \models [\neg \mathit{Hearts}_a]K_c \mathit{Clubs}_a$, we have to show that $\mathit{Hexa}, \clubsuit\heartsuit\spadesuit \models \neg\mathit{Hearts}_a$ implies $\mathit{Hexa}|\neg\mathit{Hearts}_a, \clubsuit\heartsuit\spadesuit \models K_c\mathit{Clubs}_a$. As it is indeed the case that $\mathit{Hexa}, \clubsuit\heartsuit\spadesuit \models \neg\mathit{Hearts}_a$, it remains to show that $\mathit{Hexa}|\neg\mathit{Hearts}_a, \clubsuit\heartsuit\spadesuit \models K_c\mathit{Clubs}_a$. The set of states that is equivalent to $\clubsuit\heartsuit\spadesuit$ for Cath, is the singleton set $\{\clubsuit\heartsuit\spadesuit\}$. So it is sufficient to show that $\mathit{Hexa}|\neg\mathit{Hearts}_a, \clubsuit\heartsuit\spadesuit \models \mathit{Clubs}_a$, which follows trivially from $\clubsuit\heartsuit\spadesuit \in V_{\mathit{Clubs}_a} = \{\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit\}$.

To give the reader a feel for this public announcement logic we give some of its valid principles.

Announcements are functional If an announcement can be executed, there is only one way to do it:

$$\langle\varphi\rangle\psi \to [\varphi]\psi \text{ is valid}$$

This is a simple consequence of the functionality of the state transition semantics for announcement. One might also say (from a program perspective) that announcements are *deterministic.*

Sequence of announcements A sequence of two announcements can always be replaced by a single, more complex announcement. Instead of first saying 'φ' and then saying 'ψ' you may as well have said for the first time 'φ and after that ψ'. This is expressed in

$$[\varphi \wedge [\varphi]\psi]\chi \text{ is equivalent to } [\varphi][\psi]\chi$$

This is useful when analysing announcements that are made with specific intentions; or, more generally, conversational implicatures à la Grice. Intentions can be postconditions ψ that should hold after the announcement. So the (truthful) announcement of φ with the intention of achieving ψ corresponds to the announcement $\varphi \wedge [\varphi]\psi$.

Being overtaken by one's intentions When such intentions are publicly known, this may land you into trouble. For an example in epistemic state ($Hexa, \clubsuit\heartsuit\spadesuit$), consider that:

> *An outsider says: "The deal of cards is neither $\spadesuit\clubsuit\heartsuit$ nor $\heartsuit\spadesuit\clubsuit$."*

This is formalized as $\neg(Spades_a \wedge Clubs_b \wedge Hearts_c) \wedge \neg(Hearts_a \wedge Spades_b \wedge Clubs_c)$. Abbreviate this announcement as one. See Figure 5 for the result of the announcement of one. None of the three players Anne, Bill, and Cath know the card deal as a result of this announcement! Now imagine that the players know (publicly) that the outsider made the announcement one in the happy knowledge of not revealing the deal of cards to anyone! For example, he might have been boasting about his logical prowess and the players might inadvertently have become aware of that. In other words, it becomes known that the announcement one was made *with the intention of keeping the players ignorant of the card deal.* Ignorance of the card deal (whatever the deal may have been) can be described as some long formula that is a conjunction of eighteen parts and that starts as $\neg K_a(Clubs_a \wedge Hearts_b \wedge Spades_c) \wedge \neg K_b(Clubs_a \wedge Hearts_b \wedge Spades_c) \wedge \neg K_c(Clubs_a \wedge Hearts_b \wedge Spades_c) \wedge \dots$. This formula is abbreviated as two, and this intention two can be seen as a subsequent announcement, as it is (publicly) known. In the model $(Hexa|\mathsf{one})$ resulting from the announcement of one, the formula two is *false* in all states that are a singleton equivalence class for at least one player, and *true* anywhere else. So it is only true in state $\clubsuit\heartsuit\spadesuit$. For the result of the announcement of two, see again Figure 5. In the epistemic state resulting from two all players know the card deal. So in that epistemic state two is false. What does it mean that the players have become aware of the intention of the outsider? This means that although the outsider was actually saying one, he really meant 'one, and after that two', in other words, he was saying $\mathsf{one} \wedge [\mathsf{one}]\mathsf{two}$. Unfortunately, $Hexa, \clubsuit\heartsuit\spadesuit \models [\mathsf{one} \wedge [\mathsf{one}]\mathsf{two}]\neg\mathsf{two}$. The outsider could have kept the card deal a secret, but by intending to keep it a secret—and the assumption that this intention is public knowledge—he was, after all, actually revealing the secret.

Announcement and knowledge Because $[\varphi]$ is interpreted as a partial function, $[\varphi]K_a\psi$ is not equivalent to $K_a[\varphi]\psi$. A simple counterexample is the following: in ($Hexa, \clubsuit\heartsuit\spadesuit$) it is true that after 'every' announcement of 'Anne holds hearts', Cath knows that Anne holds clubs. This is because that announcement cannot take place in that epistemic state. In other words, we have

$$Hexa, \clubsuit\heartsuit\spadesuit \models [Hearts_a]K_c Clubs_a$$

because of the peculiarity that all postconditions of $\Box$-operators (i.e., formulas bound by $\Box$-operators) are true when there are no accessible states. On the other hand, it is false that

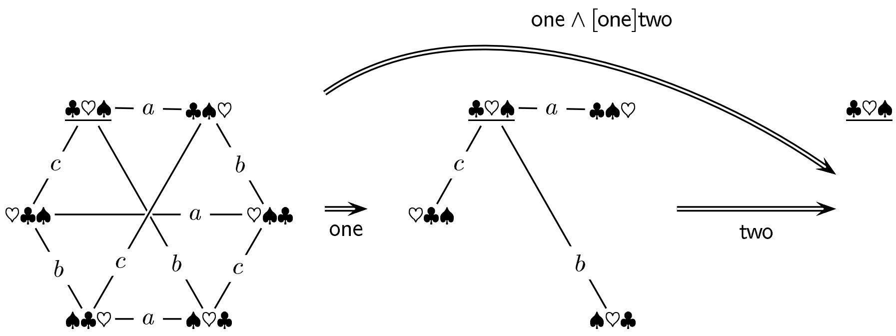

Figure 5: A sequence of two announcements can be replaced by a single announcement.

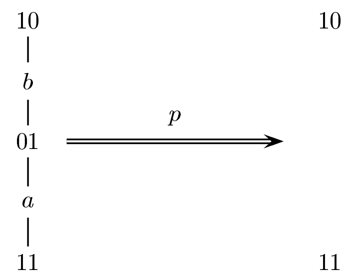

Figure 6: A state transition illustrating what a and b commonly know before and after the announcement of p.

Cath knows that after the announcement of Anne that she holds the hearts card (which she can imagine to take place), Cath knows that Anne holds the clubs card. On the contrary: Cath then knows that Anne holds hearts! So we have

$$\mathit{Hexa}, \clubsuit\heartsuit\spadesuit \not\models K_c[\mathit{Hearts}_a]\mathit{Clubs}_a$$

If we make $[\varphi]K_a\psi$ conditional to the truth of the announcement, an equivalence indeed holds:

$$[\varphi]K_a\psi \text{ is equivalent to } \varphi \rightarrow K_a[\varphi]\psi$$

Announcement and common knowledge If we restrict ourselves to the logic of announcements *without* common knowledge, every formula is logically equivalent to one in the logic without announcements. But for the logic of announcements *with* common knowledge, this is no longer the case. The principle describing the interaction between common knowledge and announcement is rather involved. The straightforward generalization of the principle $[\varphi]K_a\psi \leftrightarrow (\varphi \rightarrow K_a[\varphi]\psi)$ relating announcement and individual knowledge *would* be $[\varphi]C_A\psi \leftrightarrow (\varphi \rightarrow C_A[\varphi]\psi)$. But this is not valid! The following countermodel M demonstrates this clearly.

Consider a model M for two agents a and b and two facts p and q. Its domain is $\{11, 01, 10\}$, where 11 is the state where p and q are both true, 01 the state where p is false and q is true, and 10 the state where p is true and q is false. Agent a cannot tell 11 and 01 apart, whereas b cannot tell 01 and 10 apart. So the partition for a on the domain is $\{11, 01\}, \{10\}$ and the partition for b on the domain is $\{11\}, \{01, 10\}$. See Figure 6. Now consider the instance

$[p]C_{ab}q \leftrightarrow (p \rightarrow C_{ab}[p]q)$ of this supposed principle. The left side of the equivalence is true in state 11 of M, whereas the right side is false in that state. We show that as follows. First, $M, 11 \models [p]C_{ab}q$ is true in 11, because $M, 11 \models p$ and $M|p, 11 \models C_{ab}q$. For the result of the announcement of p in $(M, 11)$, see Figure 6. The model $M|p$ consists of two disconnected states; obviously, $M|p, 11 \models C_{ab}q$, because $M|p, 11 \models q$ and 11 is now the only reachable state from 11. On the other hand, we have that $M, 11 \not\models p \rightarrow C_{ab}[p]q$, because $M, 11 \models p$ but $M, 11 \not\models C_{ab}[p]q$. The last is because $11 \sim_{ab} 10$ (because $11 \sim_a 01$ and $01 \sim_b 10$), and $M, 10 \not\models [p]q$. When evaluating q in $M|p$, we are now in the *other* disconnected part of $M|p$, where q is false: $M|q, 10 \not\models q$.

The general principle relating announcement and common knowledge is not an axiom, but a derivation rule:

From $\chi \rightarrow [\varphi]\psi$ *and* $\chi \wedge \varphi \rightarrow E_A\chi$*, infer* $\chi \rightarrow [\varphi]C_A\psi$.

More recent developments in the area use a different modal notion, 'relativised common knowledge', of which standard common knowledge can be seen as a special case [59, 31]. This results in more expressive logics, and the relation between announcements and relativised common knowledge is again an axiom. This greatly simplifies completeness proofs for such logics.

Public announcements and interpreted systems We now *outline* the relation between 'next' operator X, as used to describe interpreted system behaviour, and the announcement operators introduced in this section. An announcement is seen as a completely observable clock tick, synchronizing the system. Announcing φ at time m is simulated in $\mathcal{I}$ by changing the value of some environmental variable p for exactly those points where φ is true, when transiting from point (r, m) to point $(r, m+1)$, and passing on that information to the local states of the agents. The static information available at time m is contained in the restriction $\mathcal{I}|m$ of the interpreted system $\mathcal{I}$ to all points for time m. This determines the meaning of purely epistemic formulas. But for formulas containing epistemic and 'next'-temporal operators the situation is more complex. Assume that for each time m there is a formula φ such that the *only* transitions allowed at m are those induced by announcement of φ. We can define a translation $*$ where, given an epistemic state and a formula, each X-operator in that formula is replaced by a corresponding dynamic operator $[\varphi]$. The following now are all equivalent

$$\begin{array}{l}
\text{if } \mathcal{I}, (r, m) \models \varphi, \text{ then } \mathcal{I}, (r, m) \models X\psi \\
\text{if } \mathcal{I}, (r, m) \models \varphi, \text{ then } \mathcal{I}, (r, m+1) \models \psi \\
\text{if } \mathcal{I}|m, (r, m) \models \varphi^*, \text{ then } \mathcal{I}|m|\varphi^*, (r, m) \models \psi^* \\
\mathcal{I}|m, (r, m) \models [\varphi^*]\psi^*
\end{array}$$

In case φ and ψ are both purely epistemic, so that $\varphi^* = \varphi$, and $\psi^* = \psi$, we have that

$$\mathcal{I}, (r, m+1) \models \psi \text{ corresponds to } \mathcal{I}|m|\varphi, (r, m) \models \psi$$

There are other ways in which non-public actions relate to runs through interpreted systems (the relation between interpreted systems, temporal logics, and dynamic logics is currently much investigated by the research community).

Notes The logic of multi-agent epistemic logic with public announcements and without common knowledge has been formulated and axiomatised by Plaza [49]. For the somewhat more general case of introspective agents this was done 'again' by Gerbrandy and Groeneveld [21], who were not aware of Plaza's work at the time. In [49], public announcement is seen as a binary operation $+$, such that $\varphi + \psi$ is equivalent to $\langle\varphi\rangle\psi$. The logic of public announcements *with* common knowledge was axiomatised by Baltag, Moss, and Solecki [8], see also [9, 6, 7], in a more general setting that will be discussed in the next section: the completeness of their proof system is a special case of the completeness of their more general logic of action models.

There are a fair number of precursors of these results. One prior line of research is in dynamic modal approaches to semantics, not necessarily also epistemic: 'update semantics'. Another prior line of research is in meta-level descriptions of epistemic change, not necessarily on the object level as in dynamic modal approaches. This relates to the temporal epistemics and interpreted systems approach for which we therefore refer to the summary discussion in the previous section.

The 'dynamic semantics' or 'update semantics' was followed in van Emde Boas, Groenendijk, and Stokhof [69], Landman [33], Groeneveld [23], and Veltman [72]. There are strong relations between that and more *PDL*-motivated work by de Rijke [12], and Jaspars [29]. As background literature to various dynamic features introduced in the 1980s and 1990s we recommend van Benthem [57, 56]. More motivated by runs in interpreted systems is van Linder, van der Hoek, and Meyer [70]. All these approaches use dynamic modal operators for information change, but (1) typically not (except [70]) in a multi-modal language that also has epistemic operators, (2) typically not for more than one agent, and (3) not necessarily such that the effects of announcements or updates are defined given the update formula and the current information state: the *PDL*-related and interpreted system related approaches *presuppose* a transition relation between information states, such as for atomic actions in *PDL*.

We outline, somewhat arbitrarily, some features of these approaches. Groeneveld's approach [23] is typical for dynamic semantics in that is has formulas $[\varphi]_a\psi$ to express that after an update of agent a's information with φ, ψ is true. His work was later merged with that of Gerbrandy, resulting in the seminal [21]. De Rijke [12] defines theory change operators $[+\varphi]$ and $[*\varphi]$ with a dynamic interpretation that link an enriched dynamic modal language to AGM-type theory revision [1] (see also Section 5 addressing dynamic epistemics for belief revision). In functionality, it is not dissimilar from Jaspars [29] φ-addition (i.e., expansion) operators $[\varphi]_u$ and φ-retraction (i.e., contraction) operators $[\varphi]_d$, called updates and downdates by Jaspars. Van Linder, van der Hoek, and Meyer [70] use a setting that combines dynamic effects with knowledge and belief, but to interpret their (various) action operators they assume an explicit transition relation as part of the Kripke structure interpreting such descriptions.

As somewhat parallel developments to [19] we also mention Lomuscio and Ryan [39]. They do not define dynamic modal operators in the language, but they define epistemic state transformers that clearly correspond to the interpretation of such operators: $M * \varphi$ is the result of refining epistemic model M with a formula φ, etc. Their semantics for updates is only an *approximation* of public announcement logic, as the operation is only defined for *finite* (approximations of) models.

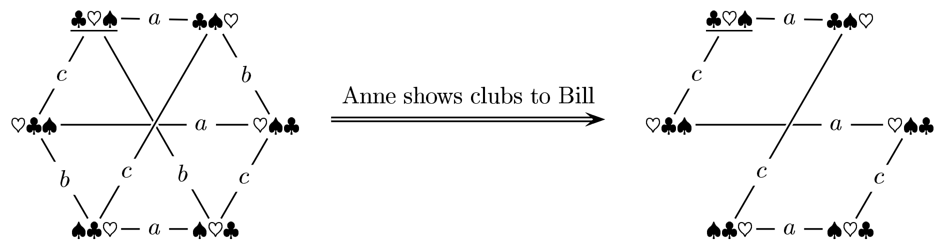

Figure 7: On the left, the Kripke model for three players each holding one card. On the right, the effect of Anne showing her clubs card to Bill.

4.2 Epistemic Actions

Some epistemic actions are more complex than public announcements, where the effect of the action is always a restriction on the epistemic model. Let us reconsider the epistemic state (*Hexa*, ♣♡♠) wherein Anne holds clubs, Bill holds hearts, and Cath holds spades. And consider again one of the example actions in the introduction:

> *Anne shows (only) to Bill her clubs card. Cath cannot see the face of the shown card, but notices that a card is being shown.*

It is assumed that it is publicly known what the players can and cannot see or hear. Call this action showclubs. The epistemic state transition induced by this action is depicted in Figure 7. Unlike after public announcements, in the showclubs action we cannot eliminate any state. Instead, all b-links between states have now been severed: whatever was the actual deal of cards, Bill now knows that card deal and cannot imagine any alternatives. We hope to demonstrate the intuitive acceptability of the resulting epistemic state. After the action showclubs, Anne considers it possible that Cath considers it possible that Anne has clubs. That much is obvious, as Anne has clubs anyway. But Anne also considers it possible that Cath considers it possible that Anne has hearts, because Anne considers it possible that Cath has spades, and so does not know whether Anne has shown clubs or hearts. It is even the case that Anne considers it possible that Cath considers it possible that Anne has spades, because Anne considers it possible that Cath does not have spades but hearts, in which case Cath would not have known whether Anne has shown clubs or spades. And in all those cases where Anne shows her card, Bill obviously would have learnt the deal of cards. Note that, even though for Cath there are only two 'possible actions'—showing clubs or showing hearts—none of the *three* possible actions can apparently be eliminated 'from public consideration'.

But it can become even more complex. Imagine the following action, rather similar to the showclubs action:

> *Anne whispers into Bill's ear that she does not have the spades card, given a (public) request from Bill to whisper into his ear one of the cards that she does not have.*

This is the action whispernospades. Given that Anne has clubs, she *could* have whispered "no hearts" or "no spades". And whatever the actual card deal was, she could always have chosen between two such options. An epistemic state results that reflects that choice, and

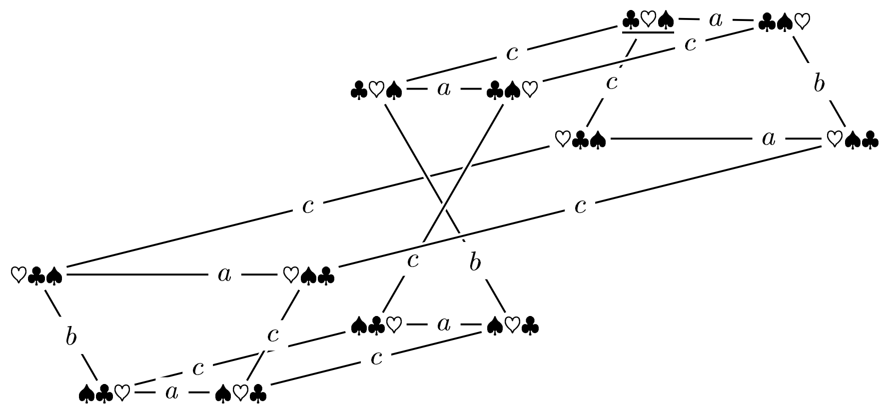

Figure 8: After Anne whispered into Bill's ear that she does not have the spades card, given a (public) request from Bill to whisper into his ear one of the cards that she does not have. Assume transitivity of the accessibility relation for Cath.

that therefore consists of $6 \times 2 = 12$ different states. It is depicted in Figure 8 (wherein we assume transitivity of the accessibility relation for c). The reader may ascertain that the desirable postconditions of the action whispernospades indeed hold. For example, given that Bill holds hearts, Bill will now have learnt from Anne what Anne's card is, and thus the entire deal of cards. So there should be no alternatives for Bill in the actual state (the underlined state ♣♡♠ 'at the back' of the figure—for convenience, different states for the same card deal have been given the same name). But Cath does not *know* that Bill knows the card deal, as Cath considers it possible that Anne actually whispered "no hearts" instead. That would have been something that Bill already knew, as he holds hearts himself—so from that action he would not have learnt very much. Except that Cath could then have imagined him to know the card deal... Note that in Figure 8 there is also another state named ♣♡♠, 'in the middle', so to speak, that is accessible for Cath from the state ♣♡♠ 'at the back', and that satisfies that Bill doesn't know that Anne has clubs.

The intuition behind action models An elegant formal way to model such actions and a large class of similar events is called 'action model logic' [8]. The basic idea is that actions can profitably be modelled in relation to other, 'similar', actions, in a way similar to how different states in a Kripke model relate to each other. When Anne shows her clubs card to Bill, this is indistinguishable for Cath from Anne showing her hearts card to Bill—if she were to have that. And, as Anne considers it possible that Cath holds hearts instead of spades, Anne also considers it possible that Cath interprets her card showing action as yet a third option, namely showing spades. These three different card showing actions are therefore, from a public perspective, all indistinguishable for Cath, but, again from a public perspective, all different for Anne and Bill. We can therefore visualise the 'epistemic action' of Anne showing clubs to Bill as some kind of Kripke structure, namely with a domain of three 'action points' standing for 'showing clubs', 'showing hearts', and 'showing spades', and accessibility relations for the three players corresponding to the observations above. We now have what is called an *action model*. What else do we need? To relate such 'action models' to the preconditions for their execution, we associate to each action point in such a model a formula in a logical language: the precondition of that action point. To execute an epistemic action, we compute what is known as the restricted modal product of the current epistemic state and the epistemic action. The result is 'the next epistemic state'. It is a *product* because the domain of the next epistemic state is a subset of the cartesian product of the domain of the current epistemic state and the domain of the action model. It is *restricted* because we restrict that full product to those (state,action) pairs such that the precondition for the action of the pair is satisfied in the state of the pair. Two states in the new epistemic state indistinguishable (accessible), if and only if the states in the previous epistemic state from which they evolved were already indistinguishable (accessible), and if the two different actions executed there were also indistinguishable. For example, Cath cannot distinguish the result of Anne showing clubs in state ♣♡♠ from Anne showing hearts in state ♡♣♠, because in the first place she could not distinguish those two card deals, and in the second place she cannot distinguish Anne showing clubs from Anne showing hearts.

Formal definitions We successively define action models and their execution, the language of action model logic, and the semantics of action model logic. An example follows, and also some explanations for those readers puzzled by a seeming mixup of syntax and semantics (we can allay their fears: there is no problem). As usual, we assume background parameters in the form of a set of agents A and a set of propositional variables P.

Definition 6 (Action model) Let $\mathcal{L}$ be a logical language. An action model U is a structure $\langle \mathsf{S}, \mathsf{R}, \mathsf{pre} \rangle$ such that S is a domain of *action points*, such that for each $a \in A$, R_a is an accessibility relation on S, and such that $\mathsf{pre} : \mathsf{S} \rightarrow \mathcal{L}$ is a preconditions function that assigns a *precondition* $\mathsf{pre}(\mathsf{s}) \in \mathcal{L}$ to each $\mathsf{s} \in \mathsf{S}$. A *pointed action model* is a structure (U, s) with $\mathsf{s} \in \mathsf{S}$.

Public announcement of φ is modelled by a singleton action model consisting of action point s, accessible to all agents, and with precondition $\mathsf{pre}(\mathsf{s}) = \varphi$.

Definition 7 (Execution) Given an epistemic state (M, s) with $M = \langle S, R, V \rangle$ and an epistemic action (U, s) with $\mathsf{U} = \langle \mathsf{S}, \mathsf{R}, \mathsf{pre} \rangle$ such that $M, s \models \mathsf{pre}(\mathsf{s})$. The result of executing (U, s) in (M, s) is the epistemic state $((M \otimes \mathsf{U}), (s, \mathsf{s}))$ where $(M \otimes \mathsf{U}) = \langle S', R', V' \rangle$ is a restricted modal product of M and U defined as

$$\begin{array}{lll} S' & \equiv & \{(s, \mathsf{s}) \mid s \in S, \mathsf{s} \in \mathsf{S}, \text{ and } M, s \models \mathsf{pre}(\mathsf{s})\} \\ R'_a((s, \mathsf{s}), (t, \mathsf{t})) & \text{iff} & R_a(s, t) \text{ and } \mathsf{R}_a(\mathsf{s}, \mathsf{t}) \\ (s, \mathsf{s}) \in V'_p & \text{iff} & s \in V_p \end{array}$$

There is only one more step to make: to give a logical language with an inductive construct for action models. For obvious reasons of well-definedness such action models are now required to be *finite*.

Definition 8 (Language) The language of action model logic is the union of the *formulas* φ and the *epistemic actions* α defined by

$$\begin{array}{lll} \varphi & ::= & p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_B\varphi \mid [\alpha]\varphi \\ \alpha & ::= & (\mathsf{U}, \mathsf{s}) \mid (\alpha \cup \alpha) \end{array}$$

where $p \in P$, $a \in A$, $B \subseteq A$, and (U, s) a *finite* pointed action model.

Definition 9 (Semantics)

$$
\begin{array}{lll}
M,s \models [\mathsf{U},\mathsf{s}]\varphi & \text{iff} & M,s \models \mathsf{pre}(\mathsf{s}) \text{ implies } (M \otimes \mathsf{U}),(s,\mathsf{s}) \models \varphi \\
M,s \models [\alpha \cup \beta]\varphi & \text{iff} & M,s \models [\alpha]\varphi \text{ and } M,s \models [\beta]\varphi
\end{array}
$$

A requirement that may be implicit in the definition of the language, is that given such a (U,s), the preconditions for all (but finitely many) action points t in its domain S should already have been constructed in a previous stage of this inductively defined hierarchy. As usual, $\langle\alpha\rangle\varphi$ is defined by notational abbreviation as $\neg[\alpha]\neg\varphi$. We also use the notational abbreviation U for $\bigcup_{\mathsf{s}\in\mathsf{S}}(\mathsf{U},\mathsf{s})$, in other words, action model U can also be seen as a non-deterministic epistemic action.

The reader may be puzzled by finding the *semantic* object of a pointed action model reappear as a primitive in the language. Strictly, in the language we are only *naming* the pointed frame underlying (U,s) as a *name* for that frame, and the precondition function pre associated with U is the way to construct a more complex expression of type 'epistemic action' from that framename and from arguments that are simpler, already defined, expressions of type 'formula'. The arity of that framename is then of course the number of actions in the domain S of U. We choose to ignore the difference, just as we ignore the difference between agent *names* a in modal operators K_a and *actual* agents a in the accessibility relations R_a that interpret such operators.

Definition 10 (Composition of action models) Let $\mathsf{U} = \langle \mathsf{S},\mathsf{R},\mathsf{pre}\rangle$ and $\mathsf{U}' = \langle \mathsf{S}',\mathsf{R}',\mathsf{pre}'\rangle$ be two action models. Their *composition* $(\mathsf{U}\ ;\ \mathsf{U}')$ is the action model $\langle \mathsf{S}'', R'', \mathsf{pre}''\rangle$ such that

$$
\begin{array}{lll}
\mathsf{S}'' & \equiv & \mathsf{S} \times \mathsf{S}' \\
\mathsf{R}''_a((\mathsf{s},\mathsf{s}'),(\mathsf{t},\mathsf{t}')) & \text{iff} & \mathsf{R}_a(\mathsf{s},\mathsf{t}) \text{ and } \mathsf{R}'_a(\mathsf{s}',\mathsf{t}') \\
\mathsf{pre}''((\mathsf{s},\mathsf{s}')) & \equiv & \langle \mathsf{U},\mathsf{s}\rangle \mathsf{pre}'(\mathsf{s}')
\end{array}
$$

The definition of composition extends in the obvious way to pointed action models: given two pointed action models (U,t) and $(\mathsf{U}',\mathsf{t}')$ as above, their composition $(\mathsf{U},\mathsf{t})\ ;\ (\mathsf{U}',\mathsf{t}')$ is the pointed action model $(\mathsf{U}'',(\mathsf{t},\mathsf{t}'))$, with U'' defined as above. Note that $\mathsf{pre}''((\mathsf{t},\mathsf{t}')) = \langle \mathsf{U},\mathsf{t}\rangle \mathsf{pre}'(\mathsf{t}')$. Compare this composition construction to the rule for a sequence of announcements. Composition of action models is essential in the completeness proof of the logic.

From a logical philosophical point of view it is not surprising to see accessibility relations on some domain of objects. But there are two surprising differences with how such accessibility relations typically appear. *First*, the alternatives in the domain of an action model do not stand for factual states of the world, in other words for 'static objects', but they stand for possible actions, i.e., 'dynamic objects'. *Second*, the domain objects appear not to be semantic, but syntactic: preconditions of actions are formulas in the logical language! Relational structures such as Kripke models are from the realm of *semantics*, that tends, in logic, to be strictly separated from the realm of *syntax* or logical language. But, we just have to say this again, action models appear not to be partitioning *semantic* objects, but *syntactic* objects: the preconditions associated with domain objects are *formulas*. The first difference is something that makes logicians happy: "Right, so we have a dynamic counterpart of a static object that we already know quite well." And we will see that in other ways it also combines well with the 'static' epistemic states that we already have. But the second difference, to the contrary, is something that makes logicians unhappy: "Pfui! One must keep syntax and semantics strictly separated, or the world will fall apart." As we have already explained, the solution is surprisingly simple, and means that we name Kripke frames, in the line of similar solutions for dynamic logics incorporating automata.

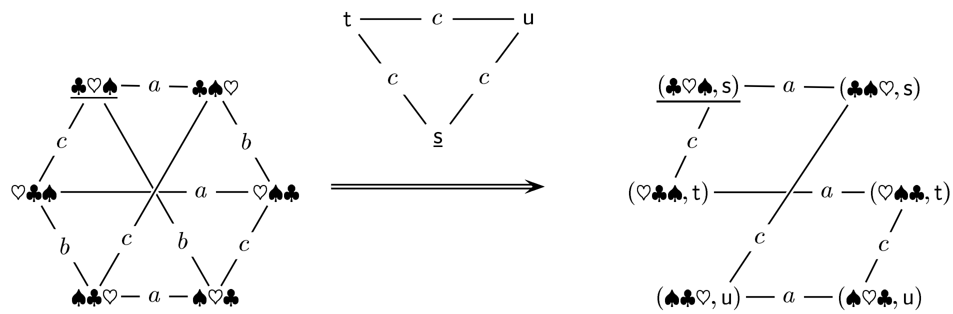

Figure 9: On the left, the Kripke model for three players each holding one card. On the right, the effect of Anne showing her clubs card to Bill, shown as a restricted modal product. The preconditions for action points $\mathsf{s}, \mathsf{t}, \mathsf{u}$ are, respectively $Clubs_a, Hearts_a, Spades_a$. Compare to Figure 7.

Example For a more concrete example, we now model the epistemic action $\mathsf{showclubs}$ where Anne shows her clubs card to Bill without Cath being able to see which card, as an action model, and we execute it in the epistemic state $(Hexa, \clubsuit\heartsuit\spadesuit)$. The action model corresponding to $\mathsf{showclubs}$ consists of three action points $\mathsf{s}, \mathsf{t}, \mathsf{u}$, with preconditions $\mathsf{pre}(\mathsf{s}) = Clubs_a$, $\mathsf{pre}(\mathsf{t}) = Hearts_a$, and $\mathsf{pre}(\mathsf{u}) = Spades_a$. These three action points are indistinguishable for Cath (so her accessibility relation on that domain is the universal relation), but they can be distinguished by both Anne and Bill from one another. The point of the model is, obviously, s, as Anne actually holds clubs. The action model is depicted in Figure 9, and also the result of its execution is shown.

Concerning accessibility relations, note that e.g. $R_a((\clubsuit\heartsuit\spadesuit, \mathsf{s}), (\clubsuit\spadesuit\heartsuit, \mathsf{s}))$ because $R_a(\mathsf{s}, \mathsf{s})$ and $R_a(\clubsuit\heartsuit\spadesuit, \clubsuit\spadesuit\heartsuit)$. Also, $R_c((\clubsuit\heartsuit\spadesuit, \mathsf{s}), (\heartsuit\clubsuit\spadesuit, \mathsf{t}))$, because $R_c(\clubsuit\heartsuit\spadesuit, \heartsuit\clubsuit\spadesuit)$ and $R_c(\mathsf{s}, \mathsf{t})$. But Bill has learnt what the card deal is! Before the execution of the action, $R_b(\clubsuit\heartsuit\spadesuit, \spadesuit\heartsuit\clubsuit)$, but afterwards these deals can now be distinguished by Bill: $(\clubsuit\heartsuit\spadesuit, \mathsf{s})$ is different for Bill from $(\spadesuit\heartsuit\clubsuit, \mathsf{t})$ because s can be distinguished by Bill from t.

Similarly, we can compute the epistemic state resulting from the action $\mathsf{whispernospades}$ where Anne whispers into Bill's ear one of the cards that she does not have. It has a similarly structured action model as the one for $\mathsf{showclubs}$, except that the preconditions for action points $\mathsf{s}, \mathsf{t}, \mathsf{u}$ are now $\mathsf{pre}(\mathsf{s}) = \neg Clubs_a$, $\mathsf{pre}(\mathsf{t}) = \neg Hearts_a$, and $\mathsf{pre}(\mathsf{u}) = \neg Spades_a$. In this case each action point is executable in four of *Hexa*'s states (unlike the card showing action where each action point was executable in only two of *Hexa*'s states). The model in Figure 8 with twelve states results.

Notes The action model framework has been developed by Baltag, Solecki, and Moss, and has appeared in various forms [8, 9, 6, 7]. The final form of their semantics is Baltag and Moss's [7]. A final publication on the completeness and expressivity results is still in preparation. A different but also rather expressive way to model epistemic actions was suggested by Gerbrandy in [19]—this generalizes the results by Gerbrandy and Groeneveld in [21]. Gerbrandy's action language can be seen as defined by relational composition, interpreted on non-wellfounded set theoretical structures corresponding to bisimilarity classes of pointed Kripke models. Van Ditmarsch explored another relational action language—but based on standard Kripke semantics—[62, 63] and was influenced by both Gerbrandy and Baltag et al. His semantics is restricted to $S5$ model transformations. Van Ditmarsch et al. later proposed *concurrent epistemic actions* in [66]. Their treatment of concurrency for dynamic operators is similar to that in the logic cPDL—for 'concurrent propositional dynamic logic'—proposed by Peleg [48] and also mentioned in, e.g., Goldblatt [22] and Harel et al. [26]. How the expressivity of these different action logics compares is unclear. Recent developments include a proposal by Economou in [14]. Relativised common knowledge, already discussed in the previous subsection, also combines well with action models [59]; for adding assignments to allow factual change as well, see the final notes to this chapter.

5 Belief Change and Dynamic Logic

Standard belief revision The traditional emphasis in what is known as the area of 'belief revision' is *theory revision*; how to change a deductively closed set of formulas $\mathcal{K}$ into another deductively closed set of formulas. Overview publications for this area are [1] and [18]. A theory typically consists of objective, i.e. non-epistemic, 'beliefs' that are changed relative to expansion, contraction, or revision, and also typically from the point of view of a single agent. Note that 'belief' means nothing but 'formula' here: there is no explicit representation of the status of such a formula as 'believed by the agent', as in epistemic logic. In the case of an *expansion*, new information described by a formula φ is incorporated into a theory $\mathcal{K}$ by somehow 'adding' φ to the theory. For the result we write $\mathcal{K} \oplus \varphi$. In case the negation was already in $\mathcal{K}$, the result will be the inconsistent theory $\mathcal{K}_\perp$ that consists of all formulas. The result of a *contraction* should be that the formula with which the theory $\mathcal{K}$ is contracted is longer be believed, for which we write $\varphi \notin \mathcal{K} \ominus \varphi$. Contraction with a validity cannot be successful, as all validities are in all theories. Contraction with a validity therefore leaves a theory unchanged. In the case of a revision, the negation $\neg\varphi$ of the revision formula φ is typically in the theory $\mathcal{K}$ but for the revision to 'succeed' the revised (consistent) theory should (also) be consistent. A process of mere expansion therefore does not work, and one first has to contract the theory in a way that removes $\neg\varphi$ from it and all its dependencies. This can then be followed by an expansion with φ. For the result of the revision of theory $\mathcal{K}$ with formula φ we write $\mathcal{K} \circledast \varphi$. The Levi-identity states that $\mathcal{K} \circledast \varphi = \mathcal{K} \ominus \neg\varphi \oplus \varphi$: a revision can be seen as a contraction followed by an expansion. The 'theories' $\mathcal{K}$ that we are changing can have any shape, such as first-order theories. Here, we restrict ourselves to propositional theories, and we investigate possibilities to extend this to theories of propositional modal formulas.

Yet another issue in traditional belief revision comes under the name of 'update'. An *update*—unfortunately a clash cannot be avoided with the more general meaning of that term in dynamic epistemic logic, where it incorporates belief revision as well—is a *factual* change, as opposed to a belief change in the three previously distinguished notions. The latter merely express a different agent stance towards a non-changing world, but in an 'update' the world itself changes. The standard reference for that is [30]. We will pay only minor attention to such updates in this section.

Belief change with dynamic non-epistemic logic The different 'theory change operators' $\oplus$, $\ominus$, and $\circledast$ can be reinterpreted as dynamic modal operators. A straightforward way to implement that, is some logic in which $[\circledast\varphi]\psi$ expresses that after revision with φ, ψ holds—where ψ actually means, as in [1], 'ψ is believed by the agent'. This approach was suggested by van Benthem in [56][2] and further developed by de Rijke in [12]. They propose a semantical counterpart of a total order on theories, in the form of an 'updating' and 'downdating' relation between states or worlds, standing for theories, and interpret the modal operator as a transition in such a structure according to these relations. 'Updating' models expansion: it relates the current state to states that result from expansion. 'Downdating' models contraction. It relates states that result from contraction to the current state. Revision is indeed downdating followed by updating. In this overview we focus on approaches that extend epistemic logics, therefore we do not give more details on this non-epistemic approach.

Belief change with dynamic epistemic logic In the approach by Segerberg and collaborators [36, 52, 51, 37] beliefs are represented explicitly. We now identify a theory $\mathcal{K}$ with the believed formulas (or some subset of the believed formulas) in an epistemic state: $\mathcal{K} = \{\psi \mid M, s \models B\psi\}$. As in [12] they express belief change with dynamic modal operators $[\oplus\varphi]$, $[\ominus\varphi]$, and $[\circledast\varphi]$. In a typical revision where we have that $\neg\varphi \in \mathcal{K}$, $\varphi \in \mathcal{K} \circledast \varphi$, and $\neg\varphi \notin \mathcal{K} \circledast \varphi$, we now get

- $M, s \models B\neg\varphi$
- $M, s \models [\circledast\varphi]B\varphi$
- $M, s \models [\circledast\varphi]\neg B\neg\varphi$

For contraction, we want that in case $M, s \models B\varphi$, after contraction φ is no longer believed, i.e., $M, s \models [\ominus\varphi]\neg B\varphi$. Similarly, for expansion we aim to achieve $M, s \models [\oplus\varphi]B\varphi$.

This approach is known as *dynamic doxastic logic* or **DDL**. Similar to [12] it presumes a transition relation between states representing theories, but this is now differently realized, namely using what is known as a Segerberg-style semantics wherein factual and epistemic information—under the terms of world component and doxastic component—are strictly separated. A dynamic operator is interpreted as a transition along the 'lines' of minimal theory change set out by this given structure, with the additional restriction that the transitions describe epistemic (doxastic) change only, and not factual change. This restriction is enforced by not allowing the 'world component' to change in the transition relation but only the 'doxastic component' [36, p.18].

There are now two options: either we restrict ourselves to beliefs in *objective* (boolean, non-epistemic) formulas, and we get what is known as basic **DDL**, as in [36, 52]. Or we allow higher-order beliefs, as in the dynamic epistemics described in previous sections. We thus get 'full' or 'unlimited' **DDL**, also discussed in [36] but mainly in [37]. Without the restriction to belief of objective formulas, a number of problems come to the fore related to higher order belief, knowledge growth, 'success' of revision, and multi-agent belief. We address these issues in the dynamic epistemic setting in this chapter, that provides a 'third way' given the two views on belief revision with dynamic logic presented so far. In dynamic epistemics, unlike the two approaches to dynamic belief revision already presented, the transition that interprets the dynamic operators is *induced* by the current information state and the revision formula, and does not *assume* such a transition relation. We demonstrate the various possible transitions by a simple example.

[2]It is only one of many topics covered in that publication, namely Section 6, pages 714–715, 'Cognitive procedures over information patterns'. Note this work is similar to a 1991 technical report.

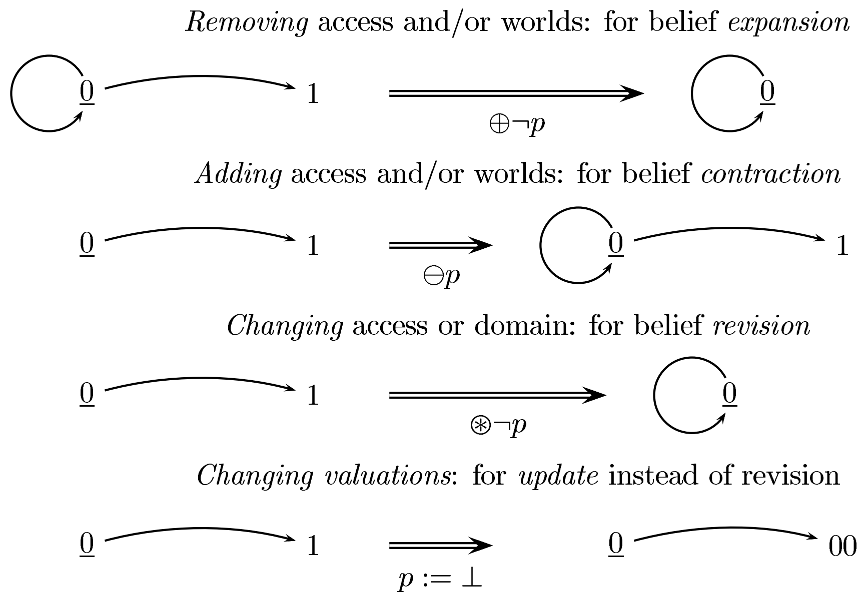

Figure 10: Possible changes of belief mirrored in Kripke structure transitions

Examples of belief change with dynamic epistemic logic Consider expressing and changing uncertainty about the truth of a single fact p, and assume an information state where the agent (whose beliefs are interpreted by the unlabeled accessibility relation depicted) may be uncertain about p and where p is actually false (indicated by 'designating' the actual state by underlining it). Figure 10 lists all conceivable sorts of belief change.

In the top structure, uncertainty about the fact p (i.e., absence of belief in p and absence of belief in $\neg p$) is changed into belief in $\neg p$. On the left, $\neg Bp$ is true, and on the right $B\neg p$. In the second from above, belief in p is weakened to uncertainty about p, and in the third from above we change from Bp to $B\neg p$. Note that also in this semantic setting of Kripke-structure transformation, belief revision can again be seen as a contraction followed by an expansion, so we may in principle consider semantic alternatives for the Levi-identity. The last information state transition in Figure 10 depicts factual change. The state with changed valuation has suggestively been renamed from 1 to 00, although formally, of course, it is only the valuation of a named state that changes. The 'assignment' or substitution $p := \bot$ indicates that the valuation of atom p is revised into the valuation of the assigned formula. As this is $\bot$, the new valuation of p (seen as a subset of the domain) is now the empty set of states.

Public announcement as belief expansion The public announcement logic already discussed in detail can be seen as an implementation of a belief expansion operator for higher-order belief (i.e., both fully introspective beliefs of a single agent but also beliefs of agents in a group about the beliefs of other agents in that group), where the next information state is computed from 'merely' the current information state and the expansion formula. This computation is straightforward for expansion, as restricting the domain or accessibility relation can easily be seen as structurally related to the existing model. The semantics of the public announcements already presented operates just like that: it restricts a model M to the submodel $M|\varphi$ consisting of the worlds where the announcement is true. From here on, for

$[\varphi]\psi$ we write $[\oplus\varphi]\psi$, and we focus on knowledge K.

Knowledge growth In such a higher-order setting we cannot maintain the expansion postulates. First, we have to revise our ideas about 'minimal change'. In particular, it can no longer be maintained that expanded theories contain their predecessors:

Identify a theory $\mathcal{K}$ as before with the set of known formulas in an information state: $\{\psi \mid M, s \models K\psi\}$. Let $M, s \models \varphi$. Suppose $\mathcal{K} \subset \mathcal{K} \oplus \varphi$, then there must be at least one formula ψ such that $\psi \in \mathcal{K} \oplus \varphi$ but $\psi \notin \mathcal{K}$. From $\psi \in \mathcal{K} \oplus \varphi$ follows by positive introspection that $K\psi \in \mathcal{K} \oplus \varphi$. From $\psi \notin \mathcal{K}$ follows by negative introspection that $\neg K\psi \in \mathcal{K}$. From $\neg K\psi \in \mathcal{K} \subset \mathcal{K} \oplus \varphi$ and $K\psi \in \mathcal{K} \oplus \varphi$ follows a contradiction.

Therefore, strict knowledge *growth* is contradictory for introspective agents (we did not use the truth axiom $K\varphi \rightarrow \varphi$, so the results hold for introspective belief as well), as observed by many authors: one does not care to preserve ignorance of the expansion formula, when expanding a theory. Therefore, one cannot adhere closely to the *expansion postulate* which states that $\mathcal{K} \subseteq \mathcal{K} \oplus \varphi$ (also known as postulate $\mathcal{K} \oplus 3$). Fortunately, knowledge *change* in a way that reflects the ideas behind expansion is still possible. And also, knowledge growth is possible for *fragments* of the public announcement language; for an example, see [67].

Success The *success postulate*, which states that $\varphi \in \mathcal{K} \oplus \varphi$ (also known as postulate $\mathcal{K} \oplus 2$), cannot be maintained either. The most basic example illustrating that, is an announcement of the Moore-sentence $p \wedge \neg Kp$. This sentence cannot be believed, or known, after its announcement [27, 43] (see also the introductory Section 1). Now this is first of all for the obvious reason that it is a Moore sentence, and by definition that is a sentence that cannot be believed, but it should be pointed out that as an announcement it can very well be true and therefore executed: after it, p is (publicly) known, which in fact entails the *negation* of $p \wedge \neg Kp$. But this means that expansion with $p \wedge \neg Kp$ *cannot* be successful: yet another barrier to satisfy the AGM postulates for higher-order belief expansion. Gerbrandy [19] calls this phenomenon an *unsuccessful update*; [20] is similar to [19]. The matter is also taken up in [65].

For truthful public announcements, the formulas φ that always become known after their announcement can be properly said to be the *successful formulas* and characterized by the validity of $[\oplus\varphi]\varphi$. This entails the validity of $\varphi \rightarrow [\oplus\varphi]C_A\varphi$ (in this setting where common knowledge of a formula entails its truth). The latter says that if φ is true, announcing φ makes it common knowledge, which more properly grasps what 'success' means in a higher-order setting. An intriguing question is: *Which* formulas are successful? An answer to that question would address knowledge expansion satisfactorily in this higher-order setting. But the answer is as yet unclear. Obvious inductive definitions fail. Even when both φ and ψ are successful, $\neg\varphi$ may be unsuccessful (for $\varphi = p \wedge \neg Kp$), $\varphi \wedge \psi$ may be unsuccessful (for $\varphi = p$ and $\psi = \neg Kp$), and as well $[\oplus\varphi]\psi$ and $\varphi \rightarrow \psi$ may be unsuccessful.

There are relevant successful fragments of the language. For example, *public knowledge formulas* are successful: $[\oplus C_A\varphi]C_A\varphi$ is valid. This follows from bisimulation invariance under point-generated submodel constructions. Another successful fragment form the *preserved formulas* (introduced for the language without announcements by van Benthem in [58]) that are inductively defined as $\varphi ::= p \mid \neg p \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid K_a\varphi \mid C_B\varphi \mid [\oplus\neg\varphi]\psi$ (where $B \subseteq A$). From $\varphi \rightarrow [\oplus\psi]\varphi$ for arbitrary ψ, follows $\varphi \rightarrow [\oplus\varphi]\varphi$ which is equivalent to $[\oplus\varphi]\varphi$; therefore preserved formulas are successful formulas. The inductive case $[\oplus\neg\varphi]\psi$ in the 'preserved

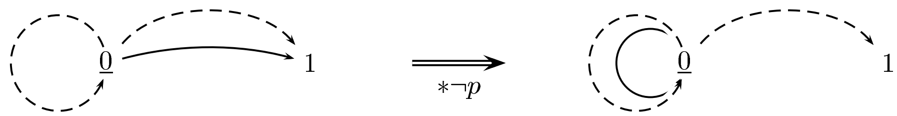

Figure 11: Belief revision related to degrees of belief

formulas' may possibly puzzle the reader. Its proof [65] is quite elementary (and proceeds by induction on formula structure) and shows that the puzzling *negation* in the announcement clause is directly related to the truth of the announcement as a *condition*:

Let $M, s \models [\oplus\neg\varphi]\psi$, and $M' \subseteq M$ such that $s \in M'$. Assume $M', s \models \neg\varphi$. Then $M, s \models \neg\varphi$ by *contraposition* of the inductive hypothesis for φ. From that and $M, s \models [\oplus\neg\varphi]\psi$ follows $M|\neg\varphi, s \models \psi$. From the inductive hypothesis for ψ follows $M'|\neg\varphi, s \models \psi$. Therefore $M', s \models [\oplus\neg\varphi]\psi$ by definition.

Dynamic doxastic logic as belief revision We now present a (different from DDL) dynamic doxastic semantics that can be seen as the implementation of a belief revision operator. Assuming that the new information state (pointed Kripke model) is constructed from the current information state and the revision formula (and does not assume an underlying transition structure), it seems harder to provide a mechanism that explains *adding* or *changing* access than one for merely deleting access—for the same reason that contraction or revision needs an entrenchment relation or something similar. A transition as in Figure 10 is hard to justify by the structure of the belief state before the revision. And another problem is that the dynamic epistemic logics presented so far do not provide a way to model 'forgetting' knowledge or beliefs, as, from an agent's point of view, belief and knowledge are indistinguishable in these logics (and 'belief' always means 'conviction').

In a setting where *degrees of belief* and possibly knowledge too are represented, one can provide such a structural justification. For a simple example, we add another degree of belief to the 'revision' example in Figure 10: in Figure 11 the dotted line interprets a stronger degree of belief. It contains (entails) the weaker, or most normal, belief that is represented as before with the solid line. There is a one-to-one correspondence between such accessibility relations satisfying inclusion, and preferences between worlds or a 'system of spheres' as in [35] and also propagated in [24, 53].

The basic idea in [35] is that, given a domain of worlds, from the perspective of a given world in that domain, some worlds may be preferable over others. The worlds for which a preference exists are the plausible worlds, and the preference is typically a partial order (plus additional constraints). In the left model in Figure 11, we have that, given world 0, the set of plausible worlds is $\{0, 1\}$, and that world 1 is preferred over world 0, for which we write $1 <^0 0$; relation $<^0$ is the preference relation associated with world 0. The set of plausible worlds given world 1 is empty; $<^1 = \emptyset$. The belief revision resulting in the model on the right now consists of changing preferences between the plausible worlds: in the resulting model world 0 is preferred over world 1: $0 <^0 1$. The relation between preferences and accessibility is fairly simple. In general, given a partial order with degrees of belief x we can define $R^x(s, s')$ if and only if the degree of world s' in the preference order $<^s$ is at most x. In Figure 11 we have two degrees of belief, and therefore two accessibility relations; the 'at most' is to ensure an inclusion relation. To accessibility relations R^x are associated 'degree of belief' modalities

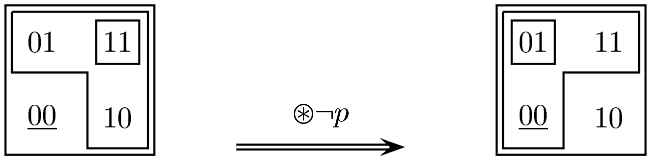

Figure 12: This depicts Spohn-like belief revision in a multimodal setting.

B^x in the obvious way.

Such a modal setting for reasoning about preferences also applies to a multi-agent situation, one can also restrict oneself to introspective belief or knowledge, and further demand additional (frame characterizable) restrictions expressing that agents are knowledgable about their own preferences. Concerning the static picture, such ideas have emerged under the name of 'dynamic interactive epistemology' in the game theoretical and philosophical community [54, 3, 11]—the word 'dynamic' refers to the conditional modal operators in those approaches that are used to model belief revision, not to the dynamic modal approach intended here.[3] The ideas have also surfaced as dynamic doxastic logic(s) in [4, 64]. We close this section with an example of the latter.

Example of introspective belief revision A proposal by Aucher [4] can be seen as an implementation of one of Spohn's proposals in [53][4]. An illustration is depicted in Figure 12. On the left in that figure, the agent believes atomic propositions p and q – the name 11 stands for the world where p and q are both true. In particular, Bp is true. Note that in this case there are three degrees of belief, let us say 0, 1, and 2. Degree of belief 0 stands for most normal belief and R^0 therefore corresponds to B. Apart from that, we have B^1 and B^2—in this case B^2 is equal to knowledge K. For example, $B^1(p \vee q)$ is valid on the model: the agent has a somewhat stronger belief in the (weaker proposition than $p \wedge q$ namely that) $p \vee q$.

On the right the agent believes $\neg p$ (and q). In other words, $B\neg p$ is true. On the left, it is therefore true that after revision with $\neg p$, the agent believes $\neg p$: $[\circledast\neg p]B\neg p$ is true. The belief revision is therefore successful. This revision is computationally achieved by the following recipe: determine the minimal world where the revision formula $\neg p$ is true. This is 01. Now deduct the degree of *that* world from the degree of *any* $\neg p$-world. We thus get a degree 0 for world 01, and a degree 1 for world 00. For worlds where the revision formula is false, i.e., where p is true, do the same, but ensure that the most normal p-world is slightly less preferable than the most normal $\neg p$. In this particular example we ensure this by 'adding 1 to the current degree'. This results in a degree 1 for world 11 and a degree 2 for world 10. This completes the computation.

[3] The relation between conditional modal operators and dynamic modal operators, and how appropriate it is to model belief revision in the former setting compared to the latter, seems to us only incompletely understood and merits further investigation. See [64], appendix A, and see [59] for the encompassing notion of 'relativized common knowledge'—a proposal to generalize conditional (individual) knowledge.

[4] Namely the revision also known as 'minimal Spohn': when revising with φ, make the minimal φ-worlds the 'most normal' worlds, such that they are believed after the revision; for details, see Definition 6 on page 117 in [53]. In Spohn's terms the revision in the example below would be called $\{00, 01\}$-1 conditionalization of the current ordinal conditional function, where $\{00, 01\}$ is the denotation of the revision formula $\neg p$ in the current epistemic state, and '1' is the decreased 'firmness α' with which the p worlds are updated.

The above presents only *one* example of *one* dynamic belief revision operator that can be seen as an implementation of the AGM postulates. This particular operator is successful on propositional formulas, and for those can also be considered as effecting minimal change. Other examples are given in [64], the list of over twenty different theory change operators in [50] seems also particularly suitable for implementation in this setting.

A final word on 'update'. As mentioned, belief update as opposed to belief revision is also investigated in dynamic epistemics under the name of 'factual change'. This is investigated in, for example, [9, 68, 59, 31]. These ideas also deserve to be properly applied to the belief revision arena.

For the further discussion of belief revision we refer to the main contribution on that topic in this encyclopedia.

References

[1] C.E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[2] S. Artemov. Evidence-based common knowledge. Technical report, CUNY, 2004. Ph.D. Program in Computer Science Technical Report TR-2004018.

[3] G.B. Asheim and Y. Søvik. Preference-based belief operators. *Mathematical Social Sciences*, 50(1):61–82, 2005.

[4] G. Aucher. A combined system for update logic and belief revision. Master's thesis, ILLC, University of Amsterdam, Amsterdam, the Netherlands, 2003.

[5] R.J. Aumann. Agreeing to disagree. *Annals of Statistics*, 4(6):1236–1239, 1976.

[6] A. Baltag. A logic for suspicious players: Epistemic actions and belief updates in games. *Bulletin of Economic Research*, 54(1):1–45, 2002.

[7] A. Baltag and L.S. Moss. Logics for epistemic programs. *Synthese*, 139:165–224, 2004. Knowledge, Rationality & Action 1–60.

[8] A. Baltag, L.S. Moss, and S. Solecki. The logic of common knowledge, public announcements, and private suspicions. In I. Gilboa, editor, *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 98)*, pages 43–56, 1998.

[9] A. Baltag, L.S. Moss, and S. Solecki. The logic of public announcements, common knowledge, and private suspicions. Technical report, Centrum voor Wiskunde en Informatica, Amsterdam, 1999. CWI Report SEN-R9922.

[10] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, Cambridge, 2001. Cambridge Tracts in Theoretical Computer Science 53.

[11] O. Board. Dynamic interactive epistemology. *Games and Economic Behaviour*, 49:49–80, 2004.

[12] M. de Rijke. Meeting some neighbours. In J. van Eijck and A. Visser, editors, *Logic and information flow*, pages 170–195, Cambridge MA, 1994. MIT Press.

[13] C. Dixon, M. Fisher, and M. Wooldridge. Resolution for temporal logics of knowledge. *Journal of Logic and Computation*, 8(3):345–372, 1998.

[14] P. Economou. Sharing beliefs about actions: A parallel composition operator for epistemic programs. In *Proceedings of the ESSLLI 2005 workshop Belief Revision and Dynamic Logic*, 2005. Available on `http://www.irit.fr/~Andreas.Herzig/Esslli05/`.

[15] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. MIT Press, Cambridge MA, 1995.

[16] H. Freudenthal. (formulation of the sum-and-product problem). *Nieuw Archief voor Wiskunde*, 3(17):152, 1969.

[17] G. Gamow and M. Stern. *Puzzle-Math*. Macmillan, London, 1958.

[18] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Bradford Books, MIT Press, Cambridge, MA, 1988.

[19] J.D. Gerbrandy. *Bisimulations on Planet Kripke*. PhD thesis, University of Amsterdam, 1999. ILLC Dissertation Series DS-1999-01.

[20] J.D. Gerbrandy. The surprise examination in dynamic epistemic logic. *Synthese*, 2006. To appear.

[21] J.D. Gerbrandy and W. Groeneveld. Reasoning about information change. *Journal of Logic, Language, and Information*, 6:147–169, 1997.

[22] R. Goldblatt. *Logics of Time and Computation*. CSLI Publications, Stanford, 2 edition, 1992. CSLI Lecture Notes No. 7.

[23] W. Groeneveld. *Logical investigations into dynamic semantics*. PhD thesis, University of Amsterdam, 1995. ILLC Dissertation Series DS-1995-18.

[24] A. Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.

[25] J.Y. Halpern, R. van der Meyden, and M.Y. Vardi. Complete axiomatizations for reasoning about knowledge and time. *SIAM Journal on Computing*, 33(3):674–703, 2004.

[26] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. MIT Press, Cambridge MA, 2000. Foundations of Computing Series.

[27] J. Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962.

[28] J. Hintikka. Reasoning about knowledge in philosophy. In J. Y. Halpern, editor, *Proceedings of the 1986 Conference on Theoretical Aspects of Reasoning About Knowledge*, pages 63–80, San Francisco, 1986. Morgan Kaufmann Publishers.

[29] J. Jaspars. *Calculi for Constructive Communication*. PhD thesis, University of Tilburg, 1994. ILLC Dissertation Series DS-1994-4, ITK Dissertation Series 1994-1.

[30] H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 387–394, 1991.

[31] B.P. Kooi. On public update logics. *Journal of Applied Non-Classical Logics*, 2006. To appear.

[32] S. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.

[33] F. Landman. *Towards a Theory of Information.* PhD thesis, University of Amsterdam, 1986.

[34] D.K. Lewis. *Convention, a Philosophical Study.* Harvard University Press, Cambridge (MA), 1969.

[35] D.K. Lewis. *Counterfactuals.* Harvard University Press, Cambridge (MA), 1973.

[36] S. Lindström and W. Rabinowicz. Belief change for introspective agents, 1999. `http://www.lucs.lu.se/spinning/`.

[37] S. Lindström and W. Rabinowicz. DDL unlimited: dynamic doxastic logic for introspective agents. *Erkenntnis*, 50:353–385, 1999.

[38] A.R. Lomuscio. *Knowledge Sharing among Ideal Agents.* PhD thesis, University of Birmingham, Birmingham, UK, 1999.

[39] A.R. Lomuscio and M.D. Ryan. An algorithmic approach to knowledge evolution. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AIEDAM)*, 13(2), 1999. Special issue on Temporal Logic in Engineering.

[40] J. McCarthy. Formalization of two puzzles involving knowledge. In Vladimir Lifschitz, editor, *Formalizing Common Sense : Papers by John McCarthy*, Ablex Series in Artificial Intelligence. Ablex Publishing Corporation, Norwood, N.J., 1990. original manuscript dated 1978–1981.

[41] J.-J.Ch. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science.* Cambridge Tracts in Theoretical Computer Science 41. Cambridge University Press, Cambridge, 1995.

[42] G.E. Moore. *Ethics.* Oxford University Press, 1912. Consulted edition: The Home University Library of Modern Knowledge, volume 54, OUP, 1947.

[43] G.E. Moore. A reply to my critics. In P.A. Schilpp, editor, *The Philosophy of G.E. Moore*, pages 535–677. Northwestern University, Evanston IL, 1942. The Library of Living Philosophers (volume 4).

[44] G.E. Moore. Russell's "theory of descriptions". In P.A. Schilpp, editor, *The Philosophy of Bertrand Russell*, pages 175–225. Northwestern University, Evanston IL, 1944. The Library of Living Philosophers (volume 5).

[45] R.C. Moore. Reasoning about knowledge and action. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence (IJCAI-77)*, Cambridge, Massachusetts, 1977.

[46] Y.O. Moses, D. Dolev, and J.Y. Halpern. Cheating husbands and other stories: a case study in knowledge, action, and communication. *Distributed Computing*, 1(3):167–176, 1986.

[47] R. Parikh. Knowledge and the problem of logical omniscience. In R. Zas and M. Zemankova, editors, *Proceedings of the 2nd international symposium on methodologies for intelligent systems*, pages 432–439, Amsterdam, 1987. North Holland.

[48] D. Peleg. Concurrent dynamic logic. *Journal of the ACM*, 34(2):450–479, 1987.

[49] J.A. Plaza. Logics of public communications. In M.L. Emrich, M.S. Pfeifer, M. Hadzikadic, and Z.W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[50] H. Rott. Adjusting priorities: Simple representations for 27 iterated theory change operators. Manuscript, 2004.

[51] K. Segerberg. Default logic as dynamic doxastic logic. *Erkenntnis*, 50:333–352, 1999.

[52] K. Segerberg. Two traditions in the logic of belief: bringing them together. In H.J. Ohlbach and U. Reyle, editors, *Logic, Language, and Reasoning*, pages 135–147, Dordrecht, 1999. Kluwer Academic Publishers.

[53] W. Spohn. Ordinal conditional functions: a dynamic theory of epistemic states. In W.L. Harper and B. Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume II, pages 105–134, 1988.

[54] R. Stalnaker. Knowledge, belief and counterfactual reasoning in games. *Economics and Philosophy*, 12:133–163, 1996.

[55] D. Steiner. A system for consistency preserving belief change. In *Proceedings of the ESSLLI Workshop on Rationality and Knowledge*, pages 133–144, 2006.

[56] J.F.A.K. van Benthem. Logic and the flow of information. In *Proceedings of the 9th International Congress of Logic, Methodology and Philosophy of Science (1991)*. Elsevier Science B.V., 1994. Also available as: Report LP-91-10, ILLC, University of Amsterdam.

[57] J.F.A.K. van Benthem. *Exploring logical dynamics.* CSLI Publications, 1996.

[58] J.F.A.K. van Benthem. One is a lonely number: on the logic of communication. Technical report, University of Amsterdam, 2002. ILLC Research Report PP-2002-27 (material presented at the Logic Colloquium 2002).

[59] J.F.A.K. van Benthem, J. van Eijck, and B.P. Kooi. Logics of communication and change, 2006. Accepted for *Information and Computation*.

[60] R. van der Meyden. Axioms for knowledge and time in distributed systems with perfect recall. In *Proceedings of the Ninth Annual IEEE Symposium on Logic in Computer Science (LICS-94)*, pages 448–457, Paris, July 1994.

[61] R. van der Meyden. Common knowledge and update in finite environments. *Information and Computation*, 140(2):115–157, 1998.

[62] H.P. van Ditmarsch. *Knowledge games.* PhD thesis, University of Groningen, 2000. ILLC Dissertation Series DS-2000-06.

[63] H.P. van Ditmarsch. Descriptions of game actions. *Journal of Logic, Language and Information*, 11:349–365, 2002.

[64] H.P. van Ditmarsch. Prolegomena to dynamic logic for belief revision. *Synthese (Knowledge, Rationality & Action)*, 147:229–275, 2005.

[65] H.P. van Ditmarsch and B.P. Kooi. The secret of my success. *Synthese*, 151:201–232, 2006.

[66] H.P. van Ditmarsch, W. van der Hoek, and B.P. Kooi. Concurrent dynamic epistemic logic. In V.F. Hendricks, K.F. Jørgensen, and S.A. Pedersen, editors, *Knowledge Contributors*, pages 45–82, Dordrecht, 2003. Kluwer Academic Publishers. Synthese Library Volume 322.

[67] H.P. van Ditmarsch, W. van der Hoek, and B.P. Kooi. Public announcements and belief expansion. In R. Schmidt, I. Pratt-Hartmann, M. Reynolds, and H. Wansing, editors, *Advances in Modal Logic*, volume 5, pages 335–346. King's College Publications, London, 2005.

[68] J. van Eijck. Guarded actions. Technical report, Centrum voor Wiskunde en Informatica, Amsterdam, 2004. CWI Report SEN-E0425.

[69] P. van Emde Boas, J. Groenendijk, and M. Stokhof. The conway paradox: Its solution in an epistemic framework. In J. Groenendijk, T. M. V. Janssen, and M. Stokhof, editors, *Truth, Interpretation and Information: Selected Papers from the Third Amsterdam Colloquium*, pages 159–182. Foris Publications, Dordrecht, 1984.

[70] B. van Linder, W. van der Hoek, and J.-J.Ch. Meyer. Actions that make you change your mind. In A. Laux and H. Wansing, editors, *Knowledge and Belief in Philosophy and Artificial Intelligence*, pages 103–146, Berlin, 1995. Akademie Verlag.

[71] G. van Tilburg. Doe wel en zie niet om (do well and don't look back). *Katholieke Illustratie (Catholic Illustrated Journal)*, 90(32):47, 1956. Breinbrouwsel 137 ('Brain Cracker' 137).

[72] F. Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25:221–261, 1996.

[73] G.H. von Wright. *An Essay in Modal Logic.* North Holland, Amsterdam, 1951.