Chapter 7

Dynamic Logics of Belief Change

Johan van Benthem and Sonja Smets

Contents

Abstract

This chapter gives an overview of current dynamic logics that describe belief update and revision, both for single agents and in multi-agent settings. We employ a mixture of ideas from AGM belief revision theory and dynamic-epistemic logics of information-driven agency. After describing the basic background, we review logics of various kinds of beliefs based on plausibility models, and then go on to various sorts of belief change engendered by changes in current models through hard and soft information. We present matching complete logics with dynamic-epistemic recursion axioms, and develop a very general perspective on belief change by the use of event models and priority

Chapter 7 of the *Handbook of Logics for Knowledge and Belief*, H. van Ditmarsch, J.Y. Halpern, W. van der Hoek and B. Kooi (eds), College Publications, 2015, pp. 299–368.

update. The chapter continues with three topics that naturally complement the setting of single steps of belief change: connections with probabilistic approaches to belief change, long-term temporal process structure including links with formal learning theory, and multi-agent scenarios of information flow and belief revision in games and social networks. We end with a discussion of alternative approaches, further directions, and windows to the broader literature, while links with relevant philosophical traditions are discussed throughout.

Human cognition and action involve a delicate art of living dangerously. Beliefs are crucial to the way we plan and choose our actions, even though our beliefs can be very wrong and refuted by new information. What keeps the benefits and dangers in harmony is our ability to revise beliefs as the need arises. In this chapter, we will look at the logical structure of belief revision, and belief change generally. But before we can do this, we need background of two kinds: (a) the pioneering AGM approach in terms of postulates governing belief revision which showed that this process has clear formal structures regulating its behavior, and (b) the basics of dynamic-epistemic logics of information flow which showed that change of attitudes for agent and the events triggering such changes are themselves susceptible to exact logical analysis. This is what we will provide in the first two sections of this chapter. With this material in place, Section 7.3 will then start our main topic, the logical treatment of belief revision.

7.1 Basics of belief revision

7.1.1 The AGM account of belief revision

What happens when an agent is confronted with a new fact φ that goes against her prior beliefs? If she is to accept the new fact φ and maintain a consistent set of beliefs, she will have to give up some of her prior beliefs. But which of her old beliefs should she give up? More generally, what policy should she follow to revise her beliefs? As we will see in this chapter, several answers to this question are possible. The standard answer in the literature says that our agent should accept the new fact and at the same time maintain as many as possible of her old beliefs without arriving at a contradiction. Making this more precise has been the driving force behind Belief Revision Theory. Standard Belief Revision Theory, also called *AGM theory* (after the pioneering authors Alchourrón, Gärdenfors and Makinson) has provided us with a series of "rationality conditions", that are meant to precisely govern the way in which a rational agent should revise her beliefs.

AGM theory The AGM theory of belief revision is built up from three basic ingredients: 1) the notion of a *theory* (or "belief set") T, which is a logically closed set of sentences $\{\psi, \gamma ...\}$ belonging to a given language $\mathcal{L}$; 2) the *input of new information*, i.e., a syntactic formula φ; and 3) a *revision operator* $*$ which is a map associating a theory $T * \varphi$ to each pair (T, φ) consisting of a theory T and an input sentence φ. The construct $T * \varphi$ is taken to represent the agent's new theory after learning φ. Hence $T * \varphi$ is the agent's new set of beliefs, given that the initial set of beliefs is T and that the agent has learnt that φ.

Expansion The AGM authors impose a number of postulates or rationality conditions on the revision operation $*$. To state these postulates, we first need an auxiliary *belief expansion operator* $+$, that is often considered an unproblematic form of basic update. Belief expansion is intended to model the simpler case in which the new incoming information φ does not contradict the agent's prior beliefs. The *expansion* $T+\varphi$ of T with φ is defined as the closure under logical consequence of the set $T \cup \{\varphi\}$. AGM provides a list of 6 postulates that exactly regulate the expansion operator, but instead of listing them here we will concentrate on belief revision. However, later on, we will see that even expansion can be delicate when complex epistemic assertions are added.

Revision Now, belief revision goes beyond belief expansion in its intricacies. It is regulated by the following famous AGM Belief Revision Postulates:

(1)	Closure	$T * \varphi$ is a belief set
(2)	Success	$\varphi \in T * \varphi$
(3)	Inclusion	$T * \varphi \subseteq T + \varphi$
(4)	Preservation	If $\neg\varphi \notin T$, then $T + \varphi \subseteq T * \varphi$
(5)	Vacuity	$T * \varphi$ is inconsistent iff $\vdash \neg\varphi$
(6)	Extensionality	If $\vdash \varphi \leftrightarrow \psi$, then $T * \varphi = T * \psi$
(7)	Subexpansion	$T * (\varphi \wedge \psi) \subseteq (T * \varphi) + \psi$
(8)	Superexpansion	If $\neg\psi \notin T * \varphi$, then $T * (\varphi \wedge \psi) \supseteq (T * \varphi) + \psi$.

These postulates look attractive, though there is more to them than meets the eye. For instance, while the success postulate looks obvious, in our later dynamic-epistemic logics, it is the most controversial one in this list. In a logical system allowing complex epistemic formulas, the truth value of the target formula can change in a revision step, and the Success Postulate would recommend incorporating a falsehood φ into the agent's theory T. One important case in which this can occur is when an introspective agent revises her beliefs on the basis of new information that refers to beliefs or higher-order beliefs (i.e., beliefs about beliefs). Because the AGM setting does not incorporate "theories about theories", i.e., it ignores an agent's higher-order beliefs, this problem is side-stepped. All the beliefs covered by AGM are so-called factual beliefs about ontic facts that do not refer to the epistemic state of the agent. However any logic for belief change that does allow explicit belief-operators in the language, will have to pay attention to success conditions for complex updates.

A final striking aspect of the Success Postulate is the heavy emphasis placed on the last incoming proposition φ, which can abruptly override long accumulated earlier experience against φ. This theme, too, will return later when we discuss connections with formal theories of inductive learning.

Contraction A third basic operation considered in AGM is that of *belief contraction* $T - \varphi$, where one removes a given assertion φ from a belief set T, while removing enough other beliefs to make φ underivable. This is harder than expansion, since one has to make sure that there is no other way within the new theory to derive the target formula after all. And while there is no unique way to construct a contracted theory, AGM prescribes the following formal postulates:

(1)	Closure	$T - \varphi$ is a belief set
(2)	Contraction	$(T - \varphi) \subseteq T$
(3)	Minimal Action	If $\varphi \notin T$, then $T - \varphi = T$
(4)	Success	If $\not\vdash \varphi$ then $\varphi \notin (T - \varphi)$
(5)	Recovery	If $\varphi \in T$, then $T \subseteq (T - \varphi) + \varphi$
(6)	Extensionality	If $\vdash \varphi \leftrightarrow \psi$, then $T - \varphi = T - \psi$
(7)	Min-conjunction	$T - \varphi \cap T - \psi \subseteq T - (\varphi \wedge \psi)$
(8)	Max-conjunction	If $\varphi \notin T - (\varphi \wedge \psi)$, then $T - (\varphi \wedge \psi) \subseteq T - \varphi$

Again, these postulates have invited discussion, with Postulate 5 being the most controversial one. The Recovery Postulate is motivated by the intuitive principle of minimal change, which prescribes that a contraction should remove as little as possible from a given theory T.

The three basic operations on theories introduced here are connected in various ways. A famous intuition is the Levi-identity

$$T * \varphi = (T - \neg\varphi) + \varphi$$

saying that a revision can be obtained as a contraction followed by an expansion.

An important result in this area is a theorem by Gärdenfors which shows that if the contraction operation satisfies postulates (1-4) and (6), while the expansion operator satisfies its usual postulates, then the revision operation defined by the Levi-identity will satisfy the revision postulates (1-6). Moreover, if the contraction operation satisfies the seventh postulate, then so does revision, and likewise for the eight postulate.

7.1.2 Conditionals and the Ramsey Test

Another important connection runs between belief revision theory and the logic of conditionals. The *Ramsey Test* is a key ingredient in any study of this link. In 1929, F.P. Ramsey wrote:

> "If two people are arguing 'If A, will B?' and are both in doubt as to A, they are adding A hypothetically to their stock of knowledge and arguing on that basis about B; so that in a sense 'If A, B' and 'If $A, \overline{B}$' are contradictories."

Clearly, this evaluation procedure for conditional sentences $A > B$ uses the notion of belief revision. Gärdenfors formalised the connection with the Ramsey Test as the following statement:

$$A > B \in T \text{ iff } B \in T * A$$

which should hold for all theories T and sentences A, B. In a famous impossibility result, he then showed that the existence of such Ramsey conditionals is essentially incompatible with the AGM postulates for the belief revision operator $*$. The standard way out of Gärdenfors' impossibility result is to weaken the axioms of $*$, or else to drop the Ramsey test.

Most discussions in this line are cast in purely syntactic terms, and in a setting of propositional logic. However, in section 7.3 we will discuss a semantic perspective which saves much of the intuitions underlying the Ramsey Test.

This is in fact a convenient point for turning to the modal logic paradigm in studying belief revision. Like we saw with belief expansion, it may help to first introduce a simpler scenario. The second part of this introductory section shows how modal logics can describe information change and its updates[1] in what agents know. The techniques found in this realm will then be refined and extended in our later treatment of belief revision.

7.2 Modal logics of belief revision

Starting in the 1980s, several authors have been struck by analogies between AGM revision theory and modal logic over suitable universes. Belief and related notions like knowledge could obviously be treated as standard modalities, while the dynamic aspect of belief change suggested the use of ideas from Propositional Dynamic Logic of programs or actions to deal with update, contraction, and revision. There is some interest in seeing how long things took to crystallise into the format used in this chapter, and hence we briefly mention a few of these proposals before introducing our final approach.

Propositional dynamic logic over information models Propositional Dynamic Logic (PDL) is a modal logic that has both static propositions φ and programs or actions π. It provides dynamic operators $[\pi]\varphi$ that one can use to reason about what will be true after an action takes place. One special operator of PDL is the "test of a proposition φ" (denoted as $\varphi?$): it takes a proposition φ into a program that tests if the current state satisfies φ. Using this machinery over tree-like models of successive information states ordered by inclusion, in 1989, van Benthem introduced dynamic operators that mirror the operations of AGM in a modal framework. One is the addition of φ (also called "update", denoted as $+\varphi$), interpreted as moving from any state to a minimal extension satisfying φ. Other operators included "downdates" $-\varphi$ moving back to the first preceding state in the ordering where φ is not true. Revision was defined via the Levi-identity. In a modification of this approach by de Rijke in the 1990s, these dynamic operators were taken to work on universes of theories.

[1] Our use of the term "update" in this chapter differs from a common terminology of "belief update" in AI, due to Katsuno and Mendelzon. The latter notion of update refers to belief change in a factually changing world, while we will mainly (though not exclusively) consider epistemic and doxastic changes but no changes of the basic ontic facts. This is a matter of convenience though, not of principle.

Dynamic doxastic logic over abstract belief worlds These developments inspired Segerberg to develop the logical system of Dynamic Doxastic Logic (DDL), which operates at a higher abstraction level for its models. DDL combines a PDL dynamics for belief change with a static logic with modalities for knowledge K and belief B. The main syntactic construct in DDL is the use of the dynamic modal operator $[*\varphi]\psi$ which reads "ψ holds after revision with φ", where $*\varphi$ denotes a relation (often a function) that moves from the current world of the model to a new one. Here φ and ψ were originally taken to be factual formulas only, but in later versions of DDL they can also contain epistemic or doxastic operators. This powerful language can express constructs such as $[*\varphi]B\psi$ stating that after revision with φ the agent believes ψ. In what follows, we will take a more concrete modal approach to DDL's abstract world, or state, changes involved in revision – but a comparison will be given in Section 7.9.1.

Degrees of belief and quantitative update rules In this chapter, we will mainly focus on qualitative logics for belief change. But historically, the next step were quantitative systems for belief revision in the style of Dynamic Epistemic Logic, where the operations change current models instead of theories or single worlds. Such systems were proposed a decade ago, using labelled operators to express degrees of belief for an agent. In 2003, van Ditmarsch and Labuschagne gave a semantics in which each agent has associated accessibility relations corresponding to labeled preferences, and a syntax that can express degrees of belief. Revision of beliefs with new incoming information was modeled using a binary relation between information states for knowledge and degrees of belief. A more powerful system by Aucher in 2003 had degrees of belief interpreted in Spohn ranking models, and a sophisticated numerical "product update rule" in the style of Baltag, Moss and Solecki (BMS, see Section 7.5.2 below) showing how ranks of worlds change under a wide variety of incoming new information.

Belief expansion via public announcement logic An early qualitative approach, due to van Ditmarsch, van der Hoek and Kooi in 2005, relates AGM belief expansion to the basic operation of public announcement in Dynamic Epistemic Logic. The idea is to work with standard relational modal models M for belief (in particular, these need not have a reflexive accessibility relation, since beliefs can be wrong), and then view the action of getting new information φ as a public announcement that takes M to its submodel consisting only of its φ-worlds. Thus, an act of belief revision is modeled by a transformation of some current epistemic or doxastic model. The system had some built-in limitations, and important changes were made later by van Benthem and Baltag & Smets to the models and update mechanism to achieve a general theory – but it was on the methodological track that we will follow now for the rest of this chapter.

Public announcement logic To demonstrate the methodology of Dynamic Epistemic Logic to be used in this chapter, we explain the basics of Public Announcement Logic (PAL). The language of PAL is built up as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [!\varphi]\varphi$$

Here we read the K_i-modality as the knowledge of agent i and we read the dynamic construct $[!\varphi]\psi$ as "ψ holds after the public announcement of φ". We think of announcements $!\varphi$ as public events where indubitable hard information that φ is the case becomes available to all agents simultaneously, whether by communication, observation, or yet other means. In what follows, we define and study the corresponding transformations of models, providing a constructive account of how information changes under this kind of update.

Semantics for PAL We start with standard modal models $M = (W, R_i, V)$ where W is a non-empty set of possible worlds. For each agent i, we have an epistemic accessibility relation R_i, while V is a valuation which assigns sets of possible worlds to each atomic sentence p. The satisfaction relation can be introduced as usual in modal logic, making the clauses of the non-dynamic fragment exactly the standard ones for the multi-agent epistemic logic $S5$. For the case of knowledge (only an auxiliary initial interest in this chapter), we take R_i to be an equivalence relation, so that the underlying base logic is a multi-agent version of the modal logic $S5$. We now concentrate on the dynamics.

The clause for the dynamic modality goes as follows:

$$(M, w) \models [!\varphi]\psi \text{ iff } (M, w) \models \varphi \text{ implies } (M|\varphi, w) \models \psi$$

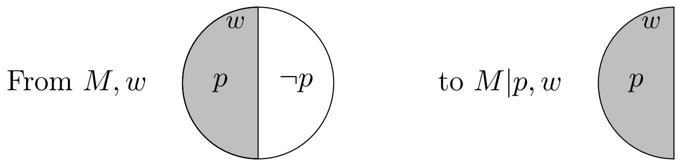
where $M|\varphi = (W', R'_i, V')$ is obtained by relativising the model M with φ as follows (here, as usual, $[\![\varphi]\!]_M$ denotes the set of worlds in M where φ is true):

$$W' = [\![\varphi]\!]_M$$

$$R'_i = R_i \cap ([\![\varphi]\!]_M \times [\![\varphi]\!]_M)$$

$$V'(p) = V(p) \cap [\![\varphi]\!]_M$$

Example 7.1 (Public announcement by world elimination)
Figure 7.1 illustrates the update effect of a public announcement in the state w of model M such that in model $M|p$ only the p-worlds survive:



⊣

Figure 7.1: A public announcement of p.

Proof system for PAL We start with the proof system for $S5$, including all its standard axioms and rules (including replacement of provable equivalents), and in addition we have the following recursion axioms:

$$\begin{array}{rcl} [!\varphi]p & \leftrightarrow & (\varphi \to p) \\ [!\varphi]\neg\psi & \leftrightarrow & (\varphi \to \neg[!\varphi]\psi) \\ [!\varphi](\psi \wedge \gamma) & \leftrightarrow & ([!\varphi]\psi \wedge [!\varphi]\gamma) \\ [!\varphi]K_i\psi & \leftrightarrow & (\varphi \to K_i[!\varphi]\psi) \end{array}$$

Completeness The axiomatic system for PAL is sound and complete. One can easily show this by using the recursion axioms to translate every sentence that contains a dynamic modality to an equivalent one without it, and then using the completeness of the static base logic.

Public announcements and belief expansion The effect of a public announcement $!\varphi$ on a given model is of a very specific type: all non-φ worlds are deleted. One can easily see that when models contract under truthful announcements, the factual knowledge of the agent expands. As we stated already, these restrictions were lifted in the work of van Benthem in 2007 and by Baltag & Smets in 2008, who deal with arbitrary updates of both plain and conditional beliefs, actions of revision and contraction, as well as other triggers for all these than public announcements. The latter systems will be the heart of this chapter, but before we go there, we also need to discuss the optimal choice of an underlying base logic in more detail than we have done so far.

Summary Modal logic approaches to belief revision bring together three traditions: 1) modal logics for static notions of knowledge and belief, 2) the AGM theory of belief revision, and 3) the modal approach to actions of Propositional Dynamic Logic. Merging these ideas opens the door to a study of knowledge update and belief change in a standard modal framework, without having to invent non-standard formalisms, allowing for smooth insertion into the body of knowledge concerning agency that has been accumulated in the modal paradigm.

7.3 Static base logics

7.3.1 Static logic of knowledge and belief

In line with the literature in philosophical logic, we want to put knowledge and belief side by side in a study of belief change. But how to do this, requires some thinking about the best models. In particular, Hintikka's original models with a binary doxastic accessibility relation that drive such well-known systems as "KD45" are not what we need – as we shall see in a moment.

Reinterpreting PAL One easy route tries to reinterpret the dynamic-epistemic logic PAL that we have presented in the previous section. Instead of knowledge, we now read the earlier K_i-operators as beliefs, placing no constraints on the accessibility relations: using just pointed arrows. One test for such an approach is that it must be possible for beliefs to be wrong:

Example 7.2 (A mistaken belief)
Consider the model in Figure 7.2 with two worlds that are epistemically accessible to each other via dashed arrows, but where the pointed arrow is the only belief relation. Here, in the actual world to the left, marked in black, the proposition p is true, but the agent mistakenly believes that $\neg p$.

With this view of doxastic modalities, PAL works exactly as before. But as pointed out by van Benthem around 2005, there is a problem, i.e., that of overly drastic belief change.

Consider an announcement $!p$ of the true fact p. The PAL result is the one-world model where p holds, with the inherited empty doxastic accessibility relation.

But on the universal quantifier reading of belief, this means the following: the agent believes that p, but also that $\neg p$, in fact $B\bot$ is true at such an end-point. Clearly, this is not what we want: agents who have their beliefs contradicted would now be shattered and they would start believing anything. ⊣

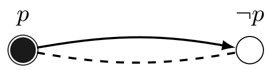


Figure 7.2: A mistaken belief.

While some ad-hoc patches persist in the literature, a better way is to change the semantics to allow for more intelligent responses, by using more structured models for conditional beliefs, as we shall show now.

7.3.2 Plausibility models

A richer view of belief follows the intuition that an agent believes the things that are true, not in all her accessible worlds, but only in those that are "best" or most relevant to her. Static models for this setting are easily defined.

Definition 7.1 (Epistemic plausibility models)
Epistemic plausibility models are structures $M = (W, \{\sim_i\}_{i\in I}, \{\leq_i\}_{i\in I}, V)$, where W is a set of states or possible worlds. The epistemic accessibility $\{\sim_i\}_{i\in I}$ relations for each agent $i \in I$ are equivalence relations. The family of plausibility relations $\{\leq_i\}_{i\in I}$ consists of binary comparison relations for agents and is read as follows, $x \leq_i y$ if agent i considers y at least as plausible as x. The plausibility relations are assumed to satisfy two conditions: (1) $\leq_i$-comparable states are $\sim_i$-indistinguishable (i.e. $s \leq_i t$ implies $s \sim_i t$) and (2) the restriction of each plausibility relation $\leq_i$ to each $\sim_i$-equivalence class is a well-preorder.[2] ⊣

[2]Here, a "preorder" is a reflexive and transitive binary relation. A "well-preorder" over W is a preorder guaranteeing that every non-empty subset of W has maximal (most plausible) elements. Note also that the conditions (1) and (2) are useful in an epistemic-doxastic context but can be relaxed in various ways to yield a more general setting. We will return to this point below.

Notation We use the notation $Max_{\leq_i}P$ for the set of $\leq_i$-maximal elements of a given set $P \subseteq W$ and use this to denote the "most plausible elements of P under the given relation". There are two other relations that are useful to name: the strict plausibility relation $s <_i t$ iff $s \leq_i t$ but $t \not\leq_i s$, and the equiplausibility relation $s \cong_i t$ iff both $s \leq_i t$ and $t \leq_i s$. We denote by $s(i)$ the $\sim_i$-equivalence class of s, i.e. $s(i) := \{t \in W : s \sim_i t\}$

Simplifying the setting The definition of epistemic plausibility models contains superfluous information. The epistemic relation $\sim_i$ can actually be recovered from the plausibility relation $\leq_i$ via the following rule:

$$s \sim_i t \text{ iff either } s \leq_i t \text{ or } t \leq_i s$$

provided that the relation $\leq_i$ is connected.[3] This makes two states epistemically indistinguishable for an agent i iff they are *comparable* with respect to $\leq_i$. Accordingly, the most economic setting are the following simplified semantic structures that we will use in the remainder of this chapter.

Definition 7.2 (Plausibility models)
A *plausibility model* $M = (W, \{\leq_i\}_{i \in I}, V)$ consists of a set of possible worlds W, a family of locally well-preordered relations $\leq_i$ and a standard valuation V.[4] ⊣

Fact 7.1
There is a bijective correspondence between Epistemic Plausibility Models and the above Plausibility Models. ⊣

Baltag & Smets show in 2006 how every plausibility model can be canonically mapped into an epistemic plausibility model and vice versa.[5]

Generalisation to arbitrary preorders The above simplification to plausibility models only works because all epistemically accessible worlds are comparable by plausibility and all plausibility comparisons are restricted to epistemically accessible worlds. These conditions are restrictive, though they can be justified in the context of belief revision.[6] However, many authors have also used plausibility models with non-connected orderings, allowing for genuinely incomparable (as opposed to indifferent) situations. With a few modifications, what we present in this chapter also applies to this more general setting that reaches beyond belief revision theory.

In this chapter we will use plausibility models as our main vehicle.

[3] A relation is "connected" if either $s \leq_i t$ or $t \leq_i s$ holds, for all $s, t \in W$.

[4] Here a *locally well-preordered* relation $\leq_i$ is a preorder whose restriction to each corresponding comparability class $\sim_i$ is well-preordered. In case the set W is *finite*, a locally well-preordered relation becomes a *locally connected preorder*: a preorder whose restriction to any comparability class is connected.

[5] In one direction this can be shown by defining the epistemic relation as $(\sim_i := \leq_i \cup \geq_i)$. In the other direction, all that is needed is a map that "forgets" the epistemic structure.

[6] One advantage is that connected pre-orders ensure the AGM rule of "rational monotonicity", capturing the intuitive idea that an agent who believes two propositions p and q will still hold a belief in q after revision with $\neg p$.

Languages and logics

One can interpret many logical operators in plausibility models. In particular, knowledge can be interpreted as usual. However, there is no need to stick to just the knowledge and belief operators handed down by the tradition. First of all, plain belief $B_i\varphi$ is a modality interpreted as follows:

Definition 7.3 (Belief as truth in the most plausible worlds)
In plausibility models, we interpret *belief* by putting $M, s \models B_i\varphi$ iff $M, t \models \varphi$ for all worlds t in $s(i)$ that are in $Max_{\leq_i}[\![\varphi]\!]$. $\dashv$

But information flow and action also involve richer conditional beliefs $B_i^\psi\varphi$, with the intuitive reading that, conditional on ψ, the agent believes that φ. Roughly speaking, conditional beliefs "pre-encode" the beliefs that agents would have if they were to learn certain things expressed by the restricting proposition.

Definition 7.4 (Interpreting conditional beliefs)
In plausibility models, $M, s \models B_i^\psi\varphi$ iff $M, t \models \varphi$ for all worlds t that are maximal for the order $x \leq_i y$ in the set $\{u \mid M, u \models \psi\} \cap s(i)$. $\dashv$

Plain belief $B_i\varphi$ can now be recovered as the special case $B_i^T\varphi$ with a trivially true condition T. It can be shown that conditional belief is not definable in terms of plain belief, so we have obtained a genuine language extension.

Example 7.3 (Conditional beliefs depicted)
Figure 7.3 shows a plausibility model containing both the plausibility relations and the indistinguishability relations (represented via the dashed arrows) for a given agent. To keep the picture simple, we draw neither reflexive arrows nor arrows that can be obtained via transitivity. In every world, the agent believes p and q_3, i.e. $B_i(p \wedge q_3)$. Note that while this agent currently holds a true belief in p, her belief in p is rather fragile because it can easily be given up when new information is received. Indeed, conditional on the information $\neg q_3$, the agent would believe that $\neg p$ was the case, i.e. $B_i^{\neg q_3}\neg p$. $\dashv$

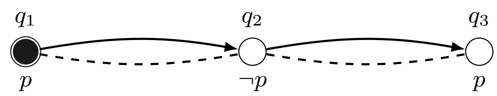


Figure 7.3: Illustrating conditional beliefs.

Epistemic-doxastic introspection We already noted that in plausibility models, epistemic accessibility is an equivalence relation, and plausibility a pre-order over the equivalence classes, the same as viewed from any world inside that class. This reflects the fact that in such models, all agents know their beliefs. Formally, it is easy to see that the following axiom is valid:

$$B_i\varphi \to K_iB_i\varphi \qquad \text{Epistemic-Doxastic Introspection}$$

This assumption is of course debatable, since not all doxastic agents may have this type of introspective capacity. Abandoning this assumption via ternary orderings will be explored briefly in Section 7.3.2.

As an illustration of the semantic framework introduced here, we return to an earlier issue about belief revision.

The Ramsey Test revisited Recall the Ramsey Test and Gärdenfors' impossibility result from Section 7.1. In 2010, Baltag & Smets gave a semantic re-analysis that uses modalities for knowledge, belief and conditional belief. In this setting, the main question becomes: "Are there any truth conditions for the conditional $A > B$ that are compatible with the Ramsey Test – given the usual modal semantics for belief, and some reasonable semantics for belief revision?". It can be shown that, if the Ramsey test is to hold for *all* theories (including those representing future belief sets, after possible revisions) and if some reasonable rationality conditions are assumed (such as full introspection of beliefs and of dispositions to believe, plus unrevisability of beliefs that are known to be true), then the answer is "no". The reason is this: the Ramsey Test treats conditional beliefs about beliefs in the same way as hypothetical beliefs about facts. The test would succeed only if, when making a hypothesis, agents revise their beliefs about their own belief revision in the same way as they revise their factual beliefs. But the latter requirement is inconsistent with the restrictions posed by introspection. Introspective agents know their own hypothetical beliefs, and so cannot accept hypotheses that go against their knowledge. Thus, beliefs about one's own belief revision policy cannot be revised.

But this is not the end of the story, and in a sense, the logics to be presented later on show a "Possibility Result". A dynamic revision of the sort pursued in the coming sections, that represents agents' revised beliefs about the situation after the revision, can consistently satisfy a version of the Ramsey Test.

World-based plausibility models

Plausibility models as defined here have uniform plausibility relations that do not vary across epistemically accessible worlds. This reflects an intuition that agents know their own plain and conditional beliefs. However, it is possible, at a small technical cost, to generalise our treatment to ternary world-dependent plausibility relations, and such relations are indeed common in current logics for epistemology. Stated equivalently, the assumption of epistemic-doxastic introspection can be changed in models that keep track of the different beliefs of the agent in every possible world.

Definition 7.5 (World-based plausibility models)
World-based plausibility models are structures of the form $M = (W, \{\sim_i\}_{i\in I}, \{\leq_i^s\}_{i\in I}, V)$ where the relations $\sim_i$ stand for epistemic accessibility as before, but the $(\leq_i^s)$ are ternary comparison relations for agents that read as: $x \leq_i^s y$ if, in world s, agent i considers y at least as plausible as x. ⊣

Models like this occur in conditional logic, logics of preference, and numerical graded models for beliefs. One can again impose natural conditions on ternary plausibility relations, such as reflexivity, or transitivity. Adding connectedness yields the well-known nested spheres from the semantics of conditional logic, with pictures of a line of equiplausibility clusters, or of concentric circles. But there are also settings that need a fourth option of incomparability: $\neg s \leq t \wedge \neg t \leq s$. This happens, for instance, when comparing worlds with conflicting criteria for preference. And sometimes also, non-connected pre-orders or partially ordered graphs are a mathematically more elegant approach.

Essentially the same truth conditions as before for plain and conditional beliefs work on these more general models, but it is easy to find models now where agents are not epistemically introspective about their own beliefs.

One thing to note in the logic of world-based plausibility models is how their treatment of conditional beliefs is very close to conditional logic:

Digression on conditionals Recall that conditional beliefs pre-encode beliefs we would have if we were to learn new things. A formal analogy is a well-known truth condition from conditional logic. A conditional $\varphi \Rightarrow \psi$ says that ψ is true in the closest worlds where φ is true, along some comparison order on worlds.[7] Thus, results from conditional logic apply to doxastic logic. For instance, on reflexive transitive plausibility models, we have a completeness theorem whose version for conditional logic is due to Burgess in 1981 and Veltman in 1985:

Theorem 7.1
The logic of conditional belief $B^{\psi}\varphi$ on world-based plausibility models is axiomatised by the laws of propositional logic plus obvious transcriptions of the following principles of conditional logic:

(a) $\varphi \Rightarrow \varphi$
(b) $\varphi \Rightarrow \psi$ implies $\varphi \Rightarrow \psi \vee \chi$
(c) $\varphi \Rightarrow \psi, \varphi \Rightarrow \chi$ imply $\varphi \Rightarrow \psi \wedge \chi$,
(d) $\varphi \Rightarrow \psi, \chi \Rightarrow \psi$ imply $(\varphi \vee \chi) \Rightarrow \psi$,
(e) $\varphi \Rightarrow \psi, \varphi \Rightarrow \chi$ imply $(\varphi \wedge \psi) \Rightarrow \chi$. ⊣

Richer attitudes: safe and strong belief

We now return to uniform plausibility models. In epistemic reality, agents have a rich repertoire of attitudes concerning information beyond just knowledge and belief, such as being certain, being convinced, assuming, etcetera. Among all options in this plethora, of special interest to us are notions whose definition has a dynamic intuition behind it. The following notion makes particular sense, intermediate between knowledge and belief. It is a modality of true belief which is stable under receiving new true information:

Definition 7.6 (Safe belief)
The modality of *safe belief* $\Box_i\varphi$ is defined as follows: $M, s \models \Box_i\varphi$ iff for all worlds t in the epistemic range of s with $t \geq_i s, M, t \models \varphi$. Thus, φ is true in all epistemically accessible worlds that are at least as plausible as the current one. ⊣

[7] On infinite models, this clause needs some modifications, but we omit details here.

The modality $\Box_i\varphi$ is stable under hard information updates, at least for factual assertions φ that do not change their truth value as the model changes. In fact, it is just the universal base modality $[\leq_i]\varphi$ for the plausibility ordering. In what follows, we will make safe belief part of the static doxastic language, treating it as a pilot for a richer theory of attitudes in the background. Pictorially, one can think of this as illustrated in the following example:

Example 7.4 (Three degrees of doxastic strength)
Consider the model in Figure 7.4, with the actual world in the middle.

$K_i\varphi$ describes what the agent knows in an absolute sense: φ must be true in all worlds in the epistemic range, less or more plausible than the current one. In Figure 7.4, the agent knows q and all tautologies in this way. $\Box_i p$ describes her safe beliefs in further investigation: p is true in all worlds of the model from the middle towards the right. Thus, we have a safe belief $\Box_i p$ at the black node and at all more plausible worlds which is not knowledge: $\neg K_i p$. Finally, $B_i\varphi$ describes the most fragile thing: her beliefs as true in all worlds in the current rightmost position. In the model of Figure 7.4, the agent holds a true belief in r, i.e. $B_i r$, which is not a safe belief, and a fortiori, not knowledge. ⊣

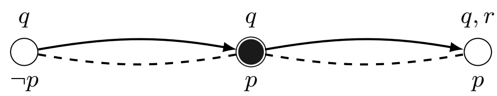


Figure 7.4: Illustrating safe beliefs.

In addition to its intuitive merits, safe belief simplifies the logic when used as a technical device, since it can define other notions of interest:

Fact 7.2
The following holds on finite connected epistemic plausibility models:

(a) Safe belief can define its own conditional variant,

(b) With a knowledge modality, safe belief can define conditional belief. ⊣

Indeed, conditional belief $B_i^\varphi\psi$ is equivalent to the modal statement

$$\widetilde{K}_i\varphi \to \widetilde{K}_i(\varphi \wedge \Box_i(\varphi \to \psi))$$

where $\widetilde{K}_i\varphi := \neg K_i\neg\varphi$ is the diamond modality for K_i. Alternative versions use modalities for the strict plausibility relation. This definition needs reformulation on non-connected plausibility models, where the following formula works:

$$K_i((\psi \wedge \varphi) \to \widetilde{K}_i(\psi \wedge \varphi \wedge \Box_i(\psi \to \varphi)))$$

Failure of negative introspection Safe belief also has less obvious features. For instance, since its accessibility relation is not Euclidean, it fails to satisfy Negative Introspection. The reason is that safe belief mixes purely epistemic information with procedural information about what may happen later on in the current process of inquiry. In the above example of Figure 7.4 it is easy to see that in the left most node $\neg\Box_i p$ holds while $\Box_i\neg\Box_i p$ does not.

Strong belief Safe belief is not the only natural new doxastic notion of interest in plausibility models. Another important doxastic attitude is this:

Definition 7.7 (Strong belief)
The proposition φ is a *strong belief* at a state w in a given model M iff φ *is epistemically possible* and also, all epistemically possible φ-states at w are strictly more plausible than all epistemically possible non-φ states. ⊣

Example 7.5 (Strong, safe, and plain belief)
Consider the model in Figure 7.5, with the actual world in the middle. In this model, the agent holds a strong belief in p but not in q. She holds a safe belief at the actual world in p but not in q, and a mere belief in $\neg q \wedge p$. ⊣

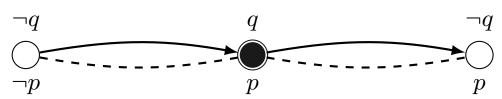


Figure 7.5: Illustrating strong beliefs.

Summary The above notions of plain, conditional, safe and strong belief can be found in various places in the logical, philosophical and economical literature. Yet there are many further epistemic and doxastic notions in our cognitive repertoire as encoded in natural language, and more notions can be interpreted on plausibility models than what we have shown.

But even with what we have shown, this section will have established its point that plausibility models are a natural way of representing a wide range of interesting doxastic notions. Hence they form a good basis for the dynamics of belief change, the main topic of this chapter, to which we now turn.

7.4 Belief revision by model transformations

Knowledge and varieties of belief express attitudes of an agent in its current epistemic-doxastic state. But belief revision is about changes in such states, and in this section, we turn to modeling this crucial feature, first on the analogy of the earlier Public Announcement Logic, and after that, with a new mechanism of changing plausibility orderings.

7.4.1 From knowledge to belief, hard and soft information

The original dynamic-epistemic logics such as PAL deal with knowledge, information and truth. Now, however, we also look at beliefs that agents form beyond the hard information that they possess. The result is a *tandem* of "jumping ahead" and "self-correction": agents believe more then they know, but this creative ability has to be kept in check by a capacity for self-correction, or stated more positively: for learning by trial and error.

Hard versus soft information With this new setting comes a richer dynamics of information. A public announcement $!\varphi$ of a fact φ was an event of *hard information* that changes irrevocably what an agent knows. Such events of hard information may also change beliefs. But in line with the greater softness and flexibility of belief over knowledge, there are also events that convey *soft information*, affecting agents' beliefs in less radical ways than world elimination.

The earlier plausibility models are well-suited for modeling the dynamics of both hard and soft information through suitable update operations.

7.4.2 Belief change under hard information

Our first dynamic logic of belief revision puts together the logic PAL of public announcement, a paradigmatic event producing hard information, with our static models for conditional belief, following the methodology explained above.

A complete axiomatic system For a start, we locate the key recursion axiom that governs belief change under hard information:

Fact 7.3
The following formula is valid for beliefs after hard information:

$$[!\varphi]B_i\psi \leftrightarrow (\varphi \to B_i^{\varphi}([!\varphi]\psi))$$

This is still somewhat like the PAL recursion axiom for knowledge. But the conditional belief in the consequent does not reduce to any obvious conditional plain belief of the form $B(\varphi \to ...)$. Therefore, we also need to state which conditional beliefs are formed after new information.[8]

Theorem 7.2
The logic of conditional belief under public announcements is axiomatised completely by (a) any complete static logic for the model class chosen, (b) the PAL recursion axioms for atomic facts and Boolean operations, (c) the following new recursion axiom for conditional beliefs:

$$[!\alpha]B_i^{\psi}\varphi \leftrightarrow (\alpha \to B_i^{\alpha\wedge[!\alpha]\psi}[!\alpha]\varphi).$$

To get a joint version with knowledge, we just combine with the PAL axioms. ⊣

[8]Classical belief revision theory says only how new plain beliefs are formed. The resulting "Iteration Problem" for consecutive belief changes cannot arise in a logic that covers revising conditional beliefs.

The Ramsey Test once more The above recursion axioms distinguish between formulas ψ before the update and after: $[!\varphi]\psi$. In this context, we can revisit our earlier discussion on why the existence of Ramsey conditionals of the form $A > B$ is incompatible with the AGM postulates.[9] Recall that the Ramsey Test says: "A conditional proposition $A > B$ is true, if, after adding A to your current stock of beliefs, the minimal consistent revision implies B." In our logic, this is ambiguous, as the consequent B *need no longer say the same thing after the revision*. As was already noted earlier, in a truly dynamic setting the truth-value of epistemic and doxastic sentences can change. Even so, the above axioms become Ramsey-like for the special case of factual propositions ψ without modal operators, that do not change their truth value under announcement:

$$[!\varphi]B_i\psi \leftrightarrow (\varphi \to B_i^{\varphi}\psi)$$
$$[!\varphi]B_i^{\alpha}\psi \leftrightarrow (\alpha \to B_i^{\varphi\wedge\alpha}\psi)$$

The reader may worry how this can be the case, given the Gärdenfors Impossibility Result. The nice thing of our logic approach, however, is that every law we formulate is sound. In other words, Ramsey-like principles do hold, provided we interpret the modalities involved in the way we have indicated here.[10]

Belief change under hard update links to an important theme in agency: variety of attitudes. We resume an earlier theme from the end of Section 7.3.

7.4.3 From dynamics to statics: safe and strong belief

Scenarios with hard information, though simple, contain some tricky cases:

Example 7.6 (Misleading with the truth)
Consider again the model in Figure 7.3 where the agent believed that p, which was indeed true in the actual world to the far right. This is a correct belief for a wrong reason: the most plausible world is not the actual world. For convenience, we assumed that each world verifies a unique proposition letter q_i.

Now giving the true information that we are not in the final world (i.e., the announcement $!\neg q_3$) updates the model to the one shown in Figure 7.6, in which the agent believes mistakenly that $\neg p$.[11] $\dashv$

Example 7.6, though tricky, is governed precisely by the complete logic of belief change under hard information. This logic is the one whose principles were stated before, with the following simple recursion law added for safe belief:

$$[!\varphi]\Box_i\psi \leftrightarrow (\varphi \to \Box_i(\varphi \to [!\varphi]\psi))$$

[9] As before, we use the notation $>$ to denote a binary conditional operator.

[10] In line with this, a weaker, but arguably the correct, version of the Ramsey test offers a way out of the semantic impossibility result. We must restrict the validity of the Ramsey test only to "theories" T that correspond to actual belief sets (in some possible world s) about the current world (s itself), excluding the application of the test to already revised theories.

[11] Cases like this occur independently in philosophy, computer science and game theory.

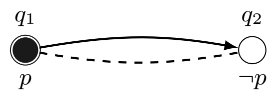


Figure 7.6: Updated belief model.

In fact, this axiom for safe belief under hard information implies the earlier more complex-looking one for conditional belief, by unpacking the modal definition of Section 7.3.2 and applying the relevant recursion laws. This shows once more how safe belief can simplify the total logic of belief and belief change.[12]

7.4.4 Belief change under soft information: radical upgrade

Soft information and plausibility change We have seen how a hard attitude like knowledge or a soft attitude like belief changes under hard information. But an agent can also take incoming signals in a softer manner, without throwing away options forever. Then public announcement is too strong:

Example 7.7 (No way back)
Consider the earlier model in Figure 7.6 where the agent believed that $\neg p$, though p was in fact the case. Publicly announcing p removes the $\neg p$-world, making any later belief revision reinstating $\neg p$ impossible. $\dashv$

We need a mechanism that just makes incoming information P more plausible, without removing the $\neg P$ worlds. There are many sources for this. Here is one from the work of Veltman in the 1990s on default rules in natural language.

Example 7.8 (Default reasoning)
A default rule $A \Rightarrow B$ of the form "A's are normally B's" does not say that all A-worlds must be B-worlds. Accepting it just makes counter-examples to the rule (the $A \wedge \neg B$-worlds) less plausible. This "soft information" does not eliminate worlds, it changes their order. More precisely, a triggering event that makes us believe that φ need only rearrange worlds making the most plausible ones φ: by "promotion" or "upgrade" rather than by elimination of worlds. $\dashv$

Thus, in our models $M = (W, \{\sim\}_i, \{\leq_i\}_i, V)$, we must now change the plausibility relations $\leq_i$, rather than the world domain W or the epistemic accessibilities $\sim_i$. Indeed, rules for plausibility change have been considered in earlier semantic "Grove models" for AGM-style belief revision as different policies that agents can adopt toward new information. We now show how the above dynamic logics deal with them in a uniform manner.

Radical revision One very strong, but widely used revision policy effects an upheaval in favor of some new proposition φ:

[12] A recursion axiom can also be found for strong belief, though we need to introduce a stronger conditional variant there. Moreover, safe and strong belief also yield natural recursion axioms under the more general dynamic operations to be discussed in the following sections.

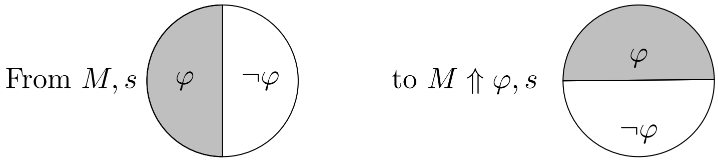


Figure 7.7: A radical revision step: all φ-worlds are moved to the top.

Definition 7.8 (Radical upgrade)
A *radical* (also called "lexicographic") *upgrade* $\Uparrow\varphi$ changes the current order $\leq_i$ between worlds in the given model M, s to a new model $M{\Uparrow}\varphi, s$ as follows: all φ-worlds in the current model become better than all $\neg\varphi$-worlds, while, within those two zones, the old plausibility order remains. $\dashv$

Language and logic With this definition in place, our earlier methodology applies. Like we did with public announcement, we introduce a corresponding "upgrade modality" in our dynamic doxastic language:

$$M, s \models [\Uparrow\varphi]\psi \text{ iff } M{\Uparrow}\varphi, s \models \psi$$

The resulting system can be axiomatised completely, in the same style as for our dynamic logics with hard information change. Here is a complete account of how agents' beliefs change under soft information, in terms of the key recursion axiom for changes in conditional belief under radical revision:

Theorem 7.3
The dynamic logic of lexicographic upgrade is axiomatised completely by

(a) any complete axiom system for conditional belief on the static models, plus

(b) the following recursion axioms[13]:

$$\begin{array}{lll}
[\Uparrow\varphi]q & \leftrightarrow & q \quad \text{for all atomic proposition letters } q \\
[\Uparrow\varphi]\neg\psi & \leftrightarrow & \neg[\Uparrow\varphi]\psi \\
[\Uparrow\varphi](\psi \wedge \alpha) & \leftrightarrow & [\Uparrow\varphi]\psi \wedge [\Uparrow\varphi]\alpha \\
[\Uparrow\varphi]K_i\varphi & \leftrightarrow & K_i[\Uparrow\varphi]\varphi \\
[\Uparrow\varphi]B_i^{\alpha}\psi & \leftrightarrow & (\widetilde{K}_i(\varphi \wedge [\Uparrow\varphi]\alpha) \wedge B_i^{\varphi \wedge [\Uparrow\varphi]\alpha}[\Uparrow\varphi]\psi) \vee \\
 & & (\neg\widetilde{K}_i(\varphi \wedge [\Uparrow\varphi]\alpha) \wedge B_i^{[\Uparrow\varphi]\alpha}[\Uparrow\varphi]\psi)
\end{array}$$

For plain beliefs $B_i\varphi$ with $\alpha =$ 'True', things simplify to:

$$[\Uparrow\varphi]B_i\psi \leftrightarrow (\widetilde{K}_i\varphi \wedge B_i^{\varphi}[\Uparrow\varphi]\psi) \vee (\neg\widetilde{K}_i\varphi \wedge B_i[\Uparrow\varphi]\psi)$$

And here is the simplified recursion axiom for factual propositions:

$$[\Uparrow\varphi]B_i^{\alpha}\psi \leftrightarrow ((\widetilde{K}_i(\varphi \wedge \alpha) \wedge B_i^{\varphi\wedge\alpha}\psi) \vee (\neg\widetilde{K}_i(\varphi \wedge \alpha) \wedge B_i^{\alpha}\psi))$$

Things get easier again with our earlier safe belief:

---

[13]Here, as before, $\widetilde{K}_i$ is the dual existential epistemic modality $\neg K_i \neg$.

Fact 7.4
The following recursion axiom is valid for safe belief under radical revision:

$$[\Uparrow\varphi]\Box_i\psi \leftrightarrow (\varphi \wedge \Box_i(\varphi \to [\Uparrow\varphi]\psi)) \vee (\neg\varphi \wedge \Box_i(\neg\varphi \to [\Uparrow\varphi]\psi) \wedge K_i(\varphi \to [\Uparrow\varphi]\psi))$$

Static pre-encoding We have shown now how complete modal logics exist for belief revision mechanisms, without any need for special-purpose formalisms. But our design in terms of recursion laws has a striking side effect, through its potential for successive reduction of dynamic modalities. As was the case with knowledge in PAL, our analysis of belief change says that any statement about epistemic-doxastic effects of hard or soft information is encoded in the initial model: the epistemic present contains the epistemic future. While this looks appealing, the latter feature may be too strong in the end.

In Section 7.7 of this chapter, we will look at extended models for belief revision encoding global informational procedures ("protocols"), that preserve the spirit of our recursion laws, but without a reduction to the static base language.

7.4.5 Conservative upgrade and other revision policies

The preceding logic was a proof-of-concept. Radical revision is just one way of taking soft information. Here is a well-known less radical policy for believing a new proposition. The following model transformation puts not all φ-worlds on top, but just the most plausible φ-worlds.

Definition 7.9 (Conservative plausibility change)
A *conservative upgrade* $\uparrow\varphi$ replaces the current order $\leq_i$ in a model M by the following: the best φ-worlds come on top, but apart from that, the old plausibility order of the model remains.[14] ⊣

The complete dynamic logic of conservative upgrade can be axiomatised in the same style as radical upgrade, this time with the following valid key recursion axiom for conditional belief:

$$[\uparrow\varphi]B_i^{\alpha}\psi \leftrightarrow (B_i^{\varphi}\neg[\uparrow\varphi]\alpha \wedge B_i^{[\uparrow\varphi]\alpha}[\uparrow\varphi]\psi) \vee (\neg B_i^{\varphi}\neg[\uparrow\varphi]\alpha \wedge B_i^{\varphi\wedge[\uparrow\varphi]\alpha}[\uparrow\varphi]\psi)$$

Clearly, this is a rather formidable principle, but then, there is no hiding the fact that belief revision is a delicate process, while it should also be kept in mind that recursion axioms like this can often be derived from general principles.

Variety of revision policies Many further changes in plausibility order can happen in response to an incoming signal. This reflects the many belief revision policies in the literature.[15] The same variety occurs in other realms of ordering

[14]Technically, $\uparrow\varphi$ is a special case of radical revision: $\Uparrow(best(\varphi))$.

[15]Maybe "policy" is the wrong term, as it suggests a persistent habit over time. But our events describe local responses to particular inputs. Moreover, speech act theory has a nice distinction between information per se (what is said) and the *uptake*, how a recipient reacts. In that sense, the softness of our scenarios is in the response, rather than in the signal itself.

change, such as the preference changes induced by the many deontic speech acts in natural language such as hard commands or soft suggestions. In fact, no uniform choice of revision policy is enforced by the logic: our language of belief change can also describe different sorts of revising behavior together, as in mixed formulas $[\Uparrow\varphi][\uparrow\psi]\alpha$. Does this dissolve logic of belief revision into a jungle of options? In Section 7.5, we will look at this situation in greater generality.

Summary This Section has extended the dynamic approach for updating knowledge to revising beliefs. The result is one merged theory of information update and belief revision, using standard modal techniques instead of ad-hoc formalisms.

7.5 General formats for belief revision

As we have seen, acts of belief revision can be modeled as transforming current epistemic-doxastic models, and there is a large variety of such transformations.

We have also seen how, given a definition of model change, one can write a matching recursion axiom, and then a complete dynamic logic. But how far does this go? In this section, we discuss two general formats that have been proposed to keep the base logic of belief change simple and clean.

7.5.1 Relation transformers as PDL programs.

One general method uses programs in propositional dynamic logic to define the new relations via standard program constructs including the test of a proposition φ (denoted as $?\varphi$), the arbitrary choice of two programs $\alpha \cup \beta$, and sequential program composition $\alpha;\beta$. Many examples fit in this format:

Fact 7.5
Radical upgrade $\Uparrow P$ is definable as the following program in propositional dynamic logic, with 'T' the universal relation between all worlds:

$$\Uparrow P(R) := (?P; T; ?\neg P) \cup (?P; R; ?P) \cup (?\neg P; R; ?\neg P)$$

Van Benthem & Liu then introduced the following general notion in 2007:

Definition 7.10 (PDL-format for relation transformers)
A definition for a new relation on models is in *PDL-format* if it can be stated in terms of the old relation R, union, composition, and tests. $\dashv$

A further example is a weaker act that has been studied in the logic of preference change, as well as that of "relevant alternatives" in formal epistemology.

Example 7.9 (Suggestions as order transformations)
A *suggestion* merely takes out R-pairs with '$\neg P$ over P'. This transformation is definable as the PDL program

$$\sharp P(R) = (?P; R) \cup (R; ?\neg P)$$

This generalises our earlier procedure with recursion axioms considerably:

Theorem 7.4
For each relation change defined in PDL-format, there is a complete set of recursion axioms that can be derived via an effective procedure. $\dashv$

Example 7.10
Instead of a proof, here are two examples of computing modalities for the new relation after the model change, using the well-known recursive program axioms of PDL. Note how the second calculation uses the existential epistemic modality $<>$ for the occurrence of the universal relation:

(a) $< \sharp P(R) >< R > \varphi \leftrightarrow < (?P; R) \cup (R; ?\neg P) > \varphi$
$\leftrightarrow < (?P; R) > \varphi \vee < (R; ?\neg P) > \varphi$
$\leftrightarrow <?P >< R > \varphi \vee < R ><?\neg P > \varphi \leftrightarrow (P \wedge < R > \varphi) \vee < R > (\neg P \wedge \varphi)$.

(b) $<\Uparrow P(R) > \varphi \leftrightarrow < (?P; T; ?\neg P) \cup (?P; R; ?P) \cup (?\neg P; R; ?\neg P) > \varphi \leftrightarrow$
$< (?P; T; ?\neg P) > \varphi \vee < (?P; R; ?P) > \varphi \vee < (?\neg P; R; ?\neg P) > \varphi \leftrightarrow$
$<?P >< T ><?\neg P > \varphi \vee <?P >< R ><?P > \varphi \vee <?\neg P >< R ><?\neg P > \varphi$
$\leftrightarrow (P \wedge E(\neg P \wedge \varphi)) \vee (P \wedge < R > (P \wedge \varphi)) \vee (\neg P \wedge < R > (\neg P \wedge \varphi))$.[16]

The final formula arrived at easily transforms into the axiom that was stated earlier for safe belief after radical upgrade $\Uparrow P$. $\dashv$

In this style of analysis, logic of belief revision becomes a form of propositional dynamic logic, and PDL then serves as the "mother logic" of belief revision. Much more complex PDL mechanisms for model change have been proposed recently, for which we refer to our Section 7.10 on further literature.

However, some forms of belief revision seem to require methods going beyond the PDL frarmework, and in order to explain the resulting general logic, we need to take a closer look at the heart of current Dynamic Epistemic Logic.

7.5.2 Product update in general dynamic-epistemic logic

Dynamic Epistemic Logic (DEL) goes far beyond public announcement logic, whose earliest version is from 1989. In the 1990s, more complex informational scenarios were investigated by Groeneveld, Gerbrandy, and van Ditmarsch, involving mixtures of public and private information. The crucial mechanism in use today is that of *product update* introduced by Baltag, Moss and Solecki, where current models need not shrink in size: they can even *grow* under update, reflecting increased complexities of an informational setting. Currently, the term DEL is used to denote a collection of logical systems that deal with complex multi-agent scenarios in which individual agents or groups of agents update their knowledge and beliefs when new information comes in a variety of public or private events.

In their original setting, these logics were designed to model only cases in which the newly received information is consistent with the agent's prior doxastic or epistemic state. But in recent years, it has become clear that ideas from Belief Revision Theory fit naturally with DEL, allowing us to model a richer set of

[16] Here the operator E stands for the existential modality in some world.

scenarios in which agents can be confronted with surprising information that may contradict their prior beliefs.

In line with our treatment so far in this chapter, DEL interprets all epistemic actions or events as "model transformers", i.e. ways to transform a given input model into an output model, where the actual transformation captures the effect of the epistemic action that took place. A powerful mechanism of this sort is the above product update that covers highly sophisticated scenarios. In this subsection, we will present its epistemic version, where it transforms a current epistemic model (usually, though not necessarily, satisfying the constraints of the modal logic $S4$ or $S5$) using a further "event model" that collects all relevant events insofar as the agents' observational abilities are concerned:

Definition 7.11 (Event models)
An *event model* over a given language $\mathcal{L}$ is a structure $\Sigma = (E, R_i, PRE)$ such that E is the set of relevant actions (or events) in our domain of investigation. For each agent $i \in I$ we have an accessibility relation R_i on E, and instead of a valuation we now have a precondition map $PRE : E \to \mathcal{L}$ which assigns the precondition $PRE(e)$ to each $e \in E$. $\dashv$

Here is what these structures represent. In public announcement logic, the event taking place was totally clear to every agent. In scenarios with private information, agents may not be sure exactly what it is they are observing, and an event model captures this epistemic horizon. The precondition map encodes the information that observed events can convey, while the accessibility relations encode, as in epistemic models, the extent to which the agents can observe what is the event actually taking place. A concrete example will follow shortly, after we have stated the update mechanism.

Product update Consider a situation in which we start from a given epistemic model $M = (W, R_i, V)$ which captures a full description of the epistemic states of all agents $i \in I$. Let the possible world s in M be our point of evaluation (alternatively, the "actual world" where the agents are) in which a given epistemic event e happens, coming from an event model $\Sigma = (E, R_i, PRE)$. To see how the epistemic event (Σ, e) affects the state (M, s), we first observe that this event can only take place when the precondition holds, i.e. $M, s \models PRE(e)$. The result of performing this event is then a new epistemic state (s, e) belonging to the direct product of the model M with the model Σ. We denote the effect of the product update by $(M \otimes \Sigma) = (W', R'_i, V')$ and define it as:

$$W' = \{(s, e) | s \in W, e \in E, \text{ and } M, s \models PRE(e)\}$$

$$R'_i = \{((s, e), (t, f)) \mid R_i(s, t) \text{ and } R_i(e, f)\}$$

$$(s, e) \in V'(p) \text{ iff } s \in V(p)$$

Here the stipulation for the valuation says that base facts do not change under information update.[17] The stipulation for the new accessibility is perhaps best

[17] There are also versions of DEL that model real factual changes by "postconditions".

understood negatively: the only ways in which an agent can distinguish situations after the update is, either the situations were already distinguishable before, or they were indistinguishable, but the signal observed makes a difference: the agent has learnt from the observation.

To see that this rule reflects our intuitions about information change, we look at a concrete example in the style of Baltag and Moss:

Example 7.11 (Private announcements)
Imagine that a coin lies on the table in such a position that two agents Alice (a) and Bob (b) cannot see the upper face of the coin. We assume that it is common knowledge that a and b are both uncertain about whether the upper face is Heads or Tails. This scenario is represented in the $S4$ model of Figure 7.8, where the pointed arrows represent the agent's epistemic uncertainty.

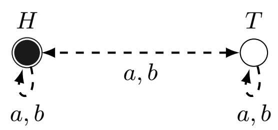


Figure 7.8: Initial coin scenario.

Assume now that the following action takes place: Bob takes a peek at the coin, while Alice doesn't notice this. We assume that Alice thinks that nothing happened and that Bob knows that Alice didn't see him take a peek at the coin. This action is a typical example of a "fully private announcement" in DEL. Bob learns the upper face of the coin, while the outsider Alice believes nothing is happening. In Figure 7.9, we depict the epistemic event model for this situation.

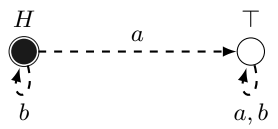


Figure 7.9: Epistemic event model.

We take two 'actions' into account: the actual action e in which Bob takes a peek and sees Heads up; and the action τ where 'nothing is really happening', which has $\top$ as precondition. The accessibility relations drawn in the picture show that τ is the only action that agent a considers possible.

The product update of the previous two models yields the result depicted in Figure 7.10. In this model, the real state of affairs is represented on top, where Heads is true. In this world b knows the upper fact of the coin is Heads while a thinks that nobody knows the upper face of the coin.

Scenarios like this may look a bit contrived, but this is a mistaken impression. Differences in information between agents are precisely what drives communication, and when we get to belief revision later on, this is often a private process, not undergone by all agents in the same manner. ⊣

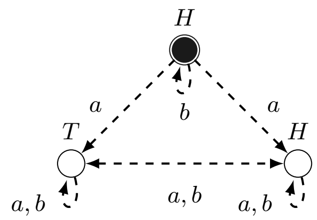


Figure 7.10: The product update result depicted.

Recursion axioms and complete logic The proof theory for DEL can be seen as a generalisation of the setting for public announcement logic. Like we saw with PAL, recursion axioms can be provided to prove the completeness of the system.[18] As in the completeness proof for PAL, we can provide a reduction of every dynamic operator in the logic to obtain an equivalent formula in the static epistemic base language. For the Boolean connectives, the recursion laws are straightforward, so let us focus on the K_i operator. The basic Action-Knowledge Axiom is the following equivalence:

$$[\Sigma, e]K_i\varphi \leftrightarrow (PRE(e) \rightarrow \bigwedge_{f \in \Sigma, eR_if} K_i[\Sigma, f]\varphi)$$

Here the events f with eR_if in Σ form the "appearance" of e to agent i in the current scenario represented by Σ. As our example of private announcement showed, not every event appears in the same way to each agent.

The PAL reduction axiom for $[!\varphi]K_i\psi$ is a special case of the general Action-Knowledge Axiom where the event model consists of one action with a precondition φ seen by all the agents. But the new axiom covers many further scenarios. The product update mechanism of DEL has been successfully applied to many different scenarios involving epistemic uncertainty, private and public communication, public learning-theoretic events, and game solution procedures. We will see a few of these later on in this chapter.

However, to handle more sophisticated scenarios involving doxastic uncertainty and belief revision, we need to extend product update to deal with plausibility order change of the kind introduced in the preceding Section 7.4.

7.5.3 Priority update

Event models as triggers of plausibility change Given the private nature of many belief changes, but also, the social settings that induce these (we often change our beliefs under pressure from others), it makes sense to use the general DEL methodology of interacting different agents. Here is a powerful idea

[18]Things are more delicate with group modalities such as common knowledge. The first completeness proofs for PAL and DEL with common knowledge were intricate because of the lack of a recursion axiom for a common knowledge operator. These matters have been solved since in terms of suitable PDL style extensions of the epistemic base language. For more discussion, see Section 7.9, as well as the references in Section 7.10.

from the work of Baltag & Smets, where the reader may now want to view the events in event models as "signals" for information strength:

Definition 7.12 (Plausibility event models)
Plausibility event models are event models as in the preceding section, now expanded with a plausibility relation over their epistemic equivalence classes. $\dashv$

This setting covers the earlier concrete examples from Section 7.4.

Example 7.12 (Radical upgrade)
A radical upgrade $\Uparrow\varphi$ can be implemented in a plausibility event model. We do not throw away worlds, so we use two 'signals' $!\varphi$ and $!\neg\varphi$ with obvious preconditions φ, $\neg\varphi$ that will produce an isomorphic copy of the input model. But we now say that signal $!\varphi$ is more plausible than signal $!\neg\varphi$, relocating the revision policy in the nature of the input. With a suitable update rule, to be defined below, this will proceed the output model depicted in Figure 7.11. $\dashv$

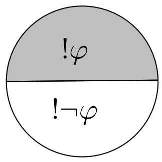


Figure 7.11: Implementing radical upgrade.

Different event models will represent a great variety of update rules. But we still need to state the update mechanism itself more precisely, since the required treatment of plausibility order is not quite the same as that for epistemic accessibility in the preceding subsection. The following proposal has been called 'One Rule To Rule Them All', i.e., one new rule that replaces earlier separate update rules for plausibility. It places the emphasis on the last event observed, but is conservative with respect to everything else:

Definition 7.13 (Priority update)
Consider an epistemic plausibility model (M, s) and a plausibility event model (Σ, e). The *priority product model* $(M \times \Sigma)$, (s, e) is defined entirely as its earlier epistemic version, with the following additional rule for the plausibility relation $\leq_i$, which also refers to its strict version $<_i$:

$$(s, e) \leq_i (t, f) \text{ iff } (s \leq_i t \wedge e \leq_i f) \vee e <_i f$$

Thus, if the new incoming information induces a strong preference between signals, that takes precedence: otherwise, we go by the old plausibility order. The emphasis on the last observation or signal received is like in belief revision theory, where receiving just one signal $*\varphi$ leads the agent to believe that φ, even if all of her life, she had been receiving evidence against φ.[19]

[19] Priority Update is also in line with "Jeffrey Update" in probability theory that imposes a new probability for some specified proposition, while adjusting all other probabilities proportionally.

Theorem 7.5
The dynamic logic of priority update is axiomatisable completely. ⊣

As before, it suffices to state the recursion axioms reflecting the above rule. We display just one case here, and to make things simple, we do so for the notion of safe belief, written in an existential format:

$$< \Sigma, e > \Diamond_i \varphi \leftrightarrow$$
$$PRE(e) \wedge (\bigvee_{e \leq f \text{ in } \Sigma} \Diamond_i < \Sigma, f > \varphi \vee (\bigvee_{e < f \text{ in } \Sigma} \tilde{K_i} < \Sigma, f > \varphi))$$

where $\tilde{K_i}$ is again the existential epistemic modality.

Example 7.13 (The coin scenario revisited)
Consider again the coin example of the preceding subsection, which showed how DEL handles public and private knowledge updates of agents. The given scenario might lead to problems if Alice finds out that Bob took a peak at the coin. So let us now also introduce beliefs. Imagine again that there is a coin on the table in such a position that the two agents Alice (a) and Bob (b) cannot see the upper face of the coin. We now assume that it is common knowledge that a and b believe that the upper face is Heads (see Figure 7.12).[20]

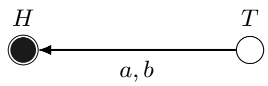


Figure 7.12: Initial coin scenario revisited.

Assume now that the following action takes place: Bob takes a peek at the coin, while Alice does not notice. We assume that Alice believes that nothing happened. Then Bob learns what is the upper face of the coin, while the outsider Alice believes nothing has happened. In Figure 7.13, we represent the plausibility event model for this situation. We take three 'actions' into account: the actual action e in which Bob takes a peek and sees Heads up; the alternative action f in which Bob sees Tails up; and the action τ in which nothing is really happening. The plausibility marking in the event model shows that τ is the most plausible action for Alice, while, if she had to choose between Bob taking a peek and seeing Heads or seeing Tails, she would give preference to him seeing Heads.

Now taking the Priority Update of the preceding two models yields the result depicted in Figure 7.14. In this model with four worlds, the most plausible world for Alice is the lower left one where Heads is true. In this model, Alice believes indeed that everyone believes that Heads is the case and nobody knows the upper face of the coin. The real world however is the top left world, in which Bob does know the upper face of the coin to be Heads.

[20]To make our drawings more transparent, reflexive plausibility arrows have been omitted as well as the epistemic uncertainty relations for agents, which can be computed on the basis of the plausibility relations as we have shown earlier.

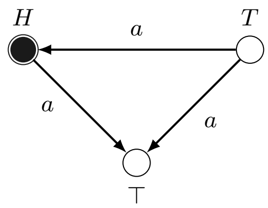
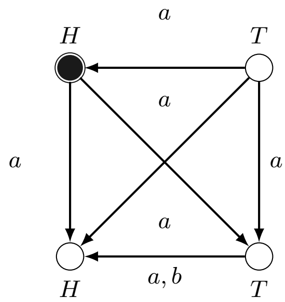


Figure 7.13: Plausibility event model for a private peep.



$\dashv$

Figure 7.14: The computed result of priority update.

The general advantage of this approach is that, instead of having to introduce many different policies for processing an incoming signal, each with its own logic, we can now put the policy in the input Σ. Accordingly, the one logic for Priority Update that we gave before is the unique dynamic logic of belief revision. Admittedly, this introduces some artificial features. The new event models are rather abstract - and, to describe even simple policies like conservative upgrade, the language of event models must be extended to event preconditions of the form 'best φ'. But the benefit is clear, too. Infinitely many policies can be encoded in the choice of plausibility event models, while belief change works with just one update rule, and the common objection that belief revision theory is non-logical and messy for its proliferation of policies evaporates.[21]

7.6 Belief revision and probability dynamics

The main technical alternative to using logic in studies of belief revision have been probabilistic methods. The interface of logic and probability is a vast and flourishing topic, well beyond the confines of this chapter. Even so, we provide a brief presentation of a way in which probabilistic perspectives merge well with

[21] More can still be said on general plausibility update rules, since it is not obvious how the stated priority format covers all PDL program changes discussed earlier. Also, one might question the emphasis on the last signal received, rather than engaging in more cautious processes of learning. These issues will return in Sections 7.6, 7.7, and 7.10.

the dynamic logical methodology of this chapter. This section presents the basics of the system of Probabilistic Epistemic Dynamic Logic, be it in a somewhat terse manner. For details, the reader may consult the literature references in Section 7.10, while Section 7.8 will contain a concrete application to the phenomenon of informational cascades.

Language and models We can think of subjective probabilities as a numerical measure for degrees of belief, though more objective interpretations as observed frequencies make sense as well in our setting of informational inquiry. The following language, due to the work of Halpern and Fagin, allows for a wide range of expressions that mix knowledge and probability.

Definition 7.14 (Epistemic-probabilistic language)
The syntax of *epistemic-probabilistic logic* has the full syntax of the standard epistemic modal language of Section 7.3 plus the probabilistic construct

$$\alpha_1 \cdot P_i(\varphi) + \ldots + \alpha_n \cdot P_i(\varphi) \geq \beta$$

where $\alpha_1, \ldots, \alpha_n, \beta$ stand for arbitrary rational numbers. $\dashv$

The reason for having these inequalities is technical, having to do with treating conditional probabilities, and a smooth completeness proof for the base logic. Standard probability assertions like $P(\varphi) \leq \alpha$ are an obvious special case. Using the new construct we can also add useful abbreviations to our language that express notions often encountered in discussing probabilistic beliefs:

$$P_i(\varphi) > P_i(\psi)(\text{ also written } [\varphi : \psi]_i > 1) \ := \ \neg(P_i(\psi) - P_i(\varphi) \geq 0),$$

It is easy then to also define $P_i(\varphi) = P_i(\psi)$ and similar comparative notions.

Next, here are the semantic structures that this language refers to.

Definition 7.15 (Probabilistic epistemic models)
A *probabilistic epistemic model* $\mathcal{M}$ is a tuple $(W, I, (\sim_i)_{i\in I}, (P_i)_{i\in I}, \Psi, [\![\bullet]\!])$ consisting of a set W of worlds; a set I of agents; and for each agent i, an equivalence relation $\sim_i \subseteq W \times W$ representing i's epistemic indistinguishability as in the above. Also, for each agent i, $P_i : W \to [0, 1]$ is a map that induces a probability measure on each $\sim_i$-equivalence class.[22] Finally, Ψ is a given set of atomic propositions and $[\![\bullet]\!] : \Psi \to \mathcal{P}(W)$ a standard valuation map, assigning to each $p \in \Psi$ some set of worlds $[\![p]\!] \subseteq W$. $\dashv$

On these models, a standard truth definition can be given for the epistemic-probabilistic language. We merely note that formulas of the form $P_i(\psi) \leq k$ are to be interpreted conditionally on agent i's knowledge at s, i.e., as

$$P_i([\![\psi]\!] \cap \{s' \in W : s' \sim_i s\}) \leq k$$

The setting also yields other useful notions. In particular, the *relative likelihood* (or "odds") of state s against state t according to agent i, $[s : t]_i$ is

$$[s : t]_i := \frac{P_i(s)}{P_i(t)}$$

[22]That is, we have $\sum\{P_i(s') : s' \sim_i s\} = 1$ for each $i \in I$ and each $s \in S$.

For concrete illustrations of these models and their transformations to follow, we refer to the references in Section 7.10, and the Urn Scenario in Section 7.8.

Dynamic updates In line with the main theme of this chapter, our next concern is how models like this change under the influence of new information. Again, the simplest case would be public announcements, or soft upgrades – but for many serious applications, the complete generality is needed of dynamic-epistemic product update in the sense of Section 7.5.

As before, the update rule requires forming products with event models that capture the relevant informational scenario as it appears toy the agents. This time, their impressions also have a probabilistic character.

Definition 7.16 (Probabilistic event models)
A *probabilistic event model* $\mathcal{E}$ is a structure $(E, I, (\sim_i)_{i\in I}, (P_i)_{i\in I}, \Phi, pre)$ consisting of a set of possible events E; a set of agents I, equivalence relations $\sim_i \subseteq E \times E$ representing agent i's epistemic indistinguishability between events; and probability assignments P_i for each agent i and each $\sim_i$-information cell. Moreover, Φ is a finite set of mutually inconsistent propositions[23] called *preconditions*. Here *pre* assigns a probability distribution $pre(\bullet|\varphi)$ over E for every proposition $\varphi \in \Phi$.[24] $\dashv$

A word of explanation is needed here. Event models crucially contain two probabilities: one for the certainty or quality of the *observation* made by the agent, and another for the probability of *occurrence* of an event given a precondition. The latter sort of information, often overlooked at first glance, is crucial to analysing famous scenarios of probabilistic reasoning such as Monty Hall or Sleeping Beauty, since it represents the agents' experience of, or belief about, the process they are in (cf. Section 7.8 for more on this theme of procedure).

Now we formulate an update rule that weighs all factors involved: prior probability of worlds, observation probabilities, and occurrence probabilities:

Definition 7.17 (Probabilistic product update)
Given a probabilistic epistemic model $\mathcal{M} = (W, I, (\sim_i)_{i\in I}, (P_i)_{i\in I}, \Psi, [\![\bullet]\!])$ and a probabilistic event model $\mathcal{E} = (E, I, (\sim_i)_{i\in I}, (P_i)_{i\in I}, \Phi, pre)$, the *probabilistic product model* $\mathcal{M} \otimes \mathcal{E} = (W', I, (\sim'_i)_{i\in I}, (P'_i)_{i\in I}, \Psi', [\![\bullet]\!]')$, is given by:

$$W' = \{(s, e) \in W \times E \mid pre(e \mid s) \neq 0\},$$

$$\Psi' = \Psi,$$

$$[\![p]\!]' = \{(s, e) \in W' : s \in [\![p]\!]\},$$

$$(s, e) \sim'_i (t, f) \text{ iff } s \sim_i t \text{ and } e \sim_i f,$$

$$P'_i(s, e) = \frac{P_i(s) \cdot P_i(e) \cdot pre(e \mid s)}{\sum\{P_i(t) \cdot P_i(f) \cdot pre(f \mid t) : s \sim_i t, e \sim_i f\}},$$

[23] Preconditions usually come from the above static probabilistic-epistemic language.
[24] Alternatively, P_i is the probabilistic odds $[e : e']_i$ for any events e, e' and agent i.

where we use the notation

$$pre(e \mid s) := \sum\{pre(e \mid \varphi) : \varphi \in \Phi \text{ such that } s \in [\![\varphi]\!]_{\mathcal{M}}\}$$

Here $pre(e \mid s)$ is either $= pre(e|\varphi_s)$ where φ_s is the unique precondition in Φ such that φ_s is true at s, or $pre(e \mid s) = 0$ if no such precondition φ_s exists.[25] ⊣

This combination of probabilistic logic and dynamic update sheds new light on many issues in probabilistic reasoning, such as the status of Bayes' Theorem. Moreover, its use of occurrence probabilities allows for careful modeling of probabilistic scenarios in areas such as learning theory (cf. Section 7.7), where we may have probabilistic information about the nature of the process giving us a stream of evidence about the world. denominator However, the above update mechanism does not solve all interface problems. For instance, like our earlier update mechanisms in Sections 7.4 and 7.5, the last signal received, represented by the event model, gets a huge weight – and alternatives have been considered, closer to standard systems of "inductive logic", that weigh the three probabilities involved differently.

Dynamic Probabilistic Logic As with our earlier systems of dynamic epistemic logic, there is a complete static and dynamic logic for the system defined here, including a complete set of recursion axioms for the new probability inequalities that hold after the application of an event model. We omit technical formulations, but the point is that our earlier logical methodology fully applies to the present more complex setting.

But there are many further logical issues about connections with our earlier sections. One is that we could add qualitative plausibility-based belief as well, giving us both quantitative and qualitative versions of beliefs. In recent years, such combinations have attracted increasing interest. Qualitative logical notions are often considered competitors to quantitative probabilistic ones, and indeed there are interesting issues as to whether the above dynamic update mechanisms can be stated entirely in terms of plausibility order.[26] But perhaps the more interesting issue for this chapter would be whether qualitative notions emerge naturally in cognition as companions to underlying probabilistic ones, a line of thought to which several recent authors have given new impetus.

Summary We have shown how probabilistic views of belief can be merged naturally with the main dynamic logic approach in this chapter. This combination is useful for applications to areas that heavily depend on probabilistic methods, such as game theory or the social sciences, but it also raises interesting broader conceptual issues that are far from resolved.

[25]Note some subtleties apply to the definition of Probabilistic Product Update, which in addition has to build in the fact that $P_i'(s, e)$ excludes the deonomator from being 0.

[26]Numbers play different roles in the above rules: as strengths of beliefs, as weights for "gluing" beliefs, and others. Thus, qualitative versions of the update rule may have to involve different mechanisms, such as "order merge" for the three probabilities above made qualitative.

7.7 Time, iterated update, and learning

Belief revision theory has mainly focused on single steps of changing beliefs, and the same is true for the events of information change in the dynamic-epistemic logics that we have used. Of course, we can iterate such steps to form longer sequences, but so far, we have not looked at the global temporal dimension per se. And yet, single steps of belief revision may lack direction, like leaves in the wind. Many serious scenarios where belief revision plays a role are global processes of inquiry or learning that have an intrinsic temporal structure, sometimes even consisting of infinite histories.

With this temporal setting come constraints on how these histories can unfold, not necessarily captured by the local preconditions of events that we have used so far. These constraints on the procedure are often called *temporal protocols*. For instance, information systems may demand that only true information is passed, or that each request is answered eventually. And civilised conversation has rules like "do not repeat yourself", or "let others speak as well".

Restricting the legitimate sequences of announcements is not just an extra, it affects our earlier dynamic logics.

Example 7.14 (Contracting consecutive assertions: admissible, or not?) PAL can suppress longer sequences of announcements into one. A well-known PAL-validity states that two consecutive announcements $!\varphi, !\psi$ have the same effect as the single announcement

$$!(\varphi \wedge [!\varphi]\psi)$$

However, the new assertion may well be more complex than what is admissible by current rules for conversation or investigation, and hence this law may fail in protocol-based models. $\dashv$

In this section, we will introduce some temporal logics of knowledge that form a natural extension to the dynamic epistemic logics used so far. After that we show how dynamic epistemic logics lie embedded here, including their protocol versions. Then we show how the same is true for logics of belief, and finally, bringing together all ideas developed so far, we provide some recent connections with formal learning theory, a natural continuation of belief revision theory.

This is a vast area, and we only provide some windows, referring the reader to the further literature referenced in Section 7.10.

7.7.1 Epistemic temporal logic

Branching temporal models are a Grand Stage view of agency, as depicted in Figure 7.15, with histories as complete runs of some information-driven process that can be described by languages with epistemic and temporal operators.

Temporal logics for epistemic and doxastic agents come in different flavors, and we will only briefly discuss one of them, interpreted over a modal universe of finite histories and indices of evaluation.

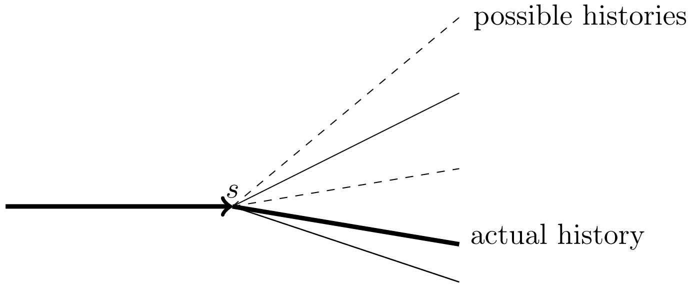


Figure 7.15: A branching temporal tree of histories.

Models and language Start with two sets I of agents and E of events. A *history* is a finite sequence of events, and E^* is the set of all histories. Here he is history h followed by event e, the unique history after e has happened in h. We write $h \leq h'$ if h is a prefix of h', and $h \leq_e h'$ if $h' = he$.

Definition 7.18 (Epistemic-temporal ETL models)
A *protocol* is a set of histories $H \subseteq E^*$ closed under prefixes. An *ETL model* is a tuple $(E, H, \{\sim_i\}_{i \in I}, V)$ with a protocol H, accessibility relations $\{\sim_i\}_{i \in I}$ plus a valuation map V sending proposition letters to sets of histories in H. $\dashv$

An ETL model describes how knowledge evolves over time in some informational process. The relations $\sim_i$ represent uncertainty of agents about the current history, due to their limited powers of observation or memory. $h \sim_i h'$ means that from agent i's point of view, history h' looks the same as history h.

An epistemic temporal language LETL for these structures is generated from a set of atomic propositions At by the following syntax:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi \mid < e > \varphi$$

where $i \in I, e \in E$, and $p \in At$. Booleans, and dual modalities $\tilde{K_i}, [e]$ are defined as usual. Let $M = (E, H, \{\sim_i\}_{i \in I}, V)$ be an ETL model. The truth of a formula φ at a history $h \in H$, denoted as $M, h \models \varphi$, is defined inductively as usual. We display the two key clauses, for knowledge and events:

(a) $M, h \models K_i\varphi$ iff for each $h' \in H$, if $h \sim_i h'$, then $M, h' \models \varphi$

(b) $M, h \models < e > \varphi$ iff there exists $h' = he \in H$ with $M, h' \models \varphi$.

This language can express many properties of agents and their long-term behavior over time. It has a base logic that we will not formulate here, though we will make a few comments on it later.

Agent properties In particular, the language LETL can express interesting properties of agents or informational processes through constraints on ETL models. These often come as epistemic-temporal axioms matched by modal frame correspondences. Here is a typical example.

	K, P, F	$K, C_G,$ $<e>$	$K, C_G, <e>,$ $PAST_e$	K, C_G, F
ETL	decidable	decidable	decidable	RE
ETL + PR	RE	RE	RE	Π_1^1-complete
ETL + NM	RE	RE	RE	Π_1^1-complete

Table 7.1: Complexity of epistemic temporal logics. ETL denotes the class of all ETL models, PR denotes Perfect Recall and NM is No Miracles.

Fact 7.6
The axiom $K_i[e]\varphi \to [e]K_i\varphi$ corresponds to Perfect Recall:

$$\text{if } he \sim_i k, \text{ then there is a history } h' \text{ with } k = h'e \text{ and } h \sim_i h'$$

This says that agents' current uncertainties can only come from previous uncertainties: a constraint on ETL models that expresses a strong form of so-called *Perfect Recall*.[27] Note that the axiom as stated presupposes perfect observation of the current event e: therefore, it does not hold in general DEL, where uncertainty can also be created by the current observation, when some event f is indistinguishable from e for the agent.

In a similar fashion, the axiom $[e]K_i\varphi \to K_i[e]\varphi$ corresponds to *No Miracles*: for all ke with $h \sim_i k$, we also have $he \sim_i ke$. This says essentially that learning only takes place for agents by observing events resolving current uncertainties.

Digression: epistemic temporal logics and complexity Epistemic-temporal logics have been studied extensively, and we cannot survey their theory here. However, one feature deserves mention. In this chapter, we will not pay attention to *computational complexity* of our logics, except for noting the following. There is a delicate balance between expressive power and computational complexity of combined logics of knowledge and time which also extends to belief. A pioneering investigation of these phenomena was made in 1989 by Halpern & Vardi. Table 7.1 lists a few observations from their work showing where dangerous thresholds occur for the complexity of validity.

In Table 7.1 complexities run from decidable through axiomatisable (RE) to Π_1^1-complete, which is the complexity of truth for universal second-order statements in arithmetic. What we see here is that complexity of the logic depends on two factors: expressive power of the language (in particular, social forms of group knowledge matter), and so do special assumptions about the type of agents involved. In particular, the property of Perfect Recall, which seems a harmless form of good behavior, increases the complexity of the logic.[28]

[27]Perfect Recall implies synchronicity: uncertainties $h \sim_i k$ only occur between h, k at the same tree level. Weaker forms of perfect memory in games also allow uncertainty links that cross between tree levels.

[28]Technically, Perfect Recall makes epistemic accessibility and future moves in time

7.7.2 Protocols in dynamic epistemic logic

Before we connect DEL and ETL in technical detail, let us see how the crucial notion of protocol in temporal logic natural enters the realm of DEL, and its extensions to belief revision.

Definition 7.19 (DEL protocols)
Let Σ be the class of all pointed event models. A *DEL protocol* is a set $P \subseteq \Sigma^*$ closed under taking initial segments. Let M be any epistemic model. A *state-dependent DEL protocol* is a map P sending worlds in M to DEL protocols. If the protocol assigned is the same for all worlds, it is called "uniform". ⊣

DEL protocols induce TL models in a simple manner. Here is an illustration.

Example 7.15 (An ETL model generated by a uniform PAL protocol)
We use a public announcement protocol for graphical simplicity. Consider the epistemic model M in Figure 7.16 with four worlds and two agents 1 and 2.

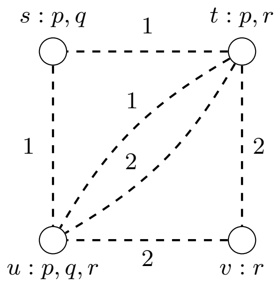


Figure 7.16: Initial epistemic model for a uniform PAL protocol.

The model comes with a protocol

$$P = \{<!p>, <!p, !q>, <!p, !r>\}$$

of available sequences of announcements or observations. The ETL forest model depicted in Figure 7.17 is then the obvious "update evolution" of M under the available announcements, with histories restricted by the event sequences in P. Note how some worlds drop out, while others 'multiply'. ⊣

The logic of PAL protocols Adding protocols changes the laws of dynamic epistemic logic. Again our example concerns public announcement.

Example 7.16 (Failures of PAL validities)
PAL had a valid axiom $<!\varphi> q \leftrightarrow \varphi \wedge q$. As a special case, this implied

$$<!\varphi> T \leftrightarrow \varphi$$

behave like a *grid* of type $IN \times IN$, with cells satisfying a confluence property. Logics of such grids are known to have very high complexity since they can encode so-called "tiling problems" of a complex geometrical nature.

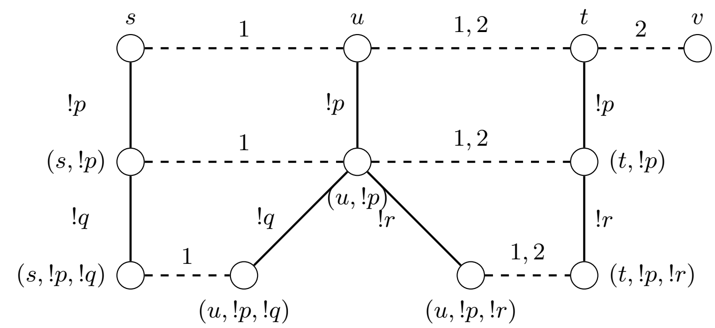


Figure 7.17: Generated ETL forest model.

From left to right, this is valid with any protocol: an announcement $!\varphi$ can only be executed when φ holds. But the direction from right to left is no longer valid: φ may be true at the current world, but the protocol need not allow a public announcement of this fact at this stage. Similar observations can be made for the PAL recursion law for knowledge. ⊣

Thus, assertions $<!\varphi > T$ come to express genuine procedural information about the informative process agents are in, and hence, they no longer reduce to basic epistemic statements. However, there is still a decidable and complete logic TPAL for PAL protocol models, be it, that we need to modify the recursion axioms. The two key cases are as follows:

$$<!\varphi > q \leftrightarrow <!\varphi > T \wedge q \quad \text{for atomic facts } q$$

$$<!\varphi > K_i\psi \leftrightarrow <!\varphi > T \wedge K_i(<!\varphi > T \rightarrow <!\varphi > \psi)$$

Similar modifications yield protocol versions of DEL with product update.

7.7.3 Representing DEL inside temporal logic.

Now we state the more general upshot of the preceding observations. The Grand Stage is also a natural habitat for the local dynamics of DEL. Epistemic temporal trees can be created through constructive unfolding of an initial epistemic model M by successive product updates, and one can determine precisely which trees arise in this way. We only give a bare outline.

A basic representation theorem Consider a scenario of "update evolution": some initial epistemic model M is given, and it gets transformed by the gradual application of event models $\Sigma = \Sigma_1, \Sigma_2, \cdots$ to form a sequence

$$M_0 = M, M_1 = M_0 \times \Sigma_1, M_2 = M_1 \times \Sigma_2, \cdots$$

where stages are horizontal, while worlds may extend downward via one or more event successors. Through successive product update, worlds in the resulting "induced ETL forest model" *Forest(M, Σ)* are finite sequences starting with one world in the initial epistemic model M followed by a finite sequence of events, inheriting their standard epistemic accessibility relations.

Induced ETL forest models have three properties making them stand out, two of which generalise the above special agent properties. In what follows, quantified variables $h, h', k, \cdots$ range only over histories present in M:

(a) If $he \sim k$, then for some f, both $k = h'f$ and $h \sim h'$ (Perfect Recall)

(b) If $h \sim k$, and $h'e \sim k'f$, then $he \sim kf$ (Uniform No Miracles)

(c) The domain of any event e is definable in the epistemic base language.

Condition (c) of "Definable Executability" ensures admissible preconditions.

Theorem 7.6
For ETL models $\mathcal{H}$, the following two conditions are equivalent:[29]

(a) $\mathcal{H}$ is isomorphic to some DEL-induced model Forest(M, Σ)

(b) $\mathcal{H}$ satisfies Perfect Recall, Uniform No Miracles, and Definable Executability. $\dashv$

DEL as an ETL-logic Now we can place DEL and understand its behavior. Its language is the $K, C_G, < e >$ slot in the earlier table, on models satisfying Perfect Recall and No Miracles. Thus, there is grid structure, but the expressive resources of DEL do not exploit it to the full, using only one-step future operators $<!P>$ or $< \Sigma, e >$. Adding unbounded future yields the same complexity as for ETL. Miller & Moss showed in 2005 that the logic of public announcement with common knowledge and Kleene iteration of assertions is Π^1_1-complete.

7.7.4 Beliefs over time

We have only discussed temporal perspectives on knowledge so far, but with a few suitable extensions, everything that we have said also applies to belief and belief revision. We show some technicalities here, and then continue with applications in later subsections.

Epistemic-doxastic temporal models So-called "DETL models" are branching event trees as before, with nodes in epistemic equivalence classes now also ordered by plausibility relations for agents. These models interpret belief modalities at finite histories, in the same style as in the earlier plausibility models.

The epistemic-doxastic language of these models is the natural temporal extension of the logics of belief discussed in earlier sections. It can express natural doxastic properties of agents (some are stated below), and it fits well with the temporal analysis of the AGM postulates that have been made by Bonanno (see his chapter on temporal belief revision in this Handbook for many further themes), as well as Dégrémont & Gierasimczuk.

[29] The theorem still holds when we replace Definable Executability by closure of event domains under all purely epistemic bisimulations of the ETL-model $\mathcal{H}$.

As with knowledge, we can ask which DETL models arise as traces of the update scenarios that we have studied before. For this purpose, one can take sequences of plausibility event models applied to some initial epistemic-doxastic model, and compute their update evolution with the earlier product Rule for epistemic accessibilities and the Priority Rule of Section 7.5 for plausibility. An analogue to the earlier representation theorem then arises, in terms of two basic properties of plausibility between histories:

Fact 7.7
The histories h, h', j, j' arising from iterated Priority Update satisfy the following two principles for any events e, f:

(a) if $je \leq j'f$, then $he \geq h'f$ implies $h \geq h'$ ("Plausibility Revelation")

(b) if $je \leq j'f$, then $h \leq h'$ implies $he \leq h'f$ ("Plausibility Propagation") $\dashv$

Theorem 7.7
A DETL model can be represented as the update evolution of an epistemic-doxastic model under a sequence of epistemic-plausibility updates iff it satisfies the structural conditions of the epistemic DEL-ETL representation theorem[30], plus Plausibility Revelation and Plausibility Propagation. $\dashv$

7.7.5 Iterated belief upgrades and limit behavior

ETL models and DETL models are very abstract and general. Much more concrete temporal scenarios arise with iterated dynamic epistemic steps of knowledge update or belief revision. Scenarios of this kind are the well-known Puzzle of the *Muddy Children*, often cited in the dynamic epistemic literature, disagreement scenarios in the epistemic foundations of game theory, or the game solution procedures of Section 7.8 below. However, in this section, we will only consider a scenario from work by Baltag and Smets on propagating beliefs in groups of agents that makes sense in social networks.

We will be concerned with iterated truthful belief revision, namely, iterations of different upgrades with true assertions that lead to sequences or streams of models. When we do this with PAL style updates of the form $!\varphi$, at least on finite models, "stabilisation" occurs in a unique model since the sequence of submodels produced is monotonically non-increasing.[31] With iterated belief changes, however, the plausibility order has a third option of "cycling", resulting in endless oscillation, as we will show a bit later on. Therefore, our main question will be whether and when an iterated belief revision process induced by truthful upgrades converges to a fixed point or not.

We start with some basic definitions, and then state a basic result on limits of belief change from the work of Baltag and Smets.

[30] Here one now needs invariance of event domains for epistemic-doxastic bisimulations.

[31] In fact, at the limit stage, only two options can occur. The sequence stabilises in a non-empty model where φ has become common knowledge, or the sequence breaks off at the first stage where φ has become false at the actual world (as happens with the Muddy Children).

Definition 7.20 (Joint upgrades)
Over the plausibility models of Section 7.4, we use the term *joint upgrade* for the effect that three different model transformers can have. We denote them in general by $\dagger\varphi$, where $\dagger \in \{!, \Uparrow, \uparrow\}$. $\dashv$

Now we define some technical types of behavior. A joint upgrade $\dagger\varphi$ is *redundant* on a pointed model M for a group of agents G if $\dagger\varphi(M)$ is bisimilar with M (written as $\dagger\varphi(M) \simeq_G M$.)[32] This means that, as far as group G is concerned, $\dagger\varphi$ does not change anything essential when applied to model M): all the group G's mutual beliefs, conditional beliefs, strong beliefs, mutual knowledge, and common knowledge stay the same after the upgrade. By contrast, an upgrade $\dagger\varphi$ is *informative* on M for group G if it is not redundant with respect to G.[33] Finally, we say that a model M is a *fixed point* of $\dagger\varphi$ if $M \simeq \dagger\varphi(M)$, i.e. if $\dagger\varphi$ is redundant on M with respect to the group of all agents.

At this point, we can capture stabilisation in the limit.

Logical characterisations Redundancy and Fixed Points can be characterised in the following logical terms, using doxastic notions that were introduced in Section 7.4.

1. $!\varphi$ is redundant with respect to a group G iff φ is common knowledge in the group G; i.e., $M \simeq_G !\varphi(M)$ iff $M \models C_G\varphi$.[34]
2. $\Uparrow\varphi$ is redundant with respect to a group G iff it is common knowledge in the group G that φ is strongly believed by all G-agents. That is, $M \simeq_G \Uparrow\varphi(M)$ iff $M \models C_G(ESb_G\varphi)$.
3. $\uparrow\varphi$ is redundant with respect to a group G iff it is common knowledge in the group G that φ is believed by all G-agents. That is, $M \simeq_G \uparrow\varphi(M)$ iff $M \models C_G(EB_G\varphi)$.

Now we set up the machinery for iterations of upgrades starting from some initial model. The following auxiliary definitions lead up to our main result.

Upgrade streams An *upgrade stream* $\dagger\vec{\varphi} = (\dagger\varphi_n)_{n\in N}$ is an infinite sequence of joint upgrades $\dagger\varphi_n$ of the same type $\dagger \in \{!, \Uparrow, \uparrow\}$. Any upgrade stream $\dagger\vec{\varphi}$ induces a function mapping every pointed model M into an infinite sequence $\dagger\vec{\varphi}(M) = (M_n)_{n\in N}$ of pointed models, defined inductively by:

$$M_0 = M, \quad \text{and} \quad M_{n+1} = \dagger\varphi_n(M_n).$$

The upgrade stream $\dagger\vec{\varphi}$ is *truthful* if every $\dagger\varphi_n$ is truthful with respect to M_n, i.e., $M_n \models \varphi_n$. Next, a *repeated truthful upgrade* is a truthful upgrade stream of the

[32] Here we mean bisimilarity in the usual sense of modal logic, with respect to all accessibility relations for all agents.

[33] As a special case, an upgrade $\dagger\varphi$ is redundant with respect to (or informative to) an agent i if it is redundant with respect to (or informative to) the singleton group $\{i\}$.

[34] Here are two special cases. $!\varphi$ is redundant with respect to an agent i iff i knows φ. Also, M *is a fixed point of* $!\varphi$ *iff* $M \models C\varphi$. Similar special cases apply for the next two clauses in the text.

form $(\dagger\varphi_n)_{n\in N}$, where $\varphi_n \in \{\varphi, \neg\varphi\}$ for some proposition φ. In other words, it consists in repeatedly learning the answer to the same question φ?

Stabilisation A stream $\dagger\vec{\varphi}$ *stabilises* a pointed model M if there exists some $n \in N$ with $M_n \simeq M_m$ for all $m > n$. A repeated truthful upgrade stabilises M if it reaches a fixed point of either $\dagger\varphi$ or of $\dagger(\neg\varphi)$. Next, we say that $\dagger\vec{\varphi}$ *stabilises all simple beliefs* (i.e., non-conditional ones) on M if the process of belief-changing induced by $\dagger\vec{\varphi}$ on M reaches a fixed point; i.e., if there exists some $n \in N$ such that $M_n \models B_i\varphi$ iff $M_m \models B_i\varphi$, for all agents i, all $m > n$, and all doxastic propositions φ. Similarly, $\dagger\vec{\varphi}$ *stabilises all conditional beliefs* on a model M if the process of conditional-belief-changing induced by $\dagger\vec{\varphi}$ on M reaches a fixed point as before, but now with respect to conditional belief $B_i^{\psi}\varphi$ for all doxastic propositions φ, ψ.[35] Finally, $\dagger\vec{\varphi}$ *stabilises all knowledge* on the model M if the knowledge-changing process induced by $\dagger\vec{\varphi}$ on M reaches a fixed point, in an obvious sense modifying the preceding two notions.

At last, we can state some precise technical results.

Lemma 7.1
The following two assertions are equivalent:

- An upgrade stream $\dagger\vec{\varphi}$ stabilises a pointed model M,
- $\dagger\vec{\varphi}$ stabilises all conditional beliefs on M. $\dashv$

Theorem 7.8
Every truthful radical upgrade stream $(\Uparrow\varphi_n)_{n\in N}$ stabilises all simple, non-conditional beliefs – even if it does not stabilise the model. $\dashv$

This result has a number of interesting consequences. For instance, every iterated truthful radical upgrade definable by a formula in doxastic-epistemic logic (i.e., in the language of simple belief and knowledge operators, without conditional beliefs) stabilises every model with respect to which it is correct, and thus stabilises all conditional beliefs. The analogue of this result is not true for conservative upgrade, where updates can keep oscillating – so limit behavior depends in subtle manners on the sort of belief revision involved.

We have shown how limit behavior of belief revision steps fits in the scope of logic, provided we place our dynamic epistemic logics in a broader temporal setting. In particular, the specific protocol of only allowing upgrades for specified assertions turned out to have a lot of interesting properties.

7.7.6 From belief revision to formal learning theory

Continuing with the theme of long-term behavior, we conclude this section by pointing out links with *formal learning theory.* In this theory, learning methods are studied for identifying an as yet unknown world on the basis of evidence streams that it generates, where identification sometimes takes place by some finite stage, but often "only in the limit", in a sense to be defined below.

[35] A similar definition can be formulated for stabilisation of strong beliefs.

Belief revision and learning theory The DEL framework is well equipped to describe local learning of facts, but for the long-term identification of a total, possibly infinite, history of evidence, temporal notions need to enter as well, as happened in the iterated upgrade scenarios of the preceding subsection. But this combination indeed matches basic features of formal learning theory. Following a suggestion by Kelly that learning theory is a natural extension of belief revision, separating bad policies from optimal ones, Baltag, Smets & Gierasimczuk have found a number of precise links, where the equivalent of "learning methods" are upgrade methods forming beliefs about the actual history on the basis of growing finite sets of data, with a correct stable belief achieved on a history as the doxastic counterpart of the learning theoretic concept of "identifiability in the limit". In what follows, we briefly present a few highlights.

Revision policies for learning Already in Section 7.1, we have seen that the AGM postulates are conservative in the sense that an agent who adopts this method will keep holding on to her old beliefs as long as possible while incorporating the new incoming information. While this looks attractive as a learning method, conservative upgrade as defined in Section 7.4 is often unable to shake a current suboptimal belief out of its complacency, and other more radical methods turn out to work better in limit identification, as we shall see.

Modeling learning methods via belief revision We start from an *epistemic space* (W, Φ) consisting of a set W of epistemic possibilities (or "possible worlds"), together with a family of observable properties $\Phi \subseteq \mathcal{P}(W)$. We work with streams of successive observations, and denote an infinite such stream as

$$\epsilon = (\epsilon_1, \epsilon_2 \ldots) \in \Phi^\omega$$

A stream of observations is said to be *sound and complete* with respect to a given world $s \in W$ if the set $\{\epsilon_n : n \in N\}$ of all properties that are observed in the stream coincides with the set $\{\varphi \in \Phi : s \in \varphi\}$ of all observable properties φ that are true in s. Now comes the crucial notion driving the analysis.

Definition 7.21 (Learning methods)
A *learning method* L for an agent is a map that associates to any epistemic space (W, Φ) and any finite sequence of observations $\sigma = (\sigma_0, \ldots, \sigma_n)$, some *hypothesis* as a result, where hypotheses are just subsets of W. A world $s \in W$ is said to be *learnable* by the method L if, for every observation stream ϵ that is sound and complete for s, there is a finite stage N such that $L(W, \Phi; \epsilon_0, \ldots, \epsilon_n) = \{s\}$ for all $n \geq N$. The epistemic space (W, Φ) itself is learnable by method L if all the worlds in W are learnable by the same method. ⊣

Now we introduce belief structure via plausibility orderings. Starting from an epistemic space (W, Φ), we define a *plausibility space* $(W, \Phi, \leq)$ by equipping it with a total preorder $\leq$ on W. A *belief-revision method* will then be a function R that associates to $(W, \Phi, \leq)$ and any observation sequence $\sigma = (\sigma_0, \ldots, \sigma_n)$, a new plausibility space

$$R(W, \Phi, \leq; \sigma) := (W^\sigma, \Phi^\sigma; \leq^\sigma),$$

with $W^\sigma \subseteq W$ and $\Phi^\sigma = \{P \cap W^\sigma : P \in \Phi\}$. Such a belief-revision method R, together with a prior plausibility ordering over W, generates in a canonical way a learning method L via the stipulation

$$L(W, \Phi, \sigma) := Min\, R(W, \Phi, \leq_W, \sigma)$$

where $Min(W', \leq')$ is the set of all the least elements of W' with respect to $\leq'$ if such least elements exist, or $\emptyset$ otherwise.

Now we say that an epistemic space (W, Φ) is *learnable by a belief-revision method* R if there exists some prior plausibility assignment on W such that (W, Φ) is learnable by the associated learning method $L(W, \Phi, \leq_W)$.

Learning methods can differ in terms of strength, ranging from weak methods to those that are *universal* in the sense of being able to learn any epistemic state that is learnable at all. Here is a major result of the belief-based analysis.

Theorem 7.9
Conditioning and radical upgrade are universal AGM-like iterated belief revision methods. Conservative upgrade is not. $\dashv$

But there is more to this style of analysis, and we mention two further issues.

Well-founded prior. The preceding theorem is based on arbitrary initial epistemic plausibility spaces, and freedom in the choice of the prior turns out very important. It can be shown that there are learning problems where only a non-wellfounded prior plausibility ordering allows the revision methods of iterated conditioning and lexicographic revision to be universal.[36]

Errors in observations. As it is standard in Learning Theory, the easiest scenario is learning under truthful observations. Much harder to analyse are learning scenarios that allow for errors in observations. In such a setting, one needs a "fairness condition" for learning, which says that that errors occur only finitely often and are always eventually corrected. Under such conditions, there still exist belief-revision methods that are universal, but now only one remains, that allows for radical shifts in hypotheses.

Fact 7.8
Iterated radical upgrade is a universal learning method for fair evidence streams. Conditioning and minimal revision are not universal in this setting. $\dashv$

Logic of learning theory Learning Theory supports the logical languages of earlier sections, which can distinguish a wide range of learning goals. For instance, the formula $FK\varphi$ or modalised variants thereof expresses for suitable φ that there comes a stage where the agent will know that φ.[37]

[36] In the class of standard well-founded plausibility spaces, which seemed rather natural in earlier sections, no AGM-like belief-revision method is universal.

[37] Stronger learning notions would be expressed by epistemic-temporal formulas like $FGK\varphi$ or $F(GK\varphi \vee GK\neg\varphi)$.

But closer to the preceding analysis are weaker belief-based success principles for learning such as the following:

$F(B\psi \to \psi)$ says that my beliefs will eventually be true,

while $F(B\psi \to K\psi)$ says that they will turn into knowledge.

Using suitably extended languages, logical definitions have been given for most standard notions of learnability, ranging from identification in the limit to stronger forms of "finite identifiability" and "conclusive learning". Thus, epistemic-doxastic-temporal logic comes to express the basics of learning theory.

Summary We have seen how dynamic logics of belief change receive a natural continuation in iterative scenarios and temporal logics of inquiry over time, leading to a better view of the role of belief revision in a general account of learning beings.

7.8 Belief, games, and social agency

Belief revision may seem an agent-internal process, with unobservable changes of mind taking place in utmost privacy. But in reality, belief revision has many important social aspects. For instance, while many people think of triggers for belief revision as some surprising fact that a person observes – very often, the trigger is something said by someone else, in a multi-agent setting of communication. Human beliefs are deeply influenced by social settings, and they enter into mutual expectations that drive behavior in scenarios such as games. Moreover, in addition to individual agents interacting, at various places in this chapter, we have even gone further, and also mentioned groups as collective actors that can have properties such as common or distributed knowledge or belief.

The logics that we have introduced in the preceding sections are well up to this extension, even though the technicalities of adding common knowledge or belief are not our main concern here.[38] Instead of developing this theory here (for which we provide literature references in Section 7.10), we discuss two samples of belief revision in basic social settings, viz. games and social networks. Some additional themes will be mentioned in Section 7.9 on further directions.

7.8.1 Iterated beliefs in games

An important area for interactive social agency are games. Games support the logics of this chapter in various ways, from modal logics of action to logics of knowledge and belief for the agents playing them. There is a fast-growing literature on these interfaces that lies far outside the scope of this chapter (see Section 7.10 for some references), but a concrete illustration is within reach.

[38] These technicalities occasionally involve a move to much more powerful static formalisms such as "epistemic PDL" – and also, some specific open questions remain, such as finding an optimal formalism with recursion laws for reasoning with common belief under Priority Update.

Beliefs play a major role in various phases of game play, from prior deliberation to actual play, and from there to post-game analysis. In this subsection, we consider the game solution method of *Backward Induction*, a procedure for creating expectations about how a game will proceed. We will construe it as an iteration scenario like in our preceding section, first in terms of public announcements and knowledge, and in our final analysis, in terms of forming beliefs. The presentation to follow is from earlier work by van Benthem.

Backward Induction and announcing rationality Our first dynamic analysis of Backward Induction is as a process of silent deliberation by players whose minds proceed in harmony. The steps are announcements $!\varphi$, now interpreted as mental reminders to players that some relevant proposition φ is true. Here the driving assertion φ is *node rationality ("rat")*, defined as follows. At a turn for player i, call a move a dominated by a sibling move b (available at the same node) if every history through a ends worse, in terms of i's preference, than every history through b.

Now the key proposition *rat* says:

"Coming to the current node, no one ever chose a strictly dominated move"

Announcing this is informative, and it will in general make a current game tree smaller by eliminating nodes. But then we get a dynamics as with the earlier-mentioned Muddy Children, where repeated true assertions of ignorance eventually solve the puzzle. For, in the smaller game tree, new nodes become dominated, and so announcing *rat* again (saying that it still holds after this round of deliberation) now makes sense.

As we saw in Section 7.8, this process of iterated announcement reaches a limit, a smallest subgame where no node is dominated any more.

Example 7.17 (Solving games through iterated assertions of rationality) Consider the game depicted in Figure 7.18, with three turns, four branches, and pay-offs for players A, E marked in that order:

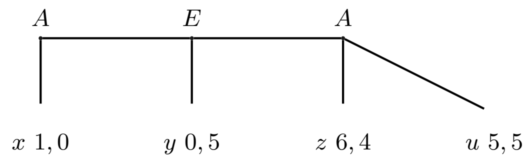


Figure 7.18: An extensive form game.

Stage 0 of the procedure rules out point u: the only point where *rat* fails, Stage 1 rules out z and the node above it: the new points where *rat* fails, and Stage 2 rules out y and the node above it. Each stage deletes part of the game tree so that in the remaining game, *rat* holds throughout. $\dashv$

The actual Backward Induction path for extensive games is obtained by repeated announcement of the assertion *rat* to its limit:

Theorem 7.10
In any game tree M, $(!rat, M)\sharp$ is the actual subtree computed by BI. $\dashv$

A more sensitive scenario: iterated plausibility change The preceding analysis of Backward Induction was in terms of knowledge. However, many foundational studies in game theory view rationality as choosing a best action given what one *believes* about the current and future behavior of the players. This suggests a refinement in terms of our soft updates of Section 7.4 that did not eliminate worlds, but rearrange their plausibility ordering. Now recall our observation that Backward Induction creates expectations for players. The information produced by the algorithm is then in the plausibility relations that it creates inductively for players among end nodes in the game, i.e., complete histories:

Solving a game via radical upgrades. Consider the preceding game once more. This time, we start with all endpoints of the game tree incomparable qua plausibility. Next, at each stage, we compare sibling nodes, using the following notion. A move x for player i *dominates its sibling* y *in beliefs* if the most plausible end nodes reachable after x along any path in the whole game tree are all better for the active player than all the most plausible end nodes reachable in the game after y. We now use the following driver for iterated upgrades:

Rationality* (rat^*): No player plays a move that is dominated in beliefs.

Then we can use an ordering change that is like a radical upgrade $\Uparrow rat^*$:

If x dominates y in beliefs, we make all end nodes from x more plausible than those reachable from y, keeping the old order inside these zones.

This changes the plausibility order, and hence the dominance pattern, so that the doxastic assertion rat^* can change truth value, and iteration can start. Figure 7.19 depicts the stages for this procedure in the preceding game example, where the letters x, y, z, u stand for the end nodes or histories of the game:

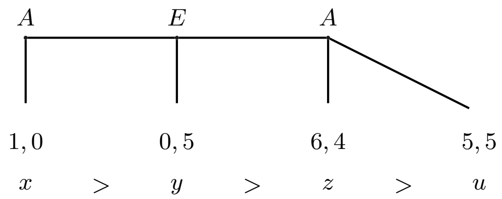


Figure 7.19: Creating plausibility order on histories of a game.

Theorem 7.11
On finite game trees, the Backward Induction strategy is encoded in the stable plausibility order for end nodes created in the limit by iterated radical upgrade with rationality-in-belief. $\dashv$

Notice that, at the end of this procedure, the players as a group have acquired *common belief in rationality*, a fundamental notion in the epistemic foundations of game theory. However, in line with the dynamic focus of this chapter, it is not so much this end result as the procedure itself that deserves attention. Rationality is not a state of grace, but a style of doing things.

7.8.2 Belief change in social networks

A further source of examples are recent studies of belief change in scenarios from social science, rather than economic game theory. Phenomena such as group polarisation, informational cascades, pluralistic ignorance, peer pressure, or epistemic bandwagoning all center around the epistemic and doxastic processes of agents in a social network. These mechanisms are more complex then those for individual agents. In a social network, the formation of opinions does not only depend, as it did in our earlier sections, on the individual agent's own prior epistemic-doxastic state, the new information it faces, and its individual belief revision mechanism. It is also crucially affected by the opinions and belief changes of other agents in the network.

The logical tools for belief change in this chapter can be adapted to incorporate the qualitative features of a social network, such as its topological configuration or its social hierarchy. We briefly present two instance of logical models for belief revision in this richer setting.

A first question that has been solved for social networks is whether informational cascades are due to irrational moves or mistakes in the reasoning steps of the agents or whether they are actually unavoidable by "rational" means.

7.8.3 Informational cascades

Informational cascades can occur when agents find themselves in a sequential network, obeying strict rules about what and how they can communicate with each other, and end up following the opinion of the preceding agents in the sequence, ignoring their own private evidence.

Example 7.18 (Choosing a restaurant)
A standard example is the choice between two restaurants A and B. You have prior private evidence that restaurant A is better than B, but you still end up choosing for B based on the fact that it has many more customers. You interpret the other customers' choice for restaurant B as conveying the information that they somehow know better. It is however very well possible that all others made their decision in exactly the same way as you. $\dashv$

Several models for informational cascades have been developed from 1992 onwards, after the term was first introduced by Bikhchandani, Hirshleifer, and Welch, with analyses usually in terms of Bayesian reasoning. Recent work by Baltag, Christoff, Hansen and Smets has used a combined probabilistic and qualitative Dynamic-Epistemic Logic. It then turns out that what might seem an irrational form of influence, manipulation or irregularity in the process of opinion formation, is actually the result of fully rational inference process. Moreover, the modeling makes explicit the agents' higher-order reasoning about their own as well as other agents' knowledge and beliefs.

Example 7.19 (An urn guessing scenario)
Consider the following information cascade based on a standard urn example used in the literature. Consider a sequence of agents $(i_1, i_2, ..., i_n)$ lined up in front of a room. The room contains one of two non-transparent urns U_W and U_B. It is common knowledge among the agents that nobody knows which urn is actually placed in the room. It is also common knowledge that Urn U_W contains two white balls and one black ball, and urn U_B contains one white ball and two black balls.

Now the agents enter the room one by one, and each agent draws a ball from the urn, looks at it, puts it back, and leaves the room. After leaving the room she publicly communicates her guess: urn U_W or U_B, to all the other agents. We assume that each agent knows the guesses of the people preceding her in the sequence before entering the room herself. It is crucial here (and realistic in many social settings) that while the agents communicate their guess about U_W or U_B, they do not communicate their private evidence, namely, the color of the ball they observed in the room. The standard Bayesian analysis of this example shows that if U_B is the real urn in the room and the first two agents i_1 and i_2 draw white balls (which happens with probability $\frac{1}{9}$), then a cascade leads everyone to the wrong guess U_A.

We can model this cascade in the Probabilistic Epistemic Dynamic Logic of Section 7.6. The reader should refer back to the notions introduced there. ⊣

Urn scenario, probabilistic epistemic model Consider the preceding urn example, with U_B the real urn in the room and the two first agents drawing white balls. The probabilistic epistemic model $\mathcal{M}_0$ of Figure 7.20 has equal odds for the initial situation, encoded in two worlds making it equally probable that U_W or U_B are true, and all agents know this. The actual state of affairs s_B satisfies the proposition U_B, while the other possible world s_W satisfies U_W. The relative likelihood of s_B versus s_W is $[1 : 1]_i$ for all agents i.

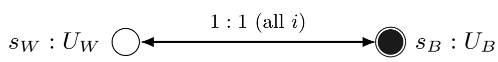


Figure 7.20: Probabilistic epistemic base model.

Urn Scenario, probabilistic event model To model the first observation of a ball by agent i_1, we use the probabilistic event model $\mathcal{E}_1$ depicted in Figure 7.21. At the moment of i_1's observation of a white ball, all other agents consider one of the following two events possible: either i_1 observes a white ball (event w_1) or she observes a black ball (event b_1). Only agent i_1 knows which event (w_1 or b_1) is actually taking place. For the event w_1 to happen, the preconditions are $pre(U_W) = \frac{2}{3}$ and $pre(U_B) = \frac{1}{3}$. All agents except i_1 consider both events equally likely and assign them the odds $1 : 1$. All this information in the successive visits by the agents is present pictorially in Figure 21.

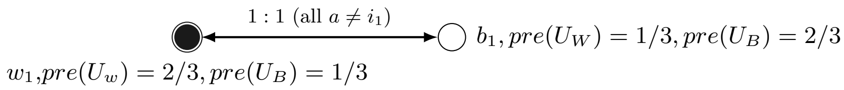


Figure 7.21: Probabilistic event model for the agent episodes.

To model the effect of the first observation, we combine the probabilistic epistemic model with the probabilistic event model using PDEL product update.

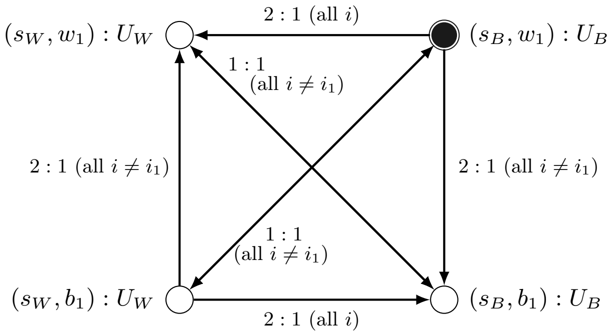


Figure 7.22: The result of probabilistic product update.

Urn scenario, the updates The product update of the initial epistemic probability model $\mathcal{M}_0$ with the event model $\mathcal{E}_1$ is the epistemic probability model $\mathcal{M}_0 \otimes \mathcal{E}_1$ illustrated in Figure 7.22, consisting of 4 possible worlds. In the world representing the actual situation, U_B is true and the first ball which has been observed was a white one w_1. The model makes it clear pictorially that agent i_1 *knows* which ball she observed: there are no a_1-arrows between the two top states in the model and the bottom states. The other agents (different from i_1) consider all four states possible and for them it is common knowledge that, if agent 1 observed a white ball (w_1), then she considers U_W to be twice as likely as U_B, and in case she would have observed a black ball b_1, then she would consider U_B twice as likely as U_W. In particular, the new model indicates that all the other agents cannot exclude any epistemic possibility as long as i_1 has not yet announced her guess publicly.

Next, the public announcement of a_1's guess is modeled as a further PAL update resulting in a third model $\mathcal{M}_1$ in which two states (those not compatible with the guess of the agent) are deleted.

The event of the second agent entering the room, drawing a ball from the urn, and announcing its guess publicly will induce further changes of the model $\mathcal{M}_1$, analogous to those in the first step, resulting in $\mathcal{M}_2$.

Things become more interesting when agent a_3 enters the room and observes a black ball. The event model $\mathcal{E}_3$ of this action will trigger an update of $\mathcal{M}_2$ to the model $\mathcal{M}_2 \otimes \mathcal{E}_3$. But in this model, a_3 still considers U_W more probable than U_B, irrespective of the result of her private observation. Hence, if i_3 then announces her guess of U_W, it deletes no more worlds: the model $\mathcal{M}_3$ after the announcement stays the same as before. From now on an informational cascade will unfold, and all further agents in the sequence announce the guess U_W.

Much more can be said about this example, but it will be clear how the machinery of this chapter can analyse agents' beliefs in social phenomena such as cascades. Moreover, it yields interesting insights such as the following. Agents who are logical omniscient and perfect reasoners, but only announce their private guesses, simply do not have the tools which can always prevent an informational cascade from happening.[39]

7.8.4 Influence in social networks

To show the reach of the methods in our chapter, we also consider a second basic phenomenon in the social realm, that of agents in social networks.

Different types of epistemic and doxastic attitudes arise in, and maintain, social settings. Seligman, Liu & Girard have recently proposed a logical framework for investigating how agents' beliefs, or knowledge, are formed and changed under the influence of the beliefs of other agents who belong to the same social community. The system of *"Facebook Logic"* designed for this has a social as well as an epistemic dimension. Both these dimensions are needed to specify basic social relations between agents, such as friendship, agents' epistemic attitudes, and their entanglement. We describe a few basics here.

Epistemic social networks We start with situations where agents can be friends, though they need not know exactly who are their friends. An *epistemic social network model* $M =< W, I, \sim_i, \asymp_w, V >$ consists of a set of states W, a set of agents I, and an epistemic relation $\sim_i$ for each agent $i \in I$. Each state $w \in W$ comes equipped with a binary irreflexive and symmetric friendship relation $\asymp_w$ over the set of agents, and, as usual, V is a valuation map assigning subsets of $W \times A$ to propositional variables.

The *epistemic friendship language* is a multimodal formalism given by:

$$\varphi ::= p \mid n \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi \mid F\varphi \mid A\varphi$$

where K is the standard epistemic knowledge modality, F is a friendship modality which is read as "for all my friends", and A is an auxiliary universal modality which quantifies over all agents.[40] These operators are interpreted in epistemic social network models as follows:

$$\begin{array}{lll} M, w, i \models p & \text{iff} & (w, i) \in V(p), \quad \text{for } p \in Prop \\ M, w, i \models K\varphi & \text{iff} & M, v, i \models \varphi \text{ for every } v \sim_i w \\ M, w, i \models F\varphi & \text{iff} & M, w, j \models \varphi \text{ for every } j \asymp_w i \\ M, w, i \models A\varphi & \text{iff} & M, w, j \models \varphi \text{ for every } j \in I \end{array}$$

[39] Baltag, Christoff, Hansen and Smets also present a second formalisation in which probabilistic reasoning is replaced by a "counting" heuristics in terms of pieces of evidence.

[40] Obvious existential modalities exist, too, such as the dual $< F >$ of the modality F, defined as $\neg F \neg$.

This simple language is quite expressive. Combinations of operators such as KFp or $FKFKp$ make complex statements about what members of a community know about their friends' knowledge. To boost power of defining social situations even further, Facebook Logic makes use of a technical device from "Hybrid Logic", the indexical "downarrow pointer" $\downarrow n$ which introduces a name n to refer to the current agent.[41] Using various defined notions in this hybrid setting, Seligman, Girard and Liu et al. can define interesting anaphoric expressions, such as $\downarrow x < F > K@_n < F > x$ which says that "I have a friend who knows that n is friends with me". While all this seems geared toward knowledge, analogous methods also apply to agents' *beliefs* in social networks.

Network dynamics This social epistemic setting supports various kinds of dynamics, such as agents learning facts that change their knowledge or beliefs. The relevant events here can be of the kinds we have discussed in our earlier hard and soft update scenarios for belief.

But the setting of Facebook Logic also includes interesting new dynamic phenomena, such as changes in agents' beliefs under social influence. For instance, a typical 'dynamic network rule' would be that

An agent comes to believe a proposition p iff all her friends believe p.

This induces changes in what is believed in the network, and hence iterations can start. As we saw in Section 7.8 on belief change over time, these iterations can either stabilise in a state where all agents have a acquired a permanent belief or a permanent disbelief in p, that could be viewed as the resulting *community opinion*. But updates can also start oscillating, as often happens in dynamical systems for population behavior, modeling cycles in 'public opinion'.

Of course, other qualitative network rules are possible, as well as more sophisticated quantitative ways of measuring the dynamics of influence.

Finally, going beyond that, there can also be changes in the network itself, like when friendship relations start changing by adding or deleting links.

Network logics All the preceding kinds of updates can be described in the same model transformation style as before. For instance, the new dynamic network rules induce dynamic updates that change the current truth value of propositions through a sort of "predicate re-assignments", or "substitutions", of the form

$$p := \varphi(p)$$

Accordingly, dynamic logics with modalities for local and global network evolution in terms of agents' knowledge and beliefs can be found, resembling the epistemic-doxastic-temporal types that we have studied before.

However, there are also interesting new technical developments here. Seligman, Girard and Liu define a very general class of dynamic operators and actions

[41] Semantically, this involves an additional 'assignment function' g with clauses (a) $M, w, i \models n$ iff $g(n) = i$, for $n \in ANom$, (b) $M, w, i \models \downarrow n\varphi$ iff $M[\overset{i}{n}], w, a \models \varphi$, where $M[\overset{i}{n}]$ is the result of changing the model M so that n names agent i.

on the above epistemic social models, leading to a system of "General Dynamic Dynamic Logic" that generalises our earlier Priority Update while still remaining reducible to PDL. Other relevant topics in this setting includes the work of Zhen and Seligman on peer pressure which investigates the effect of social relations on the logical dynamics of preference change. A few further recent references on dynamic network logics will be found in Section 7.10.

7.9 Further directions

There are many further streams in the literature on belief revision than what we have covered here, both in terms of topics and approaches. In this section, we will give a brief panorama of further directions.

7.9.1 Postulational and constructive approaches

In addition to dynamic-epistemic logics for belief change based on explicit model transformations, there is another tradition, that of Segerberg's Dynamic Doxastic Logic (DDL). It is closer to AGM by leaving the nature of the revision steps abstract, while incorporating intuitive *postulates* on belief revision as axioms in a modal logic. These are both valid styles of approach, and we briefly explore differences as well as ways in which the two complement each other.

Dynamic-doxastic logic has abstract modal operators describing update transitions in abstract universes, relational or functional.[42] The format of interpretation is as follows. Let M be a model, $[\![P]\!]$ the set of worlds in M satisfying P, and $M * [\![P]\!]$ some new updated model. Then we set

$$M, s \models [*P]\varphi \text{ iff } M * [\![P]\!], s \models \varphi$$

The minimal modal logic K is valid in this semantic framework, and further axioms constrain relation changes for special revision policies, leading to special modal logics extending K.

Instead of compiling a detailed list of differences with our dynamic-epistemic logics, we focus on an interesting bridge between the two approaches, proposed by van Benthem around 2007. It works in terms of standard "frame correspondence" for modal axioms. Given some modal axioms on an abstract update operation, to which extent will these axioms enforce the precise constructive recipes employed in dynamic-epistemic logic?

Correspondence on update universes To make a modal correspondence analysis work, we need a suitably general semantic framework behind the concrete models of earlier sections, in the spirit of DDL. Consider any family $\mathbf{M}$ of pointed epistemic models (M, s), viewed as an "update universe". Possible changes are given as a family of update relations $RP\ (M, s)(N, t)$ relating pointed models, where the index set P is a subset of M: intuitively, the proposition triggering the

[42]Concrete DDL models can be pictured graphically as Lewis spheres for conditional logic or neighborhood models.

update. One can think of the R as recording the action of some update operation $\heartsuit$ occurring in the syntax of our language that depends on the proposition P. Here different operations from earlier sections can have different effects: from hard updates $!\varphi$ to soft updates $\Uparrow\varphi$.

For each formula φ, let $[\![\varphi]\!]$ be the set of worlds in M satisfying φ. We set

$$M, s \models < \heartsuit\varphi > \psi \text{ iff there exists a model } (N, t) \text{ in } \mathbf{M}$$

where $R[\![\varphi]\!](M, s)(N, t)$ and $(N, t) \models \psi$

Now we can analyse given dynamic modal principles precisely, and we will give two examples of such a correspondence analysis.[43]

Correspondence for eliminative update One obvious choice for constraints on update operations are the earlier recursion axioms that gave us the heart of the dynamics of knowledge and belief change. Here is the outcome of a correspondence analysis for public announcement logic PAL.[44] By going through the content of all recursion axioms, with a special role for the knowledge clause, we can see that the PAL axioms essentially enforce world elimination as its interpretation of the update for $!\varphi$. Technically, however, there is some "slack", in terms of a well-known modal structure-preserving map:[45]

Theorem 7.12
An update universe satisfies the substitution-closed principles of PAL iff its transition relations F_P are partial p-morphisms defined on the sets P. $\dashv$

To force the p-morphisms to become embeddings as submodels, a few further twists are needed, analysing some further features of the treatment of propositional atoms ("context-dependent", or not), as well as a further modal recursion axiom for a global "existential modality". We omit these details here.

Discussion: refinements Update universes are reminiscent of the *protocol* version of PAL considered in Section 7.8, where available transitions can be restricted by the model. Our correspondence perspective also works for modified recursion axioms on protocol models, where update functions on domains may now become partial. Another generalisation of our semantics in Sections 7.4 and 7.5 suggested by the above is a possible context-dependent interpretation of proposition letters, not as sets of worlds, but as sets of pairs (M, w) where M is a model

[43] More precisely, we are interpreting our language in a three-index format $\mathbf{M}$,M, s, and for the accessibility relations R in this update universe $\mathbf{M}$, we have that $(M, s)R(M, t)$ iff Rst in M, without any jumps out of the model M.

[44] In what follows, we must sidestep one particularity of PAL and related dynamic-epistemic logics: its use of non-schematic axioms such as $<!\varphi > q \leftrightarrow (\varphi \wedge q)$ that is not closed under substitutions of arbitrary complex formulas for the atomic proposition q. This can be solved by going to special cases such as $<!\varphi > T \leftrightarrow \varphi$. We refer to the literature for details.

[45] The recursion axiom for negation imposes functionality of the update relations. The left and right directions of the knowledge recursion axiom enforce precisely the two central relation-preserving clauses for a p-morphism.

and w a world in M. In that case, the logic will become the substitution-closed schematic core of PAL – and similar observations holds for DEL in general.

Correspondence for belief change The same style of analysis applies to update principles for plausibility orderings. We merely state the outcome here:

Theorem 7.13
The recursion axioms of the dynamic logic of radical upgrade and conditional belief hold for an update operation on a universe of pointed plausibility models iff that operation is in fact radical upgrade. $\dashv$

Here is what this says. AGM-style postulates on changes in beliefs alone do not fix the relational transformation: we need to constrain the changes in conditional beliefs, since the new plausibility order encodes all of these. But there is an easier road as well, in line with observations in Sections 7.3 and 7.4:

Theorem 7.14
Radical upgrade is the only update operation validating the given recursion axioms for atoms, the Boolean operations, and safe belief. $\dashv$

One could also do correspondence analysis directly on the AGM postulates, but this would take us too far afield here. Likewise, a correspondence analysis of update postulates can also be performed for dynamic-epistemic logic and recursion axioms for product update as defined in Section 7.5. One reason why these techniques work is the "Sahlqvist form" of many recursion axioms for belief revision, making them amenable to standard modal techniques.

7.9.2 Belief revision versus nonstandard consequence

Van Benthem pointed out in 2008 that update of beliefs under hard or soft information is an alternative to *nonstandard notions of consequence.* Classical consequence says that all models of premises P are models for the conclusion C. McCarthy in 1980 famously pointed out how problem solving goes beyond this. A "circumscriptive consequence" from P to C says that C is true in all the *minimal* models for P, where minimality refers to a relevant comparison order $\leq$ for models, Circumscription supports non-monotonic consequence relations that show resemblances to our earlier conditional logics of Section 7.3.

This is reminiscent of plausibility models for belief. Starting from initial information we must reach a goal as new information comes in. Non-monotonic logics leave such events implicit in the background, while dynamic-epistemic logics provide an alternative in terms of the beliefs that problem solvers have as they go through their process of observation and inference.[46]

Thus, there is an interesting duality between explicit treatments of belief change in logic like the ones in this chapter, and approaching the same issues via new consequence relations.[47] The precise interplay between these two views of belief revision is far from being understood.

[46]The latter approach even suggests new notions of circumscription, for instance, versions that take the information in the premises as soft rather than hard.

[47]This triangle of perspectives has also been noted in the classical AGM survey of

7.9.3 Belief revision and evidence

Belief revision has been studied on relatively coarse semantic models in this chapter. However, it is quite feasible, as has been done in the AGM tradition and also in dynamic-doxastic logic, to work with more fine-grained models for *evidence*. Staying close to modal logic, one obvious candidate are "neighborhood models" where each world is connected to a family of sets of worlds, its neighborhoods, that may be considered as standing for pieces of evidence that we have concerning the location of the actual world. Such pieces of evidence may be consistent, but they can also contradict each other, leading to a much more realistic view of settings in which we need to form our beliefs.

In particular, van Benthem and Pacuit have recently proposed evidence models for belief that interpret notions such as having evidence for φ:

$$M, s \models \bigodot \varphi \text{ iff there is a set } X \text{ in } N \text{ with } M, t \models \varphi \text{ for all } t \in X$$

as well as cautious beliefs based on what is true in the intersections of maximal families of mutually consistent pieces of evidence.

The resulting static logics are simple, but not cheap, and combine standard modalities with generalised monotone evidence modalities. Such models lend themselves well to a generalised dynamic-epistemic treatment, in terms of actions of adding, removing, modifying, and combining evidence. For instance, the single update operation for PAL events $!\varphi$ will now decompose into several actions: adding evidence for φ, and removing evidence for $\neg\varphi$. We refer to the literature for the resulting recursion axioms. One interesting feature of this setting is that analysing belief contraction, traditionally one of the more difficult AGM operations, becomes quite perspicuous.

Fine-grained models like this have been proposed in formal epistemology recently, since they combine a finer representation of evidence and belief with the continued availability of the semantic techniques of this chapter.

7.9.4 Combining information and preference

In this chapter, we have shown how logical dynamics deals with agents' knowledge and beliefs and informational events that change these. But agency also involves a second major system, not of information but of "evaluation", as recorded in agents' *preferences* between worlds or actions. It is preference that guides our actions, rather than just possession of hard or soft information.

Preference logics can be studied with the modal and dynamic techniques that we have discussed here. In fact, one can reinterpret our plausibility models as models with worlds, or just any kind of objects, carrying a binary "betterness ordering" for each agent. These structures occur in decision theory, where worlds are outcomes of actions, and game theory, where worlds are histories, with preferences for different players.

When added to our earlier logics, a simple modality of betterness can define equilibrium notions for game solution, as well as normative deontic notions in

Gärdenfors & Rott around 1995. Also relevant here is the modal-style analysis of update operations by Ryan & Schobbens around the same time.

general agency. Moreover, our earlier relation changes now can do duty as preference changes, for instance, in response to a suggestion or a command from some agent with sufficient authority. Finally, preference, knowledge, and belief often occur *entangled*, in notions such as "ceteris paribus preference", obligations based on current information, or qualitative versions of expected utility. The above dynamic-epistemic techniques will still work for such combined systems.

7.9.5 Group belief and merge

Our forays into multi-agent scenarios in Section 7.8 have only scratched the surface of social aspects of belief change. In addition to individual belief changing acts, a social setting suggests many further dynamic operations at the level of groups that can have information and act collectively. One basic notion that has received attention in the post-AGM literature on belief revision is "belief merge" of the individual agents forming a group. Related notions occur in the area of Judgment Aggregation, while the mentioned entanglement with preference also suggests strong analogies with Social Choice Theory (see below).

There are quite a few logical challenges here since the logical structure of collectives is quite intricate, witness the semantics of plurals and collective expressions in natural language, which is by no means a simple extension of the logic of individuals and their properties. Indeed, groups bring in new features, since they are usually much more than a flat set of individual agents. There can be hierarchical structure of informational status (trustworthiness), or of preferential status (authority and power), and this structure can crucially affect how collective behavior arises, and how beliefs of individual members change.

There are promising formal models for this sort of structure that are congenial to the dynamic-epistemic approach in this chapter. Andréka, Ryan & Schobbens noted around 2002 that merge operations typically need a structured view of groups as graphs with a dominance order. The resulting "prioritised relation merge" can be defined as follows. Given an ordered priority graph $G = (G, <)$ of indices for individual relations that may have multiple occurrences in the graph, the *merged group priority relation* is:

$x \leq_G y$ iff for all indices $i \in G$, either $x \leq_i y$, or
there is some $j > i$ in G with $x <_i y$.

This is slightly reminiscent of the 'priority update' in our Section 7.5.3, and this graph setting has turned out to apply in many areas, including the above topics, but also inquiry and learning in the presence of a structured "agenda" of investigation, or preference in the presence of an ordered graph of "criteria" for judging betterness. As has been shown by Girard & Liu around 2008, priority graphs lend themselves well to algebraic and logical treatment, especially, when we focus on two major dynamic operations of merging priority graphs: "parallel composition" with no links between the two graphs, and "sequential composition" where all nodes in one graph dominate all nodes in the other.

Ideas from this circle also apply to our earlier treatment of belief change. In 2009, van Benthem analysed belief change as a process of social choice, merging signals from different sources (past experience, current observation, etc.), and made a connection with some natural intuitions from Social Choice Theory:

Theorem 7.15
A preference aggregation function is a Priority Update iff it satisfies Permutation Invariance, Locality, Abstentions, Closed Agenda, and Overruling. $\dashv$

We do not provide details here, but the point of results like this is that they capture policies for belief revision or belief merge in terms of postulates from the economics and social science literature. Many further relevant connections, relating AGM postulates to choice principles in the foundations of decision theory and economics, were found earlier on in the work of Rott.

7.9.6 Belief revision and social choice

Dynamic epistemic logic fits well with social choice theory, where group actors form collective preferences from the preferences of individual members. Merging social choice with the main concerns of this chapter can provide two things: informational structure, and more finely-grained procedure. For the first, think of Arrow's Theorem, and the horror of a dictator whose opinions are the social outcome. Even if it is common "de dicto" knowledge that there is a dictator, this does no harm if there is no person whom we know "de re" to be the dictator. Not even the dictator herself may know she is one. To see the real issues of democracy, we need social choice plus epistemic logic.

Also, rules for voting represent just a fragment of richer practices of communication and debate. One can study how groups arrive at choices by deliberating, and ways in which agents then experience preference changes. This is reminiscent of two dynamic processes that we considered in Section 7.8: deliberation, and belief adaptation in social groups. Again, the preceding combined approach seems called for here, especially when we think of groups that also have means of communication, giving them designated information channels. Two concrete sources of relevant scenarios for this broader study might be Argumentation Theory, with its studies of rules for fair debate, and the Law.

7.10 Notes

In this section we provide some major references for the material presented in this chapter, offering also some pointers to the literature for further study.

Section 7.1: Basics of belief revision The AGM theory of belief revision goes back to a classic paper by Alchourrón, Gärdenfors, and Makinson (1985). This approach has given rise to an extended series of papers and books on the topic, of which we mention (Gärdenfors, 1988), and (Gärdenfors and Rott, 1995). The AGM postulates for revision, contraction and expansion have been the subject of much philosophical discussion. One basic example is the Recovery Postulate, which, as motivated by the principle of minimal change, prescribes that a contraction should remove as little as possible from a given theory T. Extended discussions of this principle can be found in (Hansson, 1991; Fuhrmann, 1991; Levi, 1991; Niederée, 1991; Hansson, 1997). A concise up to date overview of standard belief revision theory is found in the chapter on Theory Replacement in (Kyburg and Teng, 2001), while Rott (2001) provides an in depth study of standard belief revision theory in the context of nonmonotonic reasoning.

Belief revision theory has important links with the logic of conditionals, via the "Ramsey Test". The original source is a short note by Ramsey (1990), while basic modern results are provided by Gärdenfors (1986, 1988). The mentioned semantic modal perspective on Gärdenfors' Impossibility Result comes from Baltag and Smets (2010). A different concept of belief update as involving world change, mentioned in one of the footnotes, is studied in (Katsuno and Mendelzon, 1992).

Section 7.2: Modal logics of belief revision The original modal logic approach to static notions of knowledge and belief is ascribed to Hintikka (1962). Major sources for the Propositional Dynamic Logic of actions are (Harel, 1984; Harel, Kozen, and Tiuryn, 2000), while Segerberg (1995, 1998, 1991) provides classical references for Dynamic Doxastic Logic. In the context of PDL, we refer to van Benthem (1989) for dynamic operators that mirror the AGM operations, and to (Rijke, de, 1994) for extensions of this approach.

Early quantitative systems for belief revision in the style of Dynamic Epistemic Logic were proposed by Aucher (2003) and by van Ditmarsch and Labuschagne (2003). We also mentioned the "ranking models" of Spohn (1988). The BMS notion of "product update" refers to work by Baltag, Moss, and Solecki (1998). Further details and references on Public Announcement Logic (PAL), can be found in chapter 6 on DEL in this volume as well as in work by van Ditmarsch, van der Hoek, and Kooi (2007), and by van Benthem (2011). Landmark papers on PAL and its extensions include (Plaza, 1989; Gerbrandy and Groeneveld, 1997), and the dissertation (Gerbrandy, 1998). An early PAL-style approach to AGM belief expansion is found in van (van Ditmarsch, van der Hoek, and Kooi, 2005). The much more general approaches on which this chapter is based are by van Benthem (2007a) and Baltag and Smets (2008).

Section 7.3: Static base logics Basic modal logics for belief such as "KD45" and many others can be found in the textbook by Meyer and van der Hoek (1995).

The material on connected plausibility models is based on work by Baltag and Smets (2008), while Baltag and Smets (2006b) developed the correspondence between plausibility frames and epistemic plausibility frames. Plausibility models allowing non-connected orderings are used extensively by van Benthem (2007a), and by van Eijck and Sietsma (2008). World-based plausibility models with ternary relations are discussed by van Benthem (2007a), by Baltag and Smets (2006b), and by Board (2002). However, no doubt the classic predecessor to all of this is the first semantic modeling for AGM theory given by Grove (1988). For similar models in conditional logic for AI an philosophy, see Shoham (1988) and Spohn (1988). For background in classical conditional logic, cf. (Lewis, 1973) for completeness theorems on connected orderings, and contributions by Burgess (1981), and by Veltman (1985) on preorders. An elegant setting based on partially ordered graphs is found in (Andreka, Ryan, and Schobbens, 2002).

Safe belief on plausibility models has been discovered independently in many areas, as diverse as AI (Boutilier (1994), Halpern (1997)), multi-agent systems (Shoham and Leyton-Brown (2008)), and philosophy (Stalnaker (2006)). Safe belief is related to defeasible knowledge in formal epistemology (Baltag and Smets, 2006b, 2008), and to modal preference logics (Liu, 2008; Girard, 2008) – with some related work in (Lenzen, 1980). Our definition of strong belief refers to that of Baltag and Smets (2008, 2013), but predecessors are found in work by Battigalli and Siniscalchi (2002) and Stalnaker (1998) in economic game theory.

Section 7.4: Belief revision by model transformations This section is based on proposals by van Benthem (2007a) and by Baltag and Smets (2008) on complete dynamic-epistemic logics with recursion axioms for belief revision under hard and soft announcements. Earlier sources for soft announcements are the default logic of Veltman (1996), and the minimal conservative revision of Boutillier (1993). Different rules for plausibility change in models of belief revision can be found in (Grove, 1988) and Rott (Rott, 2006), while radical and conservative upgrade in dynamic-epistemic style are highlighted by van Benthem (2007a), and by Baltag and Smets (2008).

Section 7.5: General formats for belief revision The PDL-format for relation transformers is due to van Benthem and Liu (2007). Much more complex mechanisms in this line have been studied by van Eijck and his collaborators (van Eijck and Wang (2008); van Eijck and Sietsma (2008); van Eijck (2008)), extending the methods of van Benthem, van Eijck, and Kooi (2006) to deal with a wide range of belief changes defined by PDL programs, and by Girard, Seligman, and Liu (2012), who merge the idea of PDL transformers with product models for DEL to describe information flow in complex social settings.

The classical source for the general DEL framework with product update and recursion axioms is (Baltag et al., 1998), while more examples and extensions can be found in (Baltag and Moss, 2004; Baltag and Smets, 2008, 2006b,a,c, 2007; van Benthem, 2007a). Larger treatises on the subject are offered by van Ditmarsch, van der Hoek, and Kooi (2007) and by van Benthem (2011). Recursion axioms for "relativised common knowledge" are found in (van Benthem et al., 2006). See also Chapter 6 in this volume on DEL for more material.

Our discussion of event models as triggers of soft information change is based on the Priority Update of Baltag and Smets (2006b) (with a precursor in (Aucher, 2003), and more broadly, in (van Benthem, 2002)). The link to "Jeffrey Update" refers to Jeffrey (1983). Note that the coin examples in this section are due to Baltag and Smets (2008).

Section 7.6: Belief revision and probability dynamics The probabilistic dynamic epistemic logic of this section, including probabilistic event models and a probabilistic product update mechanism, was developed by van Benthem, Gerbrandy, and Kooi (2006). A different approach by Baltag and Smets (2007) uses probabilistic models in line with the Popper-Rényi theory of conditional probabilities. A masterful overview of static logics to reason about uncertainty is by

Halpern (2003). New developments include a logic for reasoning about multi-agent epistemic probability models, by van Eijck and Schwarzentruber (2014).

Section 7.7: Time, iterated update, and learning Hodkinson and Reynolds (2006) and van Benthem and Pacuit (2006) provide surveys of the basics of epistemic temporal logics (see also Chapter5 in this volume). Agents properties like perfect recall are discussed extensively by Halpern and Vardi (1989), including a finite automata perspective that is also discussed in a dynamic-epistemic setting by Liu (2008).

For epistemic temporal logics and complexity, the classic source is provided by Halpern and Vardi (1989). Connections to further work in computational logic are given by Parikh and Ramanujam (2003), and foundations of games are surveyed by van Benthem and Pacuit (2006) in a dynamic-epistemic format. Miller and Moss (2005) prove the high complexity of PAL with a PDL-style iteration operator.

Protocol models for DEL and their connections with epistemic-temporal logics are found in van (van Benthem and Liu, 2004), and in a more general setting in (van Benthem, Hoshi, Gerbrandy, and Pacuit, 2009), including a formalisation of state dependent protocols. Hoshi (2009) axiomatises the laws of special protocol logics and makes connections with procedural information in epistemology.

The extensions to belief refer to work by van Benthem and Dégrémont (2008) and by Dégrémont (2010). For related work by Bonanno and others, see Chapter 5 of this volume.

Our treatment of iterated belief upgrades and limit behavior and its main stabilisation theorems is based on work by Baltag and Smets (2009a,b), with van van Benthem (2002) as an early predecessor. More information on bisimilarity between pointed Kripke models can be found in the textbook by Blackburn, de Rijke, and Venema (2001).

For formal learning theory, we refer to work by Kelly (1998a,b), by Kelly, Schulte, and Hendricks (1995) and by Hendricks (2003, 2001). Various links with dynamic-epistemic logic have been explored by Gierasimczuk (2010), by Baltag and Smets (2011), by Dégrémont and Gierasimczuk (2011, 2009), and by Gierasimczuk and de Jongh (2013). The major results on the learning power of different belief revision methods were taken from (Baltag, Gierasimczuk, and Smets, 2011). A brief overview of the use of temporal-epistemic logic in a learning theoretic context has been sketched in (Gierasimczuk, Hendricks, and de Jongh, 2014).

Finally, various connections with propositional dynamic logics can be found in the "knowledge programs" of Fagin, Halpern, Moses, and Vardi (1995), in the "epistemic-temporal PDL" of van Benthem and Pacuit (2007), and in the use of PDL-definable protocols by Wang, Sietsma, and van Eijck (2011).

Section 7.8: Belief, games, and social agency Logical tools and methods for solution concepts in game theory have been studied widely, see the surveys by van der Hoek and Pauly (2006), by Bonanno (1991, 1992), and by Battigalli, Brandenburger, Friedenberg, and Siniscalchi (2014). An explanation of the epistemic significance of the Muddy Children Puzzle is found in (Fagin et al., 1995).

The main sources for our treatment of Backward Induction are from van Benthem (2007b, 2011), while (van Benthem, 2014a) is an extensive exploration of the realm of logic in games. For the specific results on limits of iterated radical upgrades, see also work by Baltag, Smets, and Zvesper (2009), and by van Benthem and Gheerbrant (2010).

The logical analysis of informational cascades work presented here is from Baltag, Christoff, Ulrik Hansen, and Smets (2013). A new probabilistic logic of communication and change in cascades is proposed in Achimescu, Baltag, and Sack (2014). Facebook Logic and its recent developments can be found in a series of papers by Seligman, Liu, and Girard (2011); Liu, Seligman, and Girard (2014); Seligman, Liu, and Girard (2013a,b). Sano and Tojo (2013) apply ideas from the facebook logic setting in the context of DEL, while other merges are found in (Christoff and Ulrik Hansen, 2013). A study of the social effects on preference dynamics is presented by Liang and Seligman (2011).

Section 7.9: Further directions Classic sources for Dynamic Doxastic Logic, as mentioned earlier, are provided by Segerberg (1995, 1991, 1999). See also work by Segerberg and Leitgeb (2007) for an extensive discussion of the DDL research program. The work reported on connections with DEL via frame correspondence is from van Benthem (2014b). Another systematic approach relating modal logics and belief revision is by Ryan and Schobbens (1997).

For circumscriptive consequence, two main sources are by McCarthy (1980) and by Shoham (1988). Belief revision versus nonstandard consequence as an approach to belief is discussed by van Benthem (2011), but the interface of AGM Theory and nonmonotonic logic was already discussed extensively by Gärdenfors and Rott (1995), and by Rott (2001).

Neighborhood models for modal logic go back to Segerberg (1971). Evidence dynamics on neighborhood models was developed by van Benthem and Pacuit (2011). Further developments including links with justification logic are given by van Benthem, Fernández-Duque, and Pacuit (2014), while applications to formal epistemology are found in (Fiutek, 2013). For neighborhood models in Dynamic Doxastic Logic, see (Segerberg, 1995; Girard, 2008).

Preference change in dynamic-epistemic style has been studied in by van Benthem and Liu (2004), while the monograph of Liu (2011) is an extensive study with references to other streams in preference logic and deontic logic. For richer constraint-based views of preference in terms of ordered criteria, see also (Rott, 2001). An elegant technical framework are the priority graphs of Andreka et al. (2002). For entangled dynamic-epistemic systems that combine preference, knowledge, and belief see (Girard, 2008; Liu, 2011).

Belief merge has been studied using AGM techniques by Maynard-Reid and Shoham (1998), while List and Pettit (2004) on "judgment aggregation" is relevant too. In a DEL setting, a relevant system is the E-PDL of van van Benthem et al. (2006). The analysis of Priority Update as a form of social choice is from van Benthem (2007c).

Acknowledgements The authors thank Thomas Bolander, Barteld Kooi, Wiebe van der Hoek and Shane Steinert-Threlkeld for valuable feedback during the writing of this chapter. The contribution of Sonja Smets was funded in part by the VIDI grant 639.072.904 of the Netherlands Organisation for Scientific Research and by the European Research Council under the 7th Framework Programme FP7/2007-2013/ERC Grant no. 283963.

References

Achimescu, A., A. Baltag, and J. Sack (2014). The probabilistic logic of communication and change. Presented at LOFT 2014, Bergen.

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

Andreka, H., M. Ryan, and P. Y. Schobbens (2002). Operators and laws for combining preference relations. *Journal of Logic and Computation 12*, 13–53.

Aucher, G. (2003). A combined system for update logic and belief revision. Master's thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam.

Baltag, A., Z. Christoff, J. Ulrik Hansen, and S. Smets (2013). Logical models of informational cascades. *Studies in Logic Volume, College Publications. 47*, 405–432.

Baltag, A., N. Gierasimczuk, and S. Smets (2011). Belief revision as a truth-tracking process. In *Proceedings of the 11th conference on Theoretical aspects of rationality and knowledge*, pp. 187–190. ACM.

Baltag, A. and L. S. Moss (2004). Logics for epistemic programs. *Synthese 139*(2), 165–224.

Baltag, A., L. S. Moss, and S. Solecki (1998). The logic of public announcements, common knowledge, and private suspicions. In *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '98, San Francisco, CA, USA, pp. 43–56. Morgan Kaufmann Publishers Inc.

Baltag, A. and S. Smets (2006a). Conditional doxastic models: a qualitative approach to dynamic belief revision. *Electronic Notes in Theoretical Computer Science 165*, 5–21.

Baltag, A. and S. Smets (2006b). Dynamic belief revision over multi-agent plausibility models. In W. van der Hoek and M. Wooldridge (Eds.), *Proceedings of LOFT'06*, pp. 11–24. Liverpool: University of Liverpool.

Baltag, A. and S. Smets (2006c). The logic of conditional doxastic actions: a theory of dynamic multi-agent belief revision.

Baltag, A. and S. Smets (2007). Probabilistic dynamic belief revision. In J. van Benthem, S. Ju, and F. Veltman (Eds.), *Proceedings of LORI'07*. College Publications, London.

Baltag, A. and S. Smets (2008). A qualitative theory of dynamic interactive belief revision. In G. Bonanno, W. van der Hoek, and M. Wooldridge (Eds.), *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games, pp. 9–58. Amsterdam University Press.

Baltag, A. and S. Smets (2009a). Group belief dynamics under iterated revision: Fixed-points and cycles of joint upgrades. In *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge (TARK'09)*, pp. 41–50. ACM.

Baltag, A. and S. Smets (2009b). Learning by questions and answers: From belief- revision cycles to doxastic fixed points. *LNAI Lecture Notes in Computer Science 5514*, 124–139.

Baltag, A. and S. Smets (2010). A semantic view on ramsey's test. In X. Arrazola and M. Ponte (Eds.), *LogKCA-10, Proceedings of the Second ILCLI International Workshop on Logic and Philosophy of Knowledge, Communication and Action*, pp. 119–134. Univ. of the Basque Country Press.

Baltag, A. and S. Smets (2011). Keep changing your beliefs and aiming for the truth. *Erkenntnis 75(2)*, 255–270.

Baltag, A. and S. Smets (2013). Protocols for belief merge: Reaching agreement via communication. *Logic Journal of IGPL 21*(3), 468–487.

Baltag, A., S. Smets, and J. Zvesper (2009). Keep 'hoping' for rationality: a solution to the backward induction paradox. *Synthese 169*, 301–333.

Battigalli, P., A. Brandenburger, A. Friedenberg, and M. Siniscalchi (2014). *Epistemic Game Theory: Reasoning about Strategic Uncertainty.* Manuscript under construction.

Battigalli, P. and M. Siniscalchi (2002). Strong belief and forward induction reasoning. *Journal of Econonomic Theory 105*, 356–391.

van Benthem, J. (1989). Semantic parallels in natural language and computation. In H. D. Ebbinghaus (Ed.), *Logic Colloquium Granada 1987*, pp. 331–375. Amsterdam North Holland.

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

van Benthem, J. (2007a). Dynamic logic for belief revision. *Journal of Applied Non-classical Logics 17*(2), 129–155.

van Benthem, J. (2007b). Rational dynamics and epistemic logic in games. *International Game Theory Review 9*, 13–45.

van Benthem, J. (2007c). The social choice behind belief revision (lecture presentation). In *Workshop Dynamic Logic Montreal.*

van Benthem, J. (2011). *Logical Dynamics of Information and Interaction.* Cambridge University Press.

van Benthem, J. (2014a). *Logic in Games.* MIT Press.

van Benthem, J. (2014b). Two logical faces of belief revision. In R. Trypuz (Ed.), *Outstanding Contributions to logic: Krister Segerberg on Logic of Actions*, pp. 281–300. Springer.

van Benthem, J. and C. Dégrémont (2008). Multi-agent belief dynamics: bridges between dynamic doxastic and doxastic temporal logics. In G. Bonanno, B. Löwe, and W. van der Hoek (Eds.), *Logic and the Foundations of Game and Decision Theory (LOFT 8)*, Number 6006 in LNAI, pp. 151 – 173. Springer.

van Benthem, J., D. Fernández-Duque, and E. Pacuit (2014). Evidence and plausibility in neighborhood structures. *Annals of Pure and Applied Logic 165*, 106–133.

van Benthem, J., J. Gerbrandy, and B. Kooi (2006). Dynamic update with probabilities. In *Proceedings LOFT'06, Liverpool*. University of Liverpool.

van Benthem, J. and A. Gheerbrant (2010). Game solution, epistemic dynamics and fixed-point logics. *Fundamenta Informaticae 1*(4), 19–41.

van Benthem, J., T. Hoshi, J. Gerbrandy, and E. Pacuit (2009). Merging frameworks for interaction. *Journal of Philosophical logic 38*(5), 491–526.

van Benthem, J. and F. Liu (2004). Dynamic logic of preference upgrade. Technical report, University of Amsterdam. ILLC Research Report PP-2005-29.

van Benthem, J. and F. Liu (2007). Dynamic logic of preference upgrade. *Journal of Applied Non-classical Logics 17*(2), 157–182.

van Benthem, J., J. van Eijck, and B. Kooi (2006). Logics of communication and change. *Information and Computation 204*(11), 1620–1662.

van Benthem, J. and E. Pacuit (2006). The tree of knowledge in action: Towards a common perspective. In *Advances in Modal Logic (AiML)*.

van Benthem, J. and E. Pacuit (2007). Modelling protocols in temporal models. ILLC, University of Amsterdam.

van Benthem, J. and E. Pacuit (2011). Dynamic logics of evidence-based beliefs. *Studia Logica 99(1)*, 61–92.

Blackburn, P., M. de Rijke, and Y. Venema (2001). *Modal Logic*. Cambridge University Press.

Board, O. (2002). Dynamic interactive epistemology. *Games and Economic Behaviour 49*, 49–80.

Bonanno, G. (1991). The logic of rational play in games of perfect information. *Economics and Philosophy 7(1)*, 37–65.

Bonanno, G. (1992). Rational beliefs in extensive games. In *Theory and Decision vol. 33*. Kluwer Academic Publishers.

Boutilier, C. (1994). Conditional logics of normality: A modal approach. *Artificial Intelligence 68*, 87–154.

Boutillier, C. (1993). Revision sequences and nested conditionals. In *IJCAI'93: Proceedings of the 13th International Joint conference on AI*, pp. 519–525. Chambery, France.

Burgess, J. (1981). Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic 22:1*, 1–84.

Christoff, Z. and J. Ulrik Hansen (2013). *LORI Proceedings in Lecture Notes in Computer Science*, Chapter A Two-Tiered Formalization of Social Influence, pp. 68–81. Springer.

Dégrémont, C. (2010). *The Temporal Mind: Observations on Belief Change in Temporal Systems, Dissertation*. ILLC University of Amsterdam.

Dégrémont, C. and N. Gierasimczuk (2009). Can doxastic agents learn? on the temporal structure of learning. In X. He, J. Horty, and E. Pacuit (Eds.), *Logic, Rationality, and Interaction, Second International Workshop, LORI 2009, Chongqing, China, October 8–11, 2009, Proceedings*, Volume 5834 of *LNCS*. Springer.

Dégrémont, C. and N. Gierasimczuk (2011). Finite identification from the viewpoint of epistemic update. *Information and Computation 209*(3), 383–396.

van Ditmarsch, H. and W. A. Labuschagne (2003). A multimodal language for revising defeasible beliefs. In E. Alvarez, R. Bosch, and L. Villamil (Eds.), *Proceedings of the 12th International Congress of Logic, Methodology, and Philosophy of Science*, pp. 140–141. Oviedo university Press.

van Ditmarsch, H., W. van der Hoek, and B. Kooi (2007). *Dynamic Epistemic Logic*. Synthese Library. Springer.

van Ditmarsch, H., W. van der Hoek, and B. Kooi (2005). Public announcements and belief expansion. *Advances in Modal Logic 5*, 335–346.

van Eijck, J. (2008). Yet more modal logics of preference change and belief revision. In K. R. Apt and R. van Rooij (Eds.), *New Perspectives on Games and Interaction*. Amsterdam University Press.

van Eijck, J. and F. Schwarzentruber (2014). Epistemic probability logic simplified. In R. Goré, B. Kooi, and A. Kurucz (Eds.), *Advances in Modal Logic*, pp. 158–177. College Publications.

van Eijck, J. and F. Sietsma (2008). Multi-agent belief revision with linked plausibilities. In G. Bonanno, B. Löwe, and W. van der Hoek (Eds.), *Logic and the Foundations of Game and Decision Theory (LOFT 8)*, Number 6006 in LNAI, pp. 174–189. Springer.

van Eijck, J. and Y. Wang (2008). Propositional dynamic logic as a logic of belief revision. In W. Hodges and R. de Queiroz (Eds.), *Logic, Language, Information and Computation*, Volume 5110 of *LNAI*, pp. 136–148. Springer.

Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995). *Reasoning about knowledge*, Volume 4. MIT press Cambridge, MA.

Fiutek, V. (2013). *Playing with Knowledge and Belief, PhD Dissertation.* ILLC, University of Amsterdam.

Fuhrmann, A. (1991). Theory contraction through base contraction. *Journal of Philosophical Logic 20*, 175–203.

Gärdenfors, P. (1986). Belief revisions and the ramsey test for conditionals. *Philosophical Review 95*, 813–9.

Gärdenfors, P. (1988). *Knowledge in Flux: Modelling the Dynamics of Epistemic States.* MIT Press, Cambridge MA.

Gärdenfors, P. and H. Rott (1995). Belief revision. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson (Eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*, pp. 35–132. Oxford University Press.

Gerbrandy, J. D. (1998). *Bisimulations on Planet Kripke.* Ph. D. thesis, University of Amsterdam. ILLC Dissertation Series DS-1999-01.

Gerbrandy, J. D. and W. Groeneveld (1997). Reasoning about information change. *Journal of Logic, Language, and Information 6*, 147–169.

Gierasimczuk, N. (2010). *PhD Thesis: Knowing one's limits. logical analysis of inductive inference. Knowing one's limits. logical analysis of inductive inference.* Universiteit van Amsterdam.

Gierasimczuk, N. and D. de Jongh (2013). On the complexity of conclusive update. *The Computer Journal 56*(3), 365–377.

Gierasimczuk, N., V. Hendricks, and D. de Jongh (2014). Logic and learning. In A. Baltag and S. Smets (Eds.), *Johan F. A. K. van Benthem on Logical and Informational Dynamics.* Springer.

Girard, P. (2008). *PhD Thesis: Modal Logic for Belief and Preference Change.* Stanford University and ILLC Amsterdam.

Girard, P., J. Seligman, and F. Liu (2012). General dynamic dynamic logic. In Th. Bolander, T. Braüner, S. Ghilardi, and L. S. Moss (Eds.), *Advances in Modal Logics*, pp. 239–260. College Publications.

Grove, A. J. (1988). Two modellings for theory change. *Journal of Philosophical Logic 17*, 157–170.

Halpern, J. and M. Vardi (1989). The complexity of reasoning about knowledge and time. i. lower bounds. *Journal of Computer and System Sciences 38*(1), 195 – 237.

Halpern, J. Y. (1997). Defining relative likelihood in partially-ordered preferential structure. *Journal of Artificial Intelligence Research 7*, 1–24.

Halpern, J. Y. (2003). *Reasoning about uncertainty.* MIT Press.

Hansson, S. O. (1991). Belief contraction without recovery. *Studia Logica 50*, 251–260.

Hansson, S. O. (1997). On having bad contractions, or: no room for recovery. *Journal of Applied Non-Classical Logics 7:1-2*, 241–266.

Harel, D. (1984). Dynamic logic. In D. Gabbay and F. Guenthner (Eds.), *Handbook of Philosophical Logic*, Volume II, Dordrecht, pp. 497–604. Kluwer Academic Publishers.

Harel, D., D. Kozen, and J. Tiuryn (2000). *Dynamic Logic (Foundations of Computing).* MIT Press.

Hendricks, V. F. (2001). *The Convergence of Scientific Knowledge: A View from The Limit.* Kluwer Academic Publishers.

Hendricks, V. F. (2003). Active agents. *Journal of Logic, Language and Information 12(4)*, 469–495.

Hintikka, J. (1962). *Knowledge and Belief. An Introduction to the Logic of the Two Notions.* Ithaca, NY: Cornell University Press.

Hodkinson, I. and M. Reynolds (2006). Temporal logic. In P. Blackburn, J. van Benthem, and F. Wolter (Eds.), *Handbook of Modal Logic*, pp. 655–720. Amsterdam Elsevier.

van der Hoek, W. and M. Pauly (2006). Modal logic for games and information. *Handbook of modal logic 3*, 1077–1148.

Hoshi, T. (2009). *Epistemic Dynamics and Protocol Information.* Ph. D. thesis, Stanford University.

Jeffrey, R. (1983). *The Logic of Decision (2nd ed.).* Chicago, IL: University of Chicago Press.

Katsuno, H. and A. Mendelzon (1992). On the difference between updating a knowledge base and revising it. In *Cambridge Tracts in Theoretical Computer Science*, pp. 183–203.

Kelly, K., V. Schulte, and V. Hendricks (1995). Reliable belief revision. In *Proceedings of the 10th International Congress of Logic, Methodology, and Philosophy of Science*, pp. 383–398. Kluwer Academic Publishers.

Kelly, K. T. (1998a). Iterated belief revision, reliability, and inductive amnesia. *Erkenntnis 50*(1), 11–58.

Kelly, K. T. (1998b). The learning power of belief revision. In *TARK'98: Proceedings of the 7th conference on Theoretical aspects of rationality and knowledge*, pp. 111–124. Morgan Kaufmann Publishers Inc.

Kyburg, H. E. and C. M. Teng (2001). *Uncertain Inference*. Cambrdige University Press.

Lenzen, W. (1980). *Glauben, Wissen, und Wahrscheinlichkeit*. Springer Verlag, Library of Exact Philosophy.

Levi, I. (1991). *The Fixation of Belief and Its Undoing*. Cambridge University Press.

Lewis, D. K. (1973). *Counterfactuals*. Oxford: Blackwell Publishing.

Liang, Z. and J. Seligman (2011). A logical model of the dynamics of peer pressure. *Electronic Notes in Theoretical Computer Science 278*, 275–288.

List, C. and Ph. Pettit (2004). Aggregating sets of judgements. two impossibility results compared. *Synthese 140*, 207–235.

Liu, F. (2008). *PhD Thesis: Changing for the better: Preference Dynamics and Preference Diversity*. University of Amsterdam.

Liu, F. (2011). *Reasoning about Preference Dynamics*. Synthese Library vol. 354, Springer.

Liu, F., J. Seligman, and P. Girard (2014). Logical dynamics of belief change in the community. *Synthese 191*(11), 2403–2431.

Maynard-Reid, II, P. and Y. Shoham (1998). From belief revision to belief fusion. In *Proceedings of LOFT98, Torino.*

McCarthy, J. (1980). Circumscription - a form of non-monotonic reasoning. *Artificial Intelligence 13*, 27–39.

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

Miller, J. S. and L. S. Moss (2005). The undecidability of iterated modal relativization. *Studia Logica 79*(3), 373–407.

Niederée, R. (1991). Multiple contraction: A further case against gaerdenfors' principle of recovery. In A. Fuhrmann and M. Morreau (Eds.), *The Logic of Theory Change*, pp. 322–334. Springer-Verlag.

Parikh, R. and R. Ramanujam (2003). A knowledge based semantics of messages. *Journal of Logic, Language, and Information 12*, 453–467.

Plaza, J. (1989). Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras (Eds.), *Proceedings of the Fourth International Symposium on Methodologies for Intelligent Systems: Poster Session Program*, Oak Ridge, TN, pp. 201–216. Oak Ridge National Laboratory. Reprinted in: Synthese 158, 165–179 (2007).

Ramsey, F. (1929 (1990)). General propositions and causality. In H. A. Mellor (Ed.), *F. Ramsey, Philosophical Papers*. Cambridge University Press.

Rijke, de, M. (1994). Meeting some neighbours. In J. van Eijck and A. Visser (Eds.), *Logic and information flow*, Cambridge, Massachusetts, pp. 170–195. The MIT Press.

Rott, H. (2001). *Change, Choice and Inference*. Oxford University Press.

Rott, H. (2006). Shifting priorities: Simple representations for 27 iterated theory change operators. In H. Lagerlund, S. Lindström, and R. Sliwinski (Eds.), *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*. Uppsala Philosophical Studies.

Ryan, M. and P. Y. Schobbens (1997). Counterfactuals and updates as inverse modalities. *Journal of Logic, Language and Information 6*, 123–146.

Sano, K. and S. Tojo (2013). Dynamic epistemic logic for channel-based agent communication. In K. Lodaya (Ed.), *LNAI Proceedings of ICLA 2013*, pp. 109–120. Springer-Verlag Berlin Heidelberg.

Segerberg, K. (1971). *An Essay in Classical Modal Logic*. Uppsala,Filosofiska Föreningen Och Filosofiska Institutionen Vid Uppsala Universitet.

Segerberg, K. (1991). The basic dynamic doxastic logic of AGM. In *The Goldblatt Variations*, Volume 1, pp. 76–107. Uppsala Prints and Preprints in Philosophy.

Segerberg, K. (1995). Belief revision from the point of view of doxastic logic. *Bulletin of the IGPL 3*, 535–553.

Segerberg, K. (1998). Irrevocable belief revision in Dynamic Doxastic Logic. *Notre Dame Journal of Formal Logic 39*(3), 287–306.

Segerberg, K. (1999). Two traditions in the logic of belief: bringing them together. In H. J. Ohlbach and U. Reyle (Eds.), *Logic, Language, and Reasoning*, Dordrecht, pp. 135–147. Kluwer Academic Publishers.

Segerberg, K. and H. Leitgeb (2007). Dynamic doxastic logic - why, how and where to? *Synthese 155*(2), 167–190.

Seligman, J., F. Liu, and P. Girard (2011). Logic in the community. *Lecture Notes in Computer Science 6521*, 178–188.

Seligman, J., F. Liu, and P. Girard (2013a). Facebook and the epistemic logic of friendship. In *Proceedings of the 14th TARK Conference*, pp. 229–238.

Seligman, J., F. Liu, and P. Girard (2013b). Knowledge, friendship and social announcement. In *Logic Across the University: Foundations and Applications, Proceedings of the Tsinghua Logic Conference*. College Publications.

Shoham, Y. (1988). *Reasoning About Change: Time and Change from the standpoint of Artificial Intelligence*. MIT.

Shoham, Y. and K. Leyton-Brown (2008). *Multiagent Systems: Algorithmic, Game Theoretical and Logical Foundations*. Cambridge University Press.

Spohn, W. (1988). Ordinal conditional functions: a dynamic theory of epistemic states. In W. L. Harper and B. Skyrms (Eds.), *Causation in Decision, Belief Change, and Statistics*, Volume II, pp. 105–134.

Stalnaker, R. (1998). Belief revision in games: forward and backward induction. *Mathematical Social Sciences 36*, 31–56.

Stalnaker, R. (2006). On logics of knowledge and belief. *Philosophical Studies 128*, 169–199.

Veltman, F. (1985). *Logics for Conditionals (dissertation)*. Philosophical Institute, University of Amsterdam.

Veltman, F. (1996). Defaults in update semantics. *Journal of Philosophical Logic 25*, 221–226.

Wang, Y., F. Sietsma, and J. van Eijck (2011). Logic of information flow on communication channels. In A. Omicini, S. Sardina, and W. Vasconcelos (Eds.), *Declarative Agents Languages and Technologies VIII*, pp. 130–147. Springer.