Chapter 7 SOFT INFORMATION, CORRECTION, AND BELIEF CHANGE

So far, we developed dynamic logics that deal with knowledge, inference, and questions, all based on information and truth. Now we want to look at another pervasive attitude that agents have toward information, namely, their *beliefs*. This chapter will show how belief change fits well with our dynamic framework, and develop some of its logical theory. [131] This puts one more crucial aspect of rational agents in place: not their being right about everything, but their being wrong, and subsequent acts of self-correction.

7.1 From knowledge to belief as a trigger for actions

While knowledge is important to agency, our actions are often driven by fallible beliefs. I am riding my bicycle this evening because I believe it will get me home, even though my epistemic range includes worlds where the San Andreas Earthquake strikes. Decision theory is about choice and action on the basis of beliefs, as knowledge may not be available. Thus, our next step in the logical dynamics of rational agency is the study of beliefs, viewed as concretely as possible. Think of our scenarios so far. The cards have been dealt. I know that there are 52 of them, and I know their colors. But I have more fleeting beliefs about who holds which card, or about how the other agents will play. [132]

Hard versus soft information With this distinction in attitude comes a richer dynamics. A public announcement *!P* of a fact *P* was an event of *hard information* that changes irrevocably what I know. When I see the Ace of Spades played, I come to know that no one has it any more. This is the trigger that drove our dynamic epistemic logics in Chapters 3 and 4. Such events of hard information may also change our beliefs – and we will find a complete logical system for this. But there are also events of *soft information*, affecting my beliefs without affecting my knowledge about the cards. I see you smile. This makes it more likely that you hold a trump card, but it does not rule out that you do not. To describe this, we will use worlds with plausibility orderings supporting dynamic updates.

The tandem of jumping ahead and self-correction Here is what is most important to me in this chapter from the standpoint of rational agency. As acting agents, we are bound to form beliefs that go beyond the hard information we have. And this is not a concession to

---

[131] Later on, we discuss how this relates to the alternative *AGM* style (Gärdenfors & Rott 1995).

[132] Of course, I could even be wrong about the cards (perhaps the Devil added his visiting card) – but this worry seems morbid, and not useful in investigating normal information flow.

human frailty or to our mercurial nature. It is rather the essence of creativity, jumping ahead to conclusions we are not really entitled to, and basing our beliefs and actions on them. But there is another side to this coin, that I would dub our capacity for *self-correction*, or if you wish, for *learning*. We have an amazing capacity for standing up after we have fallen informationally, and to me, rationality is displayed at its best in intelligent responses to new evidence that contradicts what we thought so far. What new beliefs do we form, and what amended actions result? Chapter 1 saw this as a necessary *pair of skills*: jumping to conclusions (i.e., beliefs) and correcting ourselves in times of trouble. And the hallmark of a rational agent is to be good at both: it is easy to prove one theorem after another, it is hard to revise your theory when it has come crashing down. So, in pursuing the dynamic logics of this chapter, I am trying to chart this second skill.

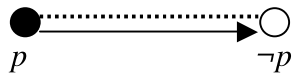
7.2 Static logic of knowledge and belief

Knowledge and belief have been studied together ever since Plato proposed his equation of knowledge with 'justified true belief', and much of epistemology is still about finding an ingredient that would turn true belief into knowledge. Without attempting this here (see Chapter 13 for our thoughts), how can we put knowledge and belief side by side?

Reinterpreting PAL One easy route reinterprets dynamic-epistemic logic so far. We read the earlier *K*-operators as beliefs, again as universal quantifiers over the accessible range, placing no constraints on the accessibility relation: just pointed arrows. One test for such an approach is that it must be possible for beliefs to be *wrong*:

Example A mistaken belief.

Consider the following model with two worlds that are epistemically accessible to each other, but the pointed arrow is the only belief relation. Here, in the actual black world to the left, the proposition p is true, but the agent mistakenly believes that $\neg p$:



■

With this view of doxastic modalities (cf. Hintikka 1962), the machinery of *DEL* works exactly as before. But there is a problem:

Example, continued

Consider a public 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. ■

In this way, agents who have their beliefs contradicted are shattered and start believing anything. Such a collapse is unworthy of a rational agent in the sense of Chapter 1, and hence we will change the semantics to allow for more intelligent responses.

World comparison by plausibility A richer view of belief follows the intuition that an agent believes the things that are true, not in all her epistemically accessible worlds, but only in those that are 'best' or most relevant to her. I believe that my bicycle will get me home on time, even though I do not know that it will not suddenly disappear in a seismic chasm. But the worlds where it stays on the road are more plausible than those where it drops down, and among the former, those where it arrives on time are more plausible than those where it does not. Static models for this setting are easily defined:

Definition Epistemic-doxastic models.

Epistemic-doxastic models are structures $\boldsymbol{M} = (W, \{\sim_i\}_{i\in I}, \{\leq_{i,\,s}\}_{i\in I}, V)$ where the relations $\sim_i$ stand for epistemic accessibility, and the $\leq_{i,\,s}$ are ternary comparison relations for agents read as follows, $x \leq_{i,\,s} y$ if, in world s, agent i considers y at least as plausible as x. ■

Now epistemic accessibility can be an equivalence relation again. Models like this occur in conditional logic, Shoham 1988 on preference relations in AI, and the 'graded models' of Spohn 1988. One can impose several conditions on the plausibility relations, depending on their intuitive reading. Burgess 1981 has *reflexivity* and *transitivity*, Lewis 1973 also imposes *connectedness*: for all worlds $s, t,$ either $s \leq t$ or $t \leq s$. The latter yields the well-known geometrical nested spheres for conditional logic. [133] As with epistemic models, our dynamic analysis works largely independently from such design decisions, important though they may be. In particular, connected orders yield nice pictures of a line of equi-plausibility clusters, in which there are only three options for worlds s, t:

either *strict precedence* $s < t$ *or* $t < s$, or *equiplausibility* $s \leq t \wedge t \leq s$.

But there are also settings that need a fourth option of *incomparability:* $\neg\, s \leq t \wedge \neg\, t \leq s$. This happens when comparing worlds with conflicting criteria, as with some preference

---

[133] The natural *strict variant* of these orderings is defined as follows: $s < t$ iff $s \leq t \wedge \neg\, t \leq s$.

logics in Chapter 9. Sometimes also, partially ordered graphs are just the mathematically more elegant approach (Andréka, Ryan, Schobbens 2002; cf. Chapter 12).

Languages and logics One can interpret many logical operators in this richer structure. In what follows, we choose intuitive maximality formulations for belief $B_i\varphi$. [134] First of all, there is plain belief, whose modality is interpreted as follows: [135]

Definition Belief as truth in the most plausible worlds.

In epistemic-doxastic models, knowledge is interpreted as usual, while we put $\boldsymbol{M}, s \models B_i\varphi$ iff $\boldsymbol{M}, t \models \varphi$ for all worlds t that are maximal in the ordering $\lambda xy.\, x \leq_{i,s} y$. [136] ■

But absolute belief does not suffice. Reasoning about information flow and action involves *conditional belief*. We write this as follows: $B^\psi\varphi$, with the intuitive reading that, conditional on ψ, the agent believes that φ. This is close to conditional logic:

Definition Conditional beliefs as plausibility conditionals.

In epistemic-doxastic models, $\boldsymbol{M}, s \models B^\psi\varphi$ iff $\boldsymbol{M}, t \models \varphi$ for all worlds t that are maximal for the ordering $\lambda xy.\, x \leq_{i,s} y$ in the set $\{u \mid \boldsymbol{M}, u \models \psi\}$. ■

Absolute belief $B\varphi$ is the special case $B^T\varphi$. It can be shown that conditional belief is not definable in terms of absolute belief, so we have a genuine language extension. [137] [138]

Digression on conditionals As with epistemic notions in Chapters 2, 3, conditional beliefs *pre-encode* beliefs that we would have if we were to learn certain things. [139] A formal

---

[134] These must be modified in non-wellfounded models allowing *infinite descent* in the ordering. This issue is orthogonal to the main thrust of this chapter, and we will largely ignore it.

[135] For convenience, henceforth we drop subscripts where they do not add insight.

[136] Here we used lambda notation to denote relations, but plain '$x \leq_{i,s} y$' would serve, too.

[137] Likewise, the binary quantifier *Most A are B* is not definable in first-order logic extended with just a unary quantifier "Most objects in the universe are *B*" (cf. Peters & Westerståhl 2005).

[138] One can also interpret richer languages on epistemic-doxastic models. E.g., maximality suggests a binary relation *best* defined as "t is maximal in $\lambda xy.\leq_s xy$". One can introduce a modality for this, defining conditional belief as $[best\ \psi]\varphi$. Dynamic extensions of our language will come below.

[139] *Static pre-encoding versus dynamics*. A conditional belief $B^\psi\varphi$ does not quite say that we would believe φ if we learnt that ψ. For an act of learning ψ *changes the current model* $\boldsymbol{M}$, and hence the truth value of φ might change, as modalities in φ now range over fewer worlds in $\boldsymbol{M}/\psi$. Similar things happened with epistemic statements after communication in Chapter 3 – and also in logic in

analogy is this. A conditional $\psi \Rightarrow \varphi$ says that φ is true in the closest worlds where ψ is true, along some comparison order on worlds. This is exactly the above clause. Thus, results from conditional logic apply. For instance, on reflexive transitive plausibility models, we have this completeness theorem (Burgess 1981, Veltman 1985):

Theorem The logic of $B^{\psi}\varphi$ is axiomatized 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* $(\varphi \wedge \psi) \Rightarrow \chi$.

Epistemic-doxastic logics In line with the general approach in this book, we do not pursue completeness theorems for static logics of knowledge and belief. But for greater ease, this chapter makes one simplification that reflects in the logic. Epistemic accessibility will be an *equivalence relation*, and plausibility a *pre-order over the equivalence classes*, the same as viewed from any world inside the class. This makes the following axiom valid:

$B\varphi \rightarrow KB\varphi$ *Epistemic-Doxastic Introspection*

While this is a debatable assumption, it helps focus on the core ideas of the dynamics.

7.3 Belief change under hard information

Our first dynamic logic of belief revision puts together the logic *PAL* with our static models for conditional belief, following the same methodology as earlier chapters. We will move fast, as the general points of Chapters 2, 3 apply, and indeed thrive here.

A complete axiomatic system For a start, we must locate the key recursion axiom for the new beliefs, something that can be done easily, using update pictures as before:

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

$[!P]B\varphi \leftrightarrow (P \rightarrow B^{P}([!P]\varphi)$.

This is like the *PAL* recursion axiom for knowledge under announcement. But note the conditional belief in the consequent, that does not reduce to a conditional absolute belief $B(P \rightarrow$ ... Still, to keep the language in harmony, this is not enough. We need to know, not

general. The relativized quantifier in "All mothers have daughters" does not say that, if we relativize to the subset of mothers, all of them have daughters who are mothers themselves.

just which beliefs are formed after new information, but which conditional beliefs. [140] What is the recursion principle for change in conditional beliefs under hard information? There might be a regress here toward conditional conditional beliefs, but in fact, we have:

Theorem The logic of conditional belief under public announcements is axiomatized 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:

$[!P]B^{\psi}\varphi \leftrightarrow (P \rightarrow B^{P \wedge [!P]\psi}[!P]\varphi)$.

Proof First we check the soundness of the new axiom. On the left hand side, it says that in the new model $(\boldsymbol{M}|P, s)$, φ is true in the best ψ-worlds. With the usual precondition for true announcement, on the right-hand side, it says that in $\boldsymbol{M}, s$, the best worlds that are P now and will become ψ after announcing that P, will also become φ after announcing P. This is indeed equivalent. The remainder of the proof is our earlier stepwise reduction analysis, noting that the above axiom is recursive, pushing announcement modalities inside. ■

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

Clarifying the Ramsey Test Our dynamic logic sharpens up the *Ramsey Test* that says: "A conditional proposition $A \Rightarrow B$ is true, if, after adding A to your current stock of beliefs, the minimal consistent revision implies B." In our perspective, this is ambiguous, as B need no longer say the same thing after the revision. That is why our recursion axiom carefully distinguishes between formulas φ before update and what happens to them after: $[!P]\varphi$. Even so, there is an interesting special case of *factual propositions* φ without modal operators (cf. Chapter 3), that do not change their truth value under announcement. In that case, with Q, R factual propositions, the above recursion axioms read as follows:

$[!P]BQ \leftrightarrow (P \rightarrow B^{P}Q)$, $[!P]B^{R}Q \leftrightarrow (P \rightarrow B^{P \wedge R}Q)$ [141] ■

---

[140] This is overlooked in classical belief revision theory, which says only how new absolute beliefs are formed. One gets stuck in one round, as the new state does not pre-encode what happens in the next round. This so-called *Iteration Problem* cannot arise in a systematic logical set-up.

[141] For some paradoxes in combining the Ramsey test with belief revision, cf. Gärdenfors 1988. The nice thing of a logic approach is that every law we formulate is automatically sound.

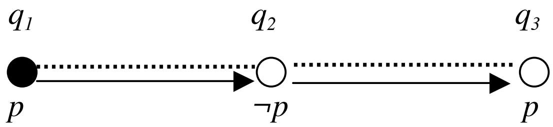
Belief change under hard update is not yet revision in the usual sense, that can be triggered by weaker information (see below). Nevertheless, we pursue it a bit further, as it links to important themes in rational agency: variety of attitudes, and of consequence relations. The first will be considered right now, the second a bit later in this chapter.

7.4 Exploring the framework: safe belief and richer attitudes

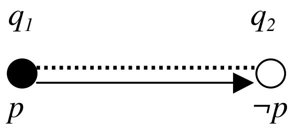
The above setting may seem simple, but it contains some tricky scenarios:

Example Misleading with the truth.

Consider a model where an agent believes that p, which is indeed true in the actual world to the far left, but for the wrong reason: she finds the most plausible world the one to the far right. For convenience, assume each world verifies a unique proposition letter q_i:



Now giving the true information that we are not in the final world $(`\neg q_3')$ updates to



in which the agent believes mistakenly that $\neg p$.[142] ■

Agents have a rich repertoire of attitudes In response, an alternative view of our task in this chapter makes sense. So far, we assumed that knowledge and belief are the only relevant attitudes. But in reality, agents have a rich repertoire of attitudes concerning information and action, witness the many terms in natural language with an epistemic or doxastic ring: being certain, being convinced, assuming, etcetera.[143]

Language extension: safe belief Among all possible options in this plethora of epistemic-doxastic attitudes, the following new notion makes particular sense, intermediate between knowledge and belief. It has stability under new true information:

---

[142] Observations like this have been made in philosophy, computer science, and game theory.

[143] Cf. Lenzen 1980 for similar views. Krista Lawlor has pointed me also at the richer repertoire of epistemic attitudes found in pre-modern epistemology.

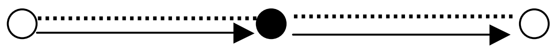
Definition Safe belief.

The modality of *safe belief* $B^+\varphi$ is defined as follows: $\boldsymbol{M}, s \models B^+\varphi$ iff for all worlds t in the epistemic range of s with $t \geq s$, $\boldsymbol{M}, t \models \varphi$. In words, φ is true in all epistemically accessible worlds that are at least as plausible as the current one. [144] ■

The modality $B^+\varphi$ is stable under hard information, at least for factual assertions φ that do not change their truth value as the model changes. [145] And it makes a lot of technical sense, as it is the *universal base modality* $[\leq]\varphi$ for the plausibility ordering. This idea occurs in Boutilier 1994, Halpern 1997 (cf. Shoham & Leyton-Brown 2008), Baltag & Smets 2006, 2007 (following Stalnaker), and independently in our Chapter 9 on preference logic. In what follows, we make safe belief part of the static doxastic language – as a pilot for a richer theory of attitudes in the background. Pictorially, one can think of this as follows:

Example Three degrees of doxastic strength.

Consider this picture, now with the actual world in the middle:



$K\varphi$ describes what we know: φ must be true in all worlds in the epistemic range, less or more plausible than the current one. $B^+\varphi$ describes our safe beliefs in further investigation: φ is true in all worlds from the middle toward the right. Finally, $B\varphi$ describes the most fragile thing: our beliefs as true in all worlds in the current rightmost position. ■

In addition, safe belief simplifies things, if only as a technical device:

Fact The following assertions hold on finite epistemic connected plausibility models:

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

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

Proof (a) is obvious, since we can conditionalize to $B^+(A \to \varphi)$ like a standard modality. (b) uses a fact about finite connected plausibility models, involving the existential dual modality $\langle B^+\rangle$ of safe belief (cf. van Benthem & Liu 2007, Baltag & Smets 2006):

Claim Conditional belief $B^\psi\varphi$ is equivalent to the iterated modal statement

$$K((\psi \wedge \varphi) \to \langle B^+\rangle(\psi \wedge \varphi \wedge B^+(\psi \to \varphi))).$$

---

[144] Safe belief in this style uses an *intersection* of epistemic accessibility and plausibility. We could also decouple the two, and introduce a modality for plausibility alone.

This claim is not for the faint-hearted, but it can be proved with a little puzzling. [146] ■

Safe belief also has some less obvious features. For instance, since its accessibility relation is transitive, it satisfies Positive Introspection, but since that relation is not Euclidean, it fails to satisfy Negative Introspection. The reason is that safe belief mixes purely epistemic information with *procedural information* (cf. Chapters 3, 11). Once we see that agents have a richer repertoire of doxastic-epistemic attitudes than K and B, old intuitions about epistemic axioms need not be very helpful in understanding the full picture.

Finally, we turn to dynamics under hard information, i.e., our key recursion axiom:

Theorem The complete logic of belief change under hard information is the one whose principles were stated before, plus the following recursion axiom for safe belief:

$[!P]\, B^{+}\varphi \leftrightarrow (P \rightarrow B^{+}(P \rightarrow [!P]\varphi))$.

This axiom for safe belief under hard information implies the earlier one for conditional belief, by unpacking the above modal definition.

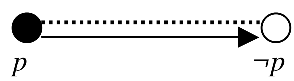
7.5 Belief change under soft information: radical upgrade

It is time to move to a much more general and flexible view of our subject.

Soft information and plausibility change Our story so far is a hybrid: we saw how a soft attitude changes under hard information. The more general scenario has an agent aware of being subject to continuous belief changes, and taking incoming signals in a softer manner, without throwing away options forever. But then, public announcement is too strong:

Example No way back.

Consider the earlier model where the agent believed that $\neg p$, though p was in fact the case:



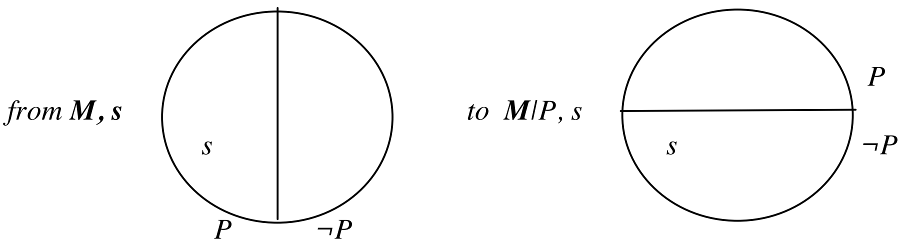
Publicly announcing p removes the $\neg p$-world, making later belief revision impossible. ■

[145] Note that new true information will never remove the actual world, our vantage point.

[146] The result generalizes to other models, and this modal translation is *itself* a good candidate for lifting the maximality account of conditional belief to infinite models, as well as non-connected ones. Alternative versions would use modalities for the strict ordering corresponding to reflexive plausibility $\leq$ to define maximal ψ-worlds directly in the format $\psi \wedge \neg <<>\psi$: cf. Girard 2008.

What we need is a mechanism that just makes incoming information P more plausible, without burning our ships behind us. An example are the *default rules* $A \Rightarrow B$ in Veltman 1996. Accepting a default conditional does not say that all A-worlds must now be B-worlds. It makes the counter-examples (the $A \wedge \neg B$-worlds) less plausible until further notice. This soft information does not eliminate worlds, it just changes their ordering. More precisely, a triggering event that makes us believe that P need only *rearrange worlds* making the most plausible ones P: by 'promotion' or 'upgrade' rather than elimination of worlds. Thus, in our models $\boldsymbol{M} = (W, \sim_i, \leq_i, V)$, we change the relations $\leq_i$, rather than the world domain W or the epistemic accessibilities $\sim_i$. Rules for plausibility change exist in models of belief revision (Grove 1988, Rott 2006) as different *policies* that agents can adopt toward new information. We now show how our dynamic logics deal with them. [147]

Radical revision One very strong policy is like a radical social revolution where some underclass P now becomes the upper class. In a picture, we get this reversal:



Definition Radical, or lexicographic upgrade.

A *lexicographic upgrade* $\Uparrow P$ changes the current ordering $\leq$ between worlds in $\boldsymbol{M}, s$ to a new model $\boldsymbol{M}\Uparrow P, s$ as follows: all P-worlds in the current model become better than all $\neg P$-worlds, while, within those two zones, the old plausibility ordering remains. [148] ■

With this definition in place, our earlier methodology applies. As for public announcement, we introduce a corresponding upgrade modality in our dynamic doxastic language:

$$\boldsymbol{M}, s \models [\Uparrow P]\varphi \quad \text{iff} \quad \boldsymbol{M}\Uparrow P, s \models \varphi$$

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 The dynamic logic of lexicographic upgrade is axiomatized completely by

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

---

[147] Alternatively, in formal learning theory (Kelly 1996), these are different learning strategies.

[148] This is known as the 'lexicographic policy' for relational belief revision.

(b) the following recursion axioms:

$[\Uparrow P]q$	$\leftrightarrow$	q	for all atomic proposition letters q
$[\Uparrow P]\neg\varphi$	$\leftrightarrow$	$\neg[\Uparrow P]\varphi$	
$[\Uparrow P](\varphi\wedge\psi)$	$\leftrightarrow$	$[\Uparrow P]\varphi \wedge [\Uparrow P]\psi$	
$[\Uparrow P]K\varphi$	$\leftrightarrow$	$K[\Uparrow P]\varphi$	
$[\Uparrow P]B^{\psi}\varphi$	$\leftrightarrow$	$(\Diamond(P \wedge [\Uparrow P]\psi) \wedge B^{P \wedge [\Uparrow P]\psi} [\Uparrow P]\varphi)$	

$$\vee\ (\neg\Diamond(P \wedge [\Uparrow P]\psi) \wedge B^{[\Uparrow P]\psi} [\Uparrow P]\varphi)$$ [149]

Proof The first four axioms are simpler than those for *PAL*, since there is no precondition for $\Uparrow P$ as there was for $!P$. The first axiom says that upgrade does not change truth values of atomic facts. The second says that upgrade is a function on models, the third is a general law of modality, and the fourth that no change takes place in epistemic accessibility.

The fifth axiom is the locus where we see the specific change in the plausibility ordering. The left-hand side says that, after the *P*-upgrade, all best ψ-worlds satisfy φ. On the right-hand side, there is a case distinction. Case (1): there are accessible *P*-worlds in the original model $\boldsymbol{M}$ that become ψ after the upgrade. Lexicographic reordering $\Uparrow P$ makes the best of these worlds the best ones over-all in $\boldsymbol{M}\Uparrow P$ to satisfy ψ. In the original $\boldsymbol{M}$ – viz. its epistemic component visible from the current world s – the worlds of Case 1 are just those satisfying the formula $P \wedge [\Uparrow P]\psi$. Therefore, the formula $B^{P \wedge [\Uparrow P]\psi} [\Uparrow P]\varphi$ says that the best among these in $\boldsymbol{M}$ will indeed satisfy φ after the upgrade. These best worlds are the same as those described earlier, as lexicographic reordering does not change order of worlds inside the *P*-area. Case (2): no *P*-worlds in the original $\boldsymbol{M}$ become ψ after upgrade. Then lexicographic reordering $\Uparrow P$ makes the best worlds satisfying ψ after the upgrade just the same best worlds over-all as before that satisfied $[\Uparrow P]\psi$. Here, the formula $B^{[\Uparrow P]\psi} [\Uparrow P]\varphi$ in the reduction axiom says that the best worlds become φ after upgrade.

The rest of the proof is the reduction argument of Chapter 3. [150] ■

The final equivalence describes which conditional beliefs agents have after soft upgrade. This may look daunting, but try to read the principles of some default logics existing

---

[149] Here, as in Chapter 2, $\Diamond$ is the dual *existential epistemic modality* $\neg K\neg$.

[150] Details on this result and the next are in van Benthem 2007B, van Benthem & Liu 2007.

today! Also, recall the earlier point that we need to describe how conditional beliefs change, not just absolute ones, to avoid getting trapped in the Iteration Problem.

Special cases Looking at special cases may help. First, consider absolute beliefs $B\varphi$. Conditioning on 'True', the key recursion axiom simplifies to:

$$([\Uparrow P]B\varphi \;\leftrightarrow\; (\Diamond P \wedge B^{P}[\Uparrow P]\varphi) \vee (\neg\Diamond P \wedge B[\Uparrow P]\varphi)$$

And here is the simplified recursion axiom for *factual propositions* that did not change their truth values under update or upgrade:

$$[\Uparrow P]B^{\psi}\varphi \;\leftrightarrow\; (\Diamond(P\wedge\psi) \wedge B^{P\wedge\psi}\varphi) \;\vee (\neg\Diamond(P\wedge\psi) \wedge B^{\psi}\varphi)$$ [151] ■

Safe belief once more As a final simplification, recall the earlier notion of safe belief, that defined conditional belief using K. We can also derive the above from the following:

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

$$[\Uparrow P]\, B^{+}\varphi \;\leftrightarrow\; (P \wedge B^{+}(P \rightarrow [\Uparrow P]\varphi)) \vee (\neg P \wedge B^{+}(\neg P \rightarrow [\Uparrow P]\varphi) \wedge K(P \rightarrow [\Uparrow P]\varphi)).$$

Proof For any world s, the two disjuncts describe its more plausible worlds after upgrade. If $\boldsymbol{M}, s \models P$ in the initial model $\boldsymbol{M}$, these are all former P-worlds that were more plausible than s. If not $\boldsymbol{M}, s \models P$, these are the old more plausible $\neg P$-worlds plus *all* P-worlds. ■

Static pre-encoding Our compositional analysis says that any statement about effects of hard or soft information is encoded in the initial model: the epistemic present contains the epistemic future. We have used this line to design the right static languages, with a crucial role for conditional belief. As in earlier Chapters, we may want to drop this reduction when considering global informational procedures: Chapter 11 shows how to do this. [152]

Radical upgrade will be used at various places in this book, especially in our study of game solution in Chapters 10, 15. This will throw further light on its semantic features.

7.6 Conservative upgrade and general revision policies

Radical revision was our pilot, but its specific plausibility change is just one way of taking soft information. A more conservative policy for believing a new proposition puts not all P-worlds on top qua plausibility, but just *the most plausible P-worlds*. After the revolution,

---

[151] To us, this is the paradox-free sense in which a Ramsey Test holds for our logic.

[152] Technically, this design involves a form of closure beyond syntactic relativization (Chapter 3). We now also need closure under syntactic *substitutions* of defined predicates for old ones.

this policy co-opts just the leaders of the underclass – the sage advice that Macchiavelli gave to rulers pondering what to do with the mob outside of their palace.

Definition Conservative plausibility change.
Conservative upgrade $\uparrow P$ replaces the current ordering $\leq$ in a model $\boldsymbol{M}$ by the following: the best P-worlds come on top, but apart from that, the old ordering remains. ■

Technically, $\uparrow P$ is a special case of radical revision: $\Uparrow(best(P))$, if we have the latter in our static language. But it seems of interest per se. Our earlier methods produce its logic:

Theorem The dynamic logic of conservative upgrade is axiomatized completely by

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

(b) the following reduction axioms:

$[\uparrow P]q$	$\leftrightarrow$	q	for all atomic proposition letters q
$[\uparrow P]\neg\varphi$	$\leftrightarrow$	$\neg[\uparrow P]\varphi$	
$[\uparrow P](\varphi\wedge\psi)$	$\leftrightarrow$	$[\uparrow P]\varphi \wedge [\uparrow P]\psi$	
$[\uparrow P]K\varphi$	$\leftrightarrow$	$K[\uparrow P]\varphi$	
$[\uparrow P]B^{\psi}\varphi$	$\leftrightarrow$	$(B^{P}\neg[\uparrow P]\psi \wedge B^{[\uparrow P]\psi}[\uparrow P]\varphi) \vee$	
		$(\neg B^{P}\neg[\uparrow P]\psi \wedge B^{P\wedge[\uparrow P]\psi}[\uparrow P]\varphi)$	

We leave a proof to the reader. Of course, one can also combine this logic with the earlier one, to combine different sorts of revising behaviour, as in mixed formulas $[\Uparrow][\uparrow]\varphi$.

Policies Many further changes in a plausibility order can be responses to an incoming signal. This reflects the host of belief revision policies in the literature: Rott 2006 has 27. General relation transformers were proposed in van Benthem, van Eijck & Frolova 1993, calling for a dynamification of preference logic. The same is true for defaults, commands (Yamada 2006), and other areas where plausibility or preference can change (cf. Chapter 9). Our approach suggests that one can take any definition of change, write a matching recursion axiom, and then a complete dynamic logic. But how far does this go? [153]

---

[153] Maybe 'policy' is the wrong term, as it suggests a persistent habit over time, like being stubborn. But our events describe local responses to particular inputs. 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.

Relation transformers in dynamic logic One general method works by inspection of the format of definition in the above examples. For instance, it is easy to see the following:

Fact Radical upgrade $\Uparrow P$ is definable as a program in propositional dynamic logic.

Proof The format is as follows, with 'T' the universal relation between all worlds:

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

Van Benthem & Liu 2007 then introduce the following format:

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

A further example is an act of 'suggestion' $\#P$ (cf. the preference logics of Chapter 9) that merely takes out R-pairs with '$\neg P$ over P':

$$\#P(R) = (?P;\, R) \cup (R;\, ?\neg P)$$

This format generalizes our earlier procedure with recursion axioms considerably:

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

Proof Here are two examples of computing modalities for the new relation after the model change, using the recursive program axioms of *PDL*. Note how the second calculation uses the existential epistemic modality $\Diamond$ for the occurrence of the universal relation:

(a) $\langle \#P(R)\rangle\langle R\rangle\varphi \leftrightarrow \langle (?P;\, R) \cup (R;\, ?\neg P)\rangle\varphi \leftrightarrow \langle (?P;\, R)\rangle\varphi \vee \langle (R;\, ?\neg P)\rangle\varphi$
$\leftrightarrow \langle ?P\rangle\langle R\rangle\varphi \vee \langle R\rangle\langle ?\neg P\rangle\varphi \leftrightarrow (P \wedge \langle R\rangle\varphi) \vee \langle R\rangle(\neg P \wedge \varphi).$

(b) $\langle \Uparrow P(R)\rangle\varphi \leftrightarrow \langle (?P;\, T\, ;\, ?\neg P) \cup (?P\, ;\, R;\, ?P) \cup (?\neg P\, ;\, R;\, ?\neg P)\rangle\, \varphi$
$\leftrightarrow \langle (?P;\, T\, ;\, ?\neg P)\rangle\varphi \vee \langle (?P\, ;\, R;\, ?P)\rangle\varphi \vee \langle (?\neg P\, ;\, R;\, ?\neg P)\rangle\varphi$
$\leftrightarrow \langle ?P\rangle\langle T\rangle\langle ?\neg P\rangle\varphi \vee \langle ?P\rangle\langle R\rangle\langle ?P\rangle\varphi \vee \langle ?\neg P\rangle\langle R\rangle\langle ?\neg P\rangle\varphi$
$\leftrightarrow (P \wedge E(\neg P \wedge \varphi)) \vee (P \wedge \langle R\rangle(P \wedge \varphi)) \vee (\neg P \wedge \langle R\rangle(\neg P \wedge \varphi)).$ [154]

This gives uniformity behind earlier cases. For instance, the latter easily transforms into an axiom for safe belief after radical upgrade $\Uparrow P$, equivalent to the one we gave before. ■

---

[154] Here, 'E' stands for the *existential modality* 'in some world'.

7.7 Conclusion

This chapter has realized the second stage of our logical analysis of agency, extending the dynamic approach for knowledge to belief. The result is one merged theory of information update and belief revision, using standard modal techniques instead of ad-hoc formalisms. We found many new topics, like using the dynamics to suggest new epistemic modalities.

After this opt-out point for the chapter, we will pursue some more themes for the interested reader, including transfer of insights between *DEL* and frameworks such as *AGM*.

7.8 Further themes: belief revision with *DEL*

DEL formats, event models as triggers Another approach uses *DEL* (Chapter 4) rather than *PAL* as a role model. Event models for information can be much more subtle than announcements, or the few specific policies we have discussed. While its motivation came from partial observation, *DEL* also applies to receiving signals with different strengths. Here is a powerful idea from Baltag & Smets 2006 (with a precursor in Aucher 2004):

Definition Plausibility event models.
Plausibility event models are event models just as in Chapter 4, but now expanded with an additional plausibility relation over their epistemic equivalence classes. ■

In this setting, radical upgrade $\Uparrow P$ can be implemented in an event model as follows: we do not throw away worlds, so we need two 'signals' $!P$ and $!\neg P$ with obvious preconditions P, $\neg P$ that will copy the old model. But we now say that signal $!P$ is more plausible than signal $!\neg P$, relocating the revision policy in the nature of the input:

$$!P \quad \geq \quad !\neg P$$

Different event models will represent a great variety of update rules. But we still need to state the update mechanism more precisely, since it is not quite that of Chapter 4:

'One Rule To Rule Them All' The product update rule radically places the emphasis on the last event observed, but it is conservative with respect to everything else:

Definition Priority Update.
Consider an epistemic plausibility model ***M***, *s* and a plausibility event model ***E***, *e*. The *product model* ***M x E***, *(s, e)* is defined entirely as in Chapter 4 – with the addition of a new

update rule for the plausibility relation, where < is the strict version of the relation:

$(s, e) \leq (t, f)$ iff $(s \leq t \wedge e \leq f) \vee e < f$. ■

Thus, if the new *P* induces a preference between worlds, that takes precedence: otherwise, we go by the old plausibility order. This rule places great weight on the last observation or signal received. This is like belief revision theory, where receiving just one signal *$*P$* leads me to believe that *P*, even if all of my life, I had been receiving evidence against *P*. It is also in line with 'Jeffrey Update' in probability (Chapter 8) that imposes a new probability for some proposition, while adjusting all other probabilities proportionally. [155] [156]

Theorem The dynamic logic of priority update is axiomatizable completely.

Proof As before, it suffices to state the crucial recursion axioms reflecting the above rule. We display just one case, for the relation of safe belief, in existential format:

$$<\boldsymbol{E}, e><\leq>\varphi \leftrightarrow (PRE_e \wedge (\bigvee_{e \leq f \, in \, \boldsymbol{E}} <\leq><\boldsymbol{E}, f>\varphi \vee (\bigvee_{e < f \, in \, \boldsymbol{E}} \Diamond<\boldsymbol{E}, f>\varphi))$$

where $\Diamond$ is again the existential epistemic modality. ■

This *shifts the locus of description*. Instead of many policies for processing a signal, each with its own logic, we now put the policy in the input ***E***. This has some artificial features: the new event models are much more abstract than those in Chapter 4. Also, even to describe simple policies like conservative upgrade, the language of event models must be extended to event preconditions of the form *most-plausible(P)*. But the benefit is clear: infinitely many policies can be encoded in event models, while belief change now 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. [157]

Digression: abrupt revision and slow learning An update rule with so much emphasis on the last signal is special. Chapter 12 brings out how, using social choice between old and new signals. Learning theory also has gentler ways of merging new with old information,

---

[155] There may be a worry that this shifts from *DEL's precondition* analysis to a forward style of thinking in terms of *postconditions*: cf. Chapter 3, but we will not pursue this possible objection.

[156] As in Chapter 4, product update with event models generalizes easily to real world change, taking on board the well-known Katsuno-Mendelzon sense of temporal 'update'.

[157] A comparison between the earlier *PDL*-style and *DEL*-style formats remains to be made.

without overriding past experience. This theme will return with probabilistic update rules in Chapter 8, and with score-based rules for preference dynamics in Chapter 9.

7.9 Further themes: belief revision versus generalized consequence

Update of beliefs under hard or soft information is also an alternative to the current world of nonstandard notions of consequence. Here is a brief illustration (van Benthem 2008D): Chapter 13 has a more extensive discussion of the issues.

Update versus inference: non-monotonic logic Classical consequence says that all models of premises $\boldsymbol{P}$ are models for the conclusion C. McCarthy 1980 pointed out how problem solving goes beyond this. A *circumscriptive* consequence from $\boldsymbol{P}$ to C says that

C is true in all the *minimal* models for $\boldsymbol{P}$

Here, minimality refers to a relevant comparison order $\leq$ for models: inclusion of object domains, extensions of predicates, and so on. The general idea is minimization over any order (Shoham 1988), supporting non-monotonic consequence relations that are like the earlier conditional logics. This is reminiscent of our plausibility models for belief, and indeed, one can question the original view of problem solving. We are given initial information and need to find the goal situation, as new information comes in. The crucial process here are our responses: solving puzzles and playing games is all about information update and belief change. Non-monotonic logics have such processes in the background, but leave them *implicit*. But making them *explicit* is the point of our dynamic logics.

Dynamic consequence on a classical base Our logic suggests two kinds of dynamic consequence. First, *(common) knowledge* may result, and we get classical consequence for factual assertions (cf. the dynamic inference of Chapter 3). [158] Or *belief* may result, referring to the minimal worlds. Thus, what is usually cast as a notion of consequence

$$P_1, \ldots, P_k \Rightarrow \varphi$$

gets several dynamic variants definable in our language:

either $[!P_1] \ldots [!P_k]\, K\varphi$ or $[!P_1] \ldots [!P_k]\, B\varphi$

whose behaviour is captured by our earlier complete logics. This suggests a truly radical point of view. Once the relevant informational events have been made explicit, there is no

[158] Factual assertions seem to drive accounts of nonstandard consequence relations. But as we saw in Chapter 3, structural rules get dynamic twists when we consider the full language.

need for 'non-standard logics'. The dynamic logic just works with a classical static notion of consequence. *Non-monotonic logic is monotonic dynamic logic of belief change*.

New styles Our logics for soft information even suggest *new* consequence relations:

$$P_1, \ldots, P_k \Rightarrow_{circ\text{-}soft} \varphi \quad \text{iff} \quad [\Uparrow P_1] \ldots [\Uparrow P_k] B\varphi$$

Fact For factual assertions P, Q, (i) $P, Q \Rightarrow_{circ\text{-}hard} P$, (ii) not $P, Q \Rightarrow_{circ\text{-}soft} P$.

Proof (i) Successive hard updates yield subsets of the P-worlds. (ii) The last upgrade with Q may have demoted all P-worlds from their former top positions. ■

Thus, we have an interesting two-way interplay between logical dynamics of belief change and the design of new non-monotonic consequence relations. [159]

7.10 Further themes: postulates and correspondence results

A brief comparison with AGM The best-known account of belief revision is *AGM theory* (Gaerdenfors 1988, Gaerdenfors & Rott 1995) that deals with three abstract operations of $+A$ ('update'), $*A$ ('revision'), $-A$ ('contraction') for factual information of a single agent. In contrast with the *DEL* mechanism for transforming plausibility models and generating complete logics, *AGM* analyzes belief change without proposing a specific construction, placing abstract algebraic postulates on the above operations instead. [160] This is the contrast between constructive and postulational approaches that we have seen in Chapter 3.

The *AGM* postulates claim to constrain all reasonable revision rules. Do they apply to ours? Here is a simple test. The 'Success Postulate' says that all new information comes to be believed in the theory revised with this information: $A \in T{*}A$. But even *PAL* fails this test, and the difference is instructive. Success follows from our axioms for *factual* propositions, but it fails for complex epistemic or doxastic ones. For instance, true Moore-sentences cannot be believed after announcement. The main intuitions of belief revision theory (or generalized consequence relations) apply to factual assertions only, whereas a logic approach like ours insists on dealing with all types of proposition at once. [161]

---

[159] Technically, this suggests new open problems about complete sets of structural rules.

[160] We refer to the cited literature for a precise statement of the postulates.

[161] The same point about complex propositions and order dependence returns with other *AGM* postulates. For instance, the 'Conjunction Postulate' that compresses two updates into one mixes events that we distinguish: processing a conjunction of propositions, and processing two new

Here are two more differences between the frameworks. *AGM* deals with single agents only, while *DEL* is essentially multi-agent, including higher-order information about what others believe. And also tellingly, *DEL* analyzes not three, but an *infinity* of triggers for belief change, from public announcements to complex informational events.

Revision postulates as frame correspondences Despite the difference in thrust, the postulational approach to belief revision makes sense. A matching modal framework at the right level of generality is Dynamic Doxastic Logic (*DDL*, Segerberg 1995, 1999, Leitgeb & Segerberg 2007). This merely assumes some relation change on the current model, functional or relational, without specifying it further. The main operator then becomes:

Definition Abstract modal logic of model change.
Let $\boldsymbol{M}$ be a model, $[[P]]$ the set of worlds in $\boldsymbol{M}$ satisfying P, and $\boldsymbol{M}*[[P]]$ some new model. For the matching modal operator, we set $\boldsymbol{M}, s \models [*P]\varphi$ iff $\boldsymbol{M}*[[P]], s \models \varphi$. ■

DDL models resemble Lewis spheres for conditional logic, or their neighbourhood versions (Girard 2008). The minimal modal logic K is valid, and on top, further axioms constrain relation changes for bona fide revision policies. In the limit, a particular set of axioms might even determine one particular policy.

We show how this ties up with our approach in terms of *frame correspondence*, just as we did for *PAL* update in Chapter 3 by postulating key recursion axioms, and then seeing which update operations qualified on an abstract universe of models and transitions.

Usually, frame correspondences analyze semantic content of given axioms in one model for a static modal language. But one can just as well take the above setting of relation changing operations ♥P over a family of models (with worlds and a ternary comparison

---

propositions successively. For factual propositions in *PAL*, this amounts to the same thing, but not for complex epistemic ones. And in our dynamic logic of belief, even the factual case is problematic. Here is why. In *PAL*, successive public announcements could still be compressed by the law $[!P][!Q]\varphi \leftrightarrow [!(P \wedge [!P]Q)]\varphi$. But two successive upgrades $\Uparrow P;\ \Uparrow Q$ rearrange a model as follows. First, P-worlds come on top of $\neg P$-ones, then the same happens with Q. The result is the order pattern $PQ \geq \neg PQ \geq P\neg Q \geq \neg P\neg Q$. No single upgrade does this, and no iteration law compresses the effect of two revision steps to just one with the same effects on conditional belief.

relation $x \leq_s y$). $\heartsuit P$ takes any model $\boldsymbol{M}$ and set of worlds P in it, [162] and yields a new model $\boldsymbol{M}\heartsuit P$ with the same worlds but a new relation $\leq_s$. Axioms may then constrain this. [163]

Analyzing a few AGM postulates For a start, the Success Postulate says something weak, that holds for both earlier operations $\Uparrow P$ and $\uparrow P$: [164]

Fact The formula $[\heartsuit p]Bp$ says that the best worlds in $\boldsymbol{M}\heartsuit p$ are all in p.

But we can also demand something stronger, that the best worlds in $\boldsymbol{M}\heartsuit p$ are precisely the best p-worlds in $\boldsymbol{M}$ (the upper class law '*UC*'). This, too, can be expressed. But we need a stronger dynamic formula, involving two different proposition letters p and q:

Fact The formula $B^p q \leftrightarrow [\heartsuit p]Bq$ expresses *UC*.

But this preoccupation with the upper classes still fails to constrain the total relation change. For that, we must look at the new social order in all classes after the Revolution, i.e., at conditional beliefs following relation upgrade. As an illustration, we consider the key reduction axiom for $\Uparrow P$, using proposition letters instead of schematic variables. [165] The following shows how this determines lexicographic reordering of models completely (again we use the earlier auxiliary existential modality E):

Theorem The formula $[\heartsuit p]\, B^r q \leftrightarrow (E(p \wedge r) \wedge B^{p \wedge r} q) \vee (\neg E(p \wedge r) \wedge B^r q)$ holds in a universe of frames iff the operation $\heartsuit p$ is lexicographic upgrade.

Proof That the principle holds for lexicographic upgrade was our earlier soundness result. Next, let the principle hold for all set values of q and r (but p is kept fixed). First, we show that, if $x \leq_s y$ in $\boldsymbol{M}\heartsuit p$, then this pair was produced by lexicographic upgrade. Let r be the

---

[162] Here we have dropped the above double denotation brackets *[[P]]* for convenience.

[163] Even more abstract spaces of models can be used here as general background for analyzing the content of dynamic axioms, but our setting suffices to make our main points.

[164] Frame correspondence has a format like this (van Benthem to appearB). The modal axiom $\Box p \rightarrow \Box\Box p$ is true at world s in frame $\boldsymbol{F} = (W, R)$ iff R is transitive at s: i.e., $\boldsymbol{F}, s \models \forall y(Rxy \rightarrow \forall z(Ryz \rightarrow Rxz))$. Frame truth is truth under all valuations on frame $\boldsymbol{F}$ for its proposition letters. Thus, it does not matter whether we use formula $\Box p \rightarrow \Box\Box p$ or schema $\Box\varphi \rightarrow \Box\Box\varphi$. Not so for *PAL* and *DEL*, where plain and *schematic* validity differ. In the following proofs, we use proposition letters for sets of worlds, by-passing issues of changes in truth value across updates.

[165] Thus we suppress the earlier dynamic modalities $[\Uparrow P]\psi$ that were sensitive to transfer effects.

set $\{x, y\}$ and $q = \{y\}$. Then the left-hand side of our axiom is true. Hence the right-hand side is true as well, and there are two cases. *Case 1*, $E(p \wedge r)$: one of x, y is in p, and hence $p \wedge r = \{x, y\}$ (1.1) or $\{x\}$ (1.2) or $\{y\}$ (1.3). Moreover, $B^{p \wedge r}\, q$ holds in $\boldsymbol{M}$ at s. If (1.1), we have $x \leq_s y$ in $\boldsymbol{M}$, with both x, y in p. If (1.2), we must have $y=x$, and by reflexivity, again $x \leq_s y$ in $\boldsymbol{M}$. Case (1.3) can only occur when $y \in p$ and $x \notin p$: the typical case for upgrade. *Case 2*, $\neg E(p \wedge r)$: x, y are not in p. The true disjunct $B^{r}\, q$ says that $x \leq_s y$ in $\boldsymbol{M}$.

Conversely, we show that all pairs satisfying the description of lexicographic upgrade in $\boldsymbol{M}$ make it into the new order in $\boldsymbol{M}\heartsuit p$. Here is one example: the other case is similar. Suppose that $y \in p$ while $x \notin p$. Set $r = \{x, y\}$ and $q = \{y\}$, whence $p \wedge r = \{y\}$. This makes $(E(p \wedge r) \wedge B^{p \wedge r}\, q)$ true at world s in $\boldsymbol{M}$, and hence also the whole disjunction to the right. The left-hand formula $[\heartsuit p]\, B^{r}\, q$ is then also true at s in $\boldsymbol{M}$. But this tells us that *in the new model* $\boldsymbol{M}\heartsuit p$, $B^{r}\, q$ holds at s. Thus, the best worlds in $\{x, y\}$ are in $\{y\}$: i.e., $x \leq_s y$ in $\boldsymbol{M}\heartsuit p$. ■

This can be generalized to abstract universes of transitions, quantifying over sets of worlds inside and across plausibility models. [166] [167] But even our simple setting shows how frame correspondence for languages with model-changing modalities is a good way of doing abstract postulational analysis of update and revision. [168]

7.11 Still more issues, and open problems

Variations on the static models We assumed agents with epistemic introspection of their plausibility order. Without this, we would need *ternary* world-dependent plausibility relations, as in conditional logic. What do our $\{K, B\}$-based systems look like then?

Also, safe belief suggests having just one primitive plausibility pre-order $\leq$, defining knowledge as truth in all worlds, whether *less* or more plausible (cf. van Eijck & Sietsma 2009 on *PDL* over such models). What happens to our themes in the latter setting?

---

[166] The above arguments then work uniformly by standard modal substitution techniques.

[167] Further *AGM*-postulates mix two operations that change models: *update !P* and *upgrade* $\heartsuit P$, with laws like *(a)* $[\heartsuit(p \wedge q)]Br \rightarrow [!q][\heartsuit p]Br$, *(b)* $([\heartsuit p]Eq \wedge [!q][\heartsuit p]Br) \rightarrow [\heartsuit(p \wedge q)]Br$. These constrain simultaneous choice of two abstract model changing operations.

[168] Still, this is not the only abstract perspective on our logics. In Chapter 12, we do an alternative postulational analysis of the earlier Priority Update rule in terms of social choice theory.

Finally, many authors have proposed basing doxastic logic on *neighbourhood models* (cf. Girard 2008, Zvesper 2009). How should we lift our theory to that setting?

Common belief and social merge We have not analyzed common belief, satisfying the fixed-point equation $CB_G\varphi \leftrightarrow \bigwedge_{i\in G} B_i\,(\varphi \wedge CB_G\varphi)$. Technically, a complete dynamic logic seems to call for a combination of relation change with the *E-PDL* techniques of Chapter 4. More generally, belief revision policies describe what a single agent does when confronted with surprising facts. But beliefs often change because *other agents* contradict us, and we need interactive settings where agents create one new plausibility ordering. These include events of *belief merge* (Maynard-Reid & Shoham 1998) and *judgment aggregation* (List & Pettit 2004). Construed either way, we need look at *groups*: Chapter 12 has more.

Model theory of plausibility models and upgrade Our static logics for belief raise standard modal issues of appropriate notions of bisimulation and frame correspondence techniques for non-standard modalities that maximize over orderings. Like for conditional logic, these model-theoretic issues seem largely unexplored. In the dynamic setting, relation-changing modalities also raise additional issues, such as respect for plausibility bisimulation. Many of the themes in Chapters 2, 3 remain to be investigated for our systems here.

Syntactic belief revision There is also belief revision in a more fine-grained inferential sense, where agents may have fallible beliefs about what follows from their data, or about the consistency of their views. Such beliefs can be refuted by events like unexpected turns in argumentation or discussion. These scenarios are not captured by the semantic models of this chapter. How to do a syntactic dynamic logic in the spirit of Chapter 5?

Proof theory of logics for revision Our logics had complete Hilbert-style axiomatizations. Still, we have not looked at more detailed proof formats dealing with dynamic modalities, say, in natural deduction or semantic tableau style. What would these look like? The same question makes sense, of course, for *PAL* and *DEL* in Chapters 3, 4.

Backward versus forward once more Recall a basic contrast from earlier chapters. *AGM* is *forward-looking*, with postconditions of coming to believe. [169] But like *DEL* (backward-looking in its version without factual change), our logics compute what agents will believe only via preconditions. More generally, we want to merge local dynamics of belief change

[169] This gets harder to maintain with complex instructions like ‘make agent *i* not believe that φ’.

with a temporal Grand Stage where an instruction $*P$ wants a *minimal* move to some future state where one believes that P. We will do this in Chapter 11. [170]

7.12 Literature

A key source for belief revision theory in the postulational style is Gärdenfors 1988. Segerberg 1995 connects this to a 'dynamic doxastic logic' on neighbourhood models. Aucher 2004 treats quantitative belief revision rules in a *DEL* format, mixing event models with Spohn-style graded models. Van Benthem 2007B is the main source for this chapter, axiomatizing dynamic logics for qualitative transformations of plausibility orders in a style similar to van Benthem & Liu 2007. Baltag & Smets 2006 gave the first general qualitative *DEL* version of belief change, with priority product update using plausibility event models. Baltag & Smets 2008 provide a mature version. Girard 2008 connects *DDL* to *DEL*, while Baltag, van Ditmarsch & Moss 2008 give broader framework comparisons and history.

[170] My first analysis of *AGM* in van Benthem 1989 was in this style, with modalities $[+P]$, $[-P]$, and $[*P]$ for update, contraction, and revision on a temporal universe of information stages.

Chapter 8 AN ENCOUNTER WITH PROBABILITY

Our dynamic logics deal with agents' knowledge and beliefs under information update. But update has long been the engine of *probability theory*, a major life-style in science and philosophy. This chapter is an intermezzo linking the two perspectives, without any pretense at completeness, and assuming the basics of probability theory without further ado. We will show how probability theory fits well with dynamic-epistemic logics, leading to a new update mechanism that merges three aspects: prior world probability, occurrence probability of events, and observation probability. The resulting logic can be axiomatized in our standard style, leading to interesting comparisons with probabilistic methods.

8.1 Probabilistic update

The absolute basics A *probability space* $\boldsymbol{M} = (W, \boldsymbol{X}, P)$ is a set of worlds W with a family $\boldsymbol{X}$ of propositions that can be true or false at worlds, plus a probability measure P on propositions. This is like single-agent epistemic models, but with refined information on agent's views of propositions. In what follows, we use high-school basics: no σ–algebras, measurable sets, and all that. Intuitions come from finite probability spaces, putting each subset in $\boldsymbol{X}$. Under this huge simplification, the function P assigns values to single worlds, while its values for larger sets are automatic by addition, subject to the usual laws:

$$P(\neg\varphi) = 1 - P(\varphi), \qquad P(\varphi \vee \psi) = P(\varphi) + P(\psi), \text{ if } \varphi, \psi \text{ are disjoint}$$

Of crucial relevance to update are *conditional probabilities* $P(\varphi \mid A)$ giving the probability for ϕ given that A is the case, using P rescaled to the set of worlds satisfying A: [171]

$$P(\varphi \mid A) \;=\; P(\varphi \wedge A) \,/\, P(A)$$

Bayes' Rule then computes conditional probabilities through the derivable equation

$$P\,(\varphi \mid A) \;=\; P\,(A \mid \varphi) \bullet P(\varphi) \;/\, P(A)$$

and more elaborate versions. This describes probability change as new factual information A comes in. Other basic mechanisms update with non-factual information. E.g., the *Jeffrey Rule* updates with probabilistic information $P_i(A) = x$ by setting the new probability for A to x, and apart from that, redistributing probabilities among worlds inside the A and $\neg A$ zones proportionally to the old probabilities. And there are even further update rules.

[171] We now use '*A*' for propositions, since the earlier '*P*' is taken for probability.

Probabilistic and logical models Conditional probability resembles the conditional belief of Chapter 7, while the informal explanations surrounding it also have a whiff of *PAL* (Chapter 3). One zooms in on the worlds where the new A holds, and re-computes probabilities. This is like eliminating $\neg A$-worlds, and re-evaluating epistemic or doxastic modalities. And Jeffrey Update is like the radical revision policy $\Uparrow A$ in Chapter 7 that fixes a belief to be achieved, but only minimally changes the plausibility order otherwise.

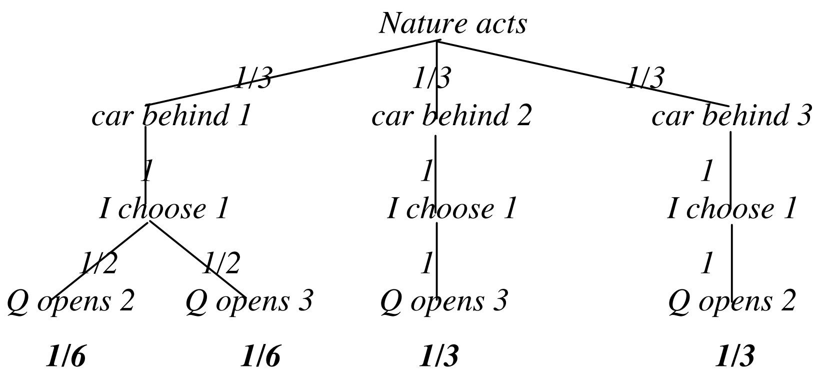
There are also differences. Crucially, dynamic logics make a difference between truth values of propositions φ before and after update, and hence a conditional belief $B^{\psi}\varphi$ was not quite the same as a *belief after update*: $[!\psi]B\varphi$. This made no difference for factual φ, ψ, but it did when formulas are more complex, as our logics want to have them. Also, our logics are about many agents, with syntactic *iterations* like "I know that your probability is high", or "My probability for your knowing φ equals y", that are scarce in probability theory. Finally, *DEL* update (cf. Chapter 4) did not just select subsets of a model: it *transforms* current models $\boldsymbol{M}$ into perhaps complex new ones. Thus, dynamic formulas $[\boldsymbol{E}, e]\varphi$ conditionalize over drastic model-transforming actions.

It is time to start with the systematic proposal of this chapter. Model constructions are not alien to a probabilistic perspective. Competent practitioners make such moves implicitly. Consider the following famous puzzle:

Example Monty Hall dilemma.
There is a car behind one of three doors: the quizmaster knows which one, you do not. You choose a door, say *1*, and then the quizmaster opens another door that has no car behind it. Say, he opens Door *3*. Now you are offered the chance to switch doors. Should you?

If you conditionalize on the new information that the car is not behind Door 2, Bayes' Rule tells you it does not matter: Doors *1* and *3* have equal probability now. But the real information is more complex: if the car is behind *1*, the quizmaster can open either Door *2* or Door *3* with probability ½. But if it is behind Door *3*, then he must open Door *2*.

This drives a construction of the probability space, as pictured in the following tree, where probabilities of branches arise by multiplying those of their successive transitions:

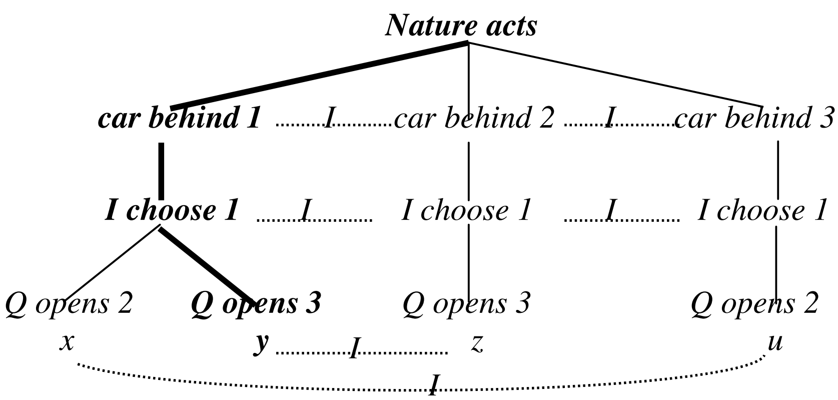


I chose *1*, *Q* opened *3*. We want the conditional probability that the car is behind Door *1*, given what we saw. The tree suggests *A = Q opens 3,* leaving two branches. Analyzing that subspace, we find that *P(Car behind 1| A) = 1/3* – and hence we should switch. ■

Now we look at the same model changes using dynamic epistemic product update.

Example Monty Hall in *DEL*.

Nature's actions are indistinguishable for me (*I*), but not for the quizmaster *Q*. The result is the epistemic model at the second level. Now I choose Door *1,* which copies the same uncertainties to the third level. Then come public events of *Q*'s opening a door, each with preconditions (a) I did not choose that door, and (b) *Q knows that* the car is not behind it. In principle this could generate *3x3 = 9* worlds, but the preconditions leave only *4*:



The actual world is *y* on the bold-face branch. In the bottom-most epistemic model, I know the world is either *y* or *z*. Throughout the tree, Quizmaster knows exactly where he is. ■

All this suggests finding a richer logic on top of this that can accommodate probabilistic fine-structure. This requires the usual steps from earlier chapters: a suitably expressive static language, a product update rule, and a complete dynamic logic. This is not totally routine, and we will be forced to think about which probabilities play a role in update.

8.2 Static epistemic probabilistic logic

Epistemic probabilistic languages describing what agents know plus the probabilities they assign were introduced by Halpern and Tuttle 1993, Fagin and Halpern 1993:

Definition Epistemic probability models.

An *epistemic probability model* $\boldsymbol{M} = (W, \boldsymbol{\sim}, \boldsymbol{P}, V)$ has a set of worlds W, a family $\boldsymbol{\sim}$ of equivalence relations $\sim_i$ on W for each agent i, a set $\boldsymbol{P}$ of probability functions P_i that assign probability distributions for each agent i at each world $w \in W$, and finally, a valuation V assigning sets of worlds to proposition letters. [172] ■

These models represent both non-probabilistic and probabilistic information of agents.

Definition Static epistemic probabilistic language.

The *static epistemic-probabilistic language* has the following inductive syntax:

$\varphi ::= p \mid \neg\varphi \mid (\varphi\wedge\psi) \mid K_i\varphi \mid P_i(\varphi) = q$, where q is a rational number, [173] plus linear inequalities $\alpha_1 \bullet P_i(\varphi_1) + \ldots + \alpha_n \bullet P_i(\varphi_n) \geq \beta$ with $\alpha_1, \ldots, \alpha_n, \beta$ rational numbers. ■

This allows formulas like $K_iP_j(\varphi) = k$, or $P_i(K_j\varphi) = k$ talking about agents' knowledge of the others' probabilities, or probabilities they give to someone knowing some fact. [174] Formulas $P_i(\varphi) = q$ are evaluated by summing over the worlds where φ holds:

Definition Semantics for epistemic-probabilistic logic.

The clauses for proposition letters, Boolean operations and epistemic modalities are as in Chapter 2. Here is the key clause for the probabilistic modality:

$$\boldsymbol{M}, s \models P_i(\varphi) = q \text{ iff } \sum_{t \text{ with } \boldsymbol{M}, t \models \varphi} P_i(s)(t) = q$$

The semantic explanation for the linear inequalities then follows immediately. ■

Epistemic probability models suggest constraints linking probability with knowledge (cf. Halpern 2003). Say, one can let $P_i(s)$ assign positive probabilities only to worlds $\sim_i$-linked to s. Such constraints define models with special logics. In particular, epistemically indistinguishable worlds often get the same probability distribution. Thus, agents will know the probabilities they assign to propositions, and hence we have a valid principle

---

[172] An important and intuitive special case is when the $P_i(w)$ are probability distributions defined only on the equivalence class of epistemically accessible worlds $\{v \mid v \sim_i w\}$.

[173] Another widely used notation has superscripts for agents: $P^i(\varphi) = q$.

[174] For convenience, we will drop agent indices whenever they do not help the presentation.

$P_i(\varphi) = q \rightarrow K_i\, P_i(\varphi) = q$ *Probabilistic Introspection*

This may be compared with our introspective treatment of beliefs in Chapter 7. The reader should feel free to assume this special setting in what follows, if it helps for concreteness – but as always, our analysis of probabilistic update mostly does not hinge on such options. [175] What matters is expressive harmony between static and dynamic languages. This *pre-encoding* uses probabilistic linear inequalities, whose purpose will become clear later.

8.3 Three roles of probability in update scenarios

Earlier update rules Merges of *PAL* with probabilistic update occur in Kooi 2003. In line with Chapter 3, there are prior world probabilities in an initial model ***M***, and one then conditionalizes to get the new probabilities in the model ***M***/*A* after a public announcement *!A*. [176] This validates the following key *recursion axiom*:

$$[!A]\, P_i(\varphi) = q \;\leftrightarrow\; P_i\,([!A]\varphi \mid A) = q$$

reducing a probability after update to a conditional probability before. [177]

This cannot deal with Monty Hall, as the Quizmaster's actions had different probabilities, depending on the world where they occur. To deal with this new feature, van Benthem 2003 introduced 'occurrence probabilities' for publicly observable events in event models, and assigning the following probabilities to new worlds *(s, e)*:

$P^{\boldsymbol{MxE}}(s, e) = P^{\boldsymbol{M}}(s) \bullet P^{\boldsymbol{E}}_s(e),$ followed by a normalization step.

More precisely, when all events are public, and their occurrence probabilities are common knowledge, the product rule reads as follows:

$$P_{i,\,(s,e)}\,(t, e) \quad = \quad \frac{P_{i,\,s}\,(t) \bullet P_t\,(e)}{\Sigma_{\,u \sim i\ s\ in\ M}\, P_{i,s}(u) \bullet P_u\,(e)}$$ [178]

---

[175] There are more general probabilistic models in the literature, such as Dempster-Shaefer theory, but we feel confident that our dynamic style of analysis will work there, too.

[176] This update rule, as well as the others in this chapter, works as long as the new proposition does not have *probability zero*. For more on the latter scenario, see the final section of this chapter.

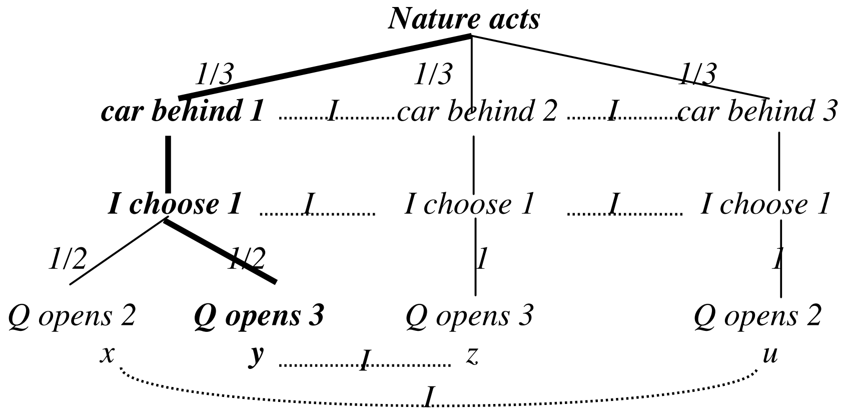
[177] Kooi's paper also has a notion of *probabilistic bisimulation* that fits the language.

This subsumes Kooi's rule for the special event of public announcement:

$$P_{i,\ (s,\,!A)}(\varphi) \quad = \quad \frac{\Sigma\{P_{i,\,s}(u) \mid s \sim_i u \ \&\ \boldsymbol{M}, u \models A \wedge [!A]\varphi\}}{\Sigma\{P_{i,\,s}(u) \mid s \sim_i u \ \&\ \boldsymbol{M}, u \models A\}}$$

Example Monty Hall via product update.

Here is how one computes the earlier tree probabilities for Monty Hall:



It is easy to check that the probabilities in the final set *{x, y}* work out to

for *y*: $(1/3 \bullet 1/2)\,/\,(1/3 \bullet 1/2 + 1/3 \bullet 1) \quad = \quad 1/3$

for *z*: $(1/3 \bullet 1)\,/\,(1/3 \bullet 1/2 + 1/3 \bullet 1) \quad = \quad 2/3$ ■

Repercussions: Bayes' Rule This dynamic perspective has some features that may be controversial. One difference with probability theory is that we conditionalize on *events*. We observe events, and hence we have probabilities with heterogeneous arguments

$P(\varphi \mid e)$ where φ is a proposition, and e is an event. [179]

Now consider Bayes' Law. Is this principle plausible in a dynamic reading? Obviously, its static form $P(\varphi \mid A) \ = \ P(A \mid \varphi) \bullet P(\varphi) \ / P(A)$ holds for probabilities in a fixed model. It is also valid if we restrict attention to update with purely factual assertions. But the Rule turns out problematic as an update principle relating a new probability model to an old one, since formulas can change truth values as new information comes in:

---

[178] In general encoding information about occurrence probabilities was done via 'generalized preconditions', that map worlds in arbitrary models to probabilities. This is a huge leap away from *DEL*, but later on, we will see that we can often make do with finite definable versions.

[179] Events of public announcement *!A* reduced to a propositional *precondition A*, and *Q*'s opening Door *3* to the *postcondition* '*Q* opened Door *3*'. But there need not be a general reduction.

Proposition Bayes' Rule fails as a law of epistemic public announcement.

Proof Consider the following epistemic model ***M*** with two agents:

$$\begin{tikzcd} \boldsymbol{p, q} \arrow[r, dash, "you"] \arrow[d, dash, "me"'] & \neg p, \neg q \\ \neg p, q & \end{tikzcd}$$

The actual world on the top left has p, q both true. Now consider the assertion

A You do not know whether *p* is the case

This is true in the two uppermost worlds, but not at the bottom. Next, take

φ I know whether *p* is the case

that only holds in the world to the right. Let each world have probability *1/3*. [180] In the initial model, $P_M(A) = 2/3$, while $P_M(\varphi) = 1/3$. A public announcement of the true fact *A* updates this model to the new model ***M****/A*:

$$\begin{tikzcd} \boldsymbol{p, q} \arrow[r, dash, "you"] & \neg p, \neg q \end{tikzcd}$$

where φ has become true everywhere. In that new model, then, $P_{M/A}(\varphi) = 1$. By contrast, an update that first announces φ would take ***M*** to the one-world model ***M****/*φ*:*

$$\neg p, \neg q$$

Here *A* is false everywhere, and we get a probability $P_{M/\varphi}(A) = 0$. Substituting all these values, we see that Bayes' Rule fails in its dynamic reading:

$$P(\varphi / A) = 1 \neq (0 \bullet 1/3) / 2/3 \qquad \blacksquare$$

In a dynamic perspective, order inversions are invalid, though they may work for special formulas. Still, Bayes' Rule has lived a useful life for centuries without logical blessing. Romeijn 2009 gives a modified Bayesian counter-analysis of the above reasoning.

Now we turn to our general analysis, based on van Benthem, Gerbrandy & Kooi 2009.

Three sources of probability The preceding approaches have performed a two-fold 'probabilization' of *DEL* product update, distinguishing two factors:

---

[180] One can think of probabilities *P* for a third person *3* who holds all worlds equi-possible.

(a) *prior probabilities of worlds* in the current epistemic-probabilistic model ***M***, representing agents' current informational attitudes,

(b) *occurrence probabilities for events* from the event model ***E*** encoding agents' views on what sort of process produces the new information.

But there is also a third type of uncertainty that plays in many realistic scenarios:

(c) *observation probability of events*, reflecting agents' uncertainty as to which event is actually being observed.

Recall the motivation for *DEL* in terms of observational access. I see you read a letter, and I know it is a rejection or an acceptance. You know the actual event (reading "Yes", or reading "No"), I do not. Here product update gives a new epistemic model without probabilities. To compute the latter, I may know about frequency of acceptance versus rejection letters, a type (b) occurrence probability. But there may also be more information *in the observation itself*! Perhaps I saw a glimpse of your letter – or you looked smug, and I think you were probably reading an acceptance. This would be an observation probability in sense (c). The latter notion is also known from scenario's motivating the Jeffrey Rule, where one is uncertain about evidence received under partial observation. [181]

A simple scenario where all three kinds of probability come together is as follows:

Example The hypochondriac.

You read about a disease, and start wondering. The chances of having the disease are slight, *1* in *100.000*. You read that one symptom is a certain gland being swollen. With the disease, the chance of this is *97%*, without the disease, it is *0*. You take a look. It is the first time you examine the gland and you do not know its proper size. You think chances are *50%* that the gland is swollen. What chance should you assign to having the disease? ■

We will now define a general mechanism for computing the answer.

8.4 Probabilistic product update

In what follows, our static epistemic-probabilistic models ***M*** are as before, and so is our language. We will use the *DEL*-notation $[\boldsymbol{E}, e]\varphi$ from Chapter 4 to describe the effect of executing event model ***E****, e* in a current model ***M****, s*. But our event models are a bit special,

[181] Occurrence probability is often an objective frequency, and observation probability a subjective chance. Our distinction lets both major views of probability co-exist within the same scenario.

to make them look like processes with uniformly specified occurrence probabilities:

Definition Probabilistic event models.

Probabilistic event models are structures $\boldsymbol{E} = (E, \sim, \Phi, Pre, P)$ with (a) E a non-empty finite set of events, (b) $\sim$ a set of equivalence relations $\sim_i$ on E for each agent i, (c) Φ a set of pairwise inconsistent sentences ('preconditions'), (d) Pre assigns to each precondition $\varphi \in \Phi$ a probability distribution over E (we write $Pre(\varphi, e)$, the chance that e occurs given φ), and (e) for each i, P_i assigns each event e a probability distribution over E. ■

The language for preconditions is given below: as in Chapter 4, there is a simultaneous recursion. Models work as follows. The *Pre* specifies occurrence probabilities of a process that makes events occur with probabilities depending on conditions Φ. Diseases and quizmasters are examples, with rules of the form "if P holds, then do a with probability q", and so are Markov processes. Models also have observation probabilities, represented by the functions P_i. The probability $P_i(e)(e')$ is the probability assigned by an agent i to event e' taking place, given that e actually takes place. This adds probabilistic structure to the uncertainty relations $\sim_i$ in much the same way as happened in our static models. [182]

Our next goal is a dynamic update rule for these models. Merging input from all three sources of probability, its mechanism is a direct generalization of earlier rules:

Definition Probabilistic Product Update Rule.

Let $\boldsymbol{M}$ be an epistemic-probabilistic model and let $\boldsymbol{E}$ be an event model. If s is a state in $\boldsymbol{M}$, write $Pre(s, e)$ for the value of $pre(\varphi, e)$ with φ the unique element of Φ true at $\boldsymbol{M}, s$. If no such φ exists, set $pre(s, e) = 0$. The *product model* $\boldsymbol{M} \times \boldsymbol{E} = (S', \sim', P', V')$ is defined by:

(a) $S' = \{ (s, e) \mid s \in S, e \in E \text{ and } pre(s, e) > 0\}$

(b) $(s, e) \sim_i (s', e')$ iff $s \sim_i s'$ and $e \sim_i e'$

(c) $P'_i((s, e), (s', e')) :=$

$$\frac{P_i(s)(s') \bullet Pre(s'; e') \bullet P_i(e)(e')}{\sum_{s'' \in S, e'' \in E} P_i(s)(s'') \bullet Pre(s'', e'') \bullet P_i(e)(e'')}$$ if the denominator > 0 and 0 otherwise.

(d) $V'((s, e)) = V(s)$ ■

[182] Each epistemic event model $\boldsymbol{E}$ in Chapter 4 can be expanded to a probabilistic event model.

The new state space after the update consists of all pairs *(s, e)* where event *e* occurs with positive probability in *s* (as specified by *Pre*). The crucial part are the new probabilities $P'_i\,(s, e)$ for *(s', e')*. These are a product of the prior probability for *s'*, the probability that *e'* actually occurs in *s'*, and the probability that *i* assigns to observing *e'*. To get a proper probability measure, we normalize this value. [183] Here is how this rule works in practice:

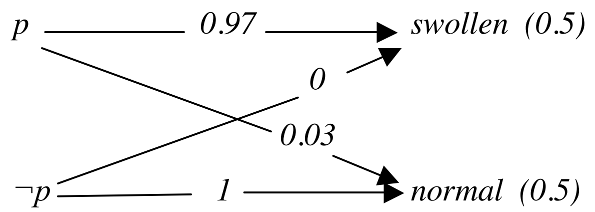
Example The Hypochondriac again.

The initial hypothesis about having disease *p* is captured by a prior probability distribution

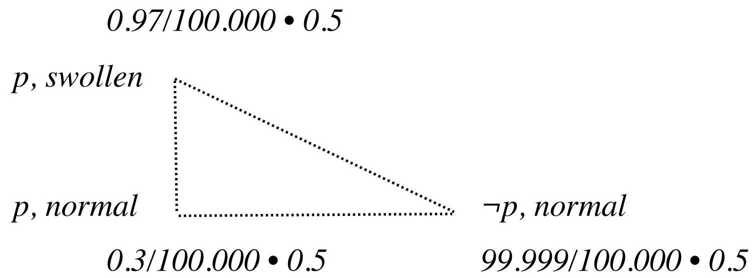
1/100.000 *99.999/100.000*

p ········· *¬p*

Then the hypochondriac examines the gland, with an occurrence probability (if he has the disease, the gland is swollen with probability *0.97*) and an observation probability (he thinks he is seeing a swollen gland with probability *0.5*) as in the given scenario. This is encoded in the following epistemic probabilistic event model:



The product of our initial state with this model is as follows:



This diagram is our new information state after the episode. Renormalizing values, the new probability that the Hypochondriac should assign to having the disease is still *1* in *100.000*. His inconclusive observation has not produced any information about having the disease. Had he found it more probable that the gland was swollen, the probability of the disease

---

[183] If the denominator in our rule sums to *0*, we stipulated a total value *0*. Thus ***M x E*** need not be a probabilistic epistemic model: $P_i(s, e)$ may assign probability *0* to all worlds. Bacchus 1990 has a probabilistic defense, but there also are ways of circumventing this feature.

would have come out higher than before by our product rule, and had he found it more probable that it was not swollen, that probability would have been lower. ■

We also see a typical *DEL* feature. Initially, we only had 2 options: having the disease or not. The update created 3 worlds, now with information if the gland is swollen or not.

8.5 Discussion and further developments

Systematic model construction Our epistemic probabilistic update rule starts from a simple probability space and, step by step, builds more complex product spaces with informational events encoded by event models. This control over probability spaces may be useful in practice, where management of relevant spaces, rather than applying the probability calculus, is the main difficulty in reasoning with uncertainty. Repeated over time, the new possibilities form all runs of a total informational process, linking up with more global epistemic probabilistic temporal logics (Chapter 11).

Model theory and probabilistic bisimulation The model theory of epistemic-probabilistic logic can be developed like its earlier counterparts, using the earlier-mentioned epistemic-probabilistic bisimulation for our static language. It is easy to see that our product update rule respects such bisimulations between input models, and we are on our way.

Shifting loci of probabilistic information A natural question is if our three components: prior world probabilities, occurrence, and observation probabilities are really independent. In modeling real scenarios one can *choose* where to locate things. Van Benthem, Gerbrandy & Kooi 2009 give constructions on update models showing how under redefinition of events, occurrence probabilities can absorb observation probabilities, and vice versa. Such tricks do not endanger the intuitive appeal of our three-source scheme.

Processes and protocols Our mechanism is more powerful than may appear at first sight. Consider temporal processes over time (cf. Chapters 3, 4, 11). Many of their *protocols* are probabilistic: say, an agent whose assertions have a certain probabilistic reliability.

Example Coins and Liars.

We know that some coin is either fair, or yields only heads. We represent observation of a throw of the coin with an initial model ***M*** with two options *Fair, Heads-only,* plus an event model ***E*** with two events *Heads* and *Tails*, related with the obvious probabilities:

Pre (Fair, Heads) = Pre (Fair, Tails) = 1/2 ,

Pre (Heads-only, Heads) = 1, Pre (Heads-only, Tails) = 0.

By our product update rule, one observation of *Tails* rules out the unfair coin, each observation of *Heads* makes it more likely. In the same mood, we meet a stranger in a logic puzzle, who might be a *Truth Teller* or a *Liar*. We have to find out what is what. We encode the options as abstract pair events inside the event model (cf. Chapter 4):

(Truth Teller, !A), (Liar, !A)

encoding both the assertion made, and the type of agent making it. After that, update is exactly as in our earlier examples, and we read off the new values for the agent types. ■

A general construction in terms of pairs *(Process type, Observed event)* is easily stated.

8.6 A complete dynamic probabilistic logic

Language and semantics To reason explicitly about probabilistic information change, we extend static epistemic-probabilistic logics with the dynamics of Chapters 3, 4.

Definition Dynamic-epistemic-probabilistic language.
The *dynamic-epistemic-probabilistic language* extends our static language with a dynamic modality $[\boldsymbol{E}, e]\varphi$, with $\boldsymbol{E}$ a probabilistic event model, and e an event from its domain. ■

Note again the recursion: the formulas that define preconditions come from this language, but through the new dynamic modalities, such models themselves enter the language.

Definition Semantics of probabilistic event models.
In an epistemic probability model $\boldsymbol{M} = (W, \sim, P, V)$ with $s \in W$, $\boldsymbol{M}, s \models [\boldsymbol{E}, e]\varphi$ iff for the unique $\psi \in \Phi$ with $\boldsymbol{M}, s \models \psi$, we have $\boldsymbol{M} \text{ x } \boldsymbol{E}, (s, e) \models \varphi$ in the product $\boldsymbol{M} \text{ x } \boldsymbol{E}$ as above. ■

A complete axiomatic system With all this in place, here is our main result:

Theorem The dynamic-epistemic probabilistic logic of update by probabilistic event models is completely axiomatizable over the chosen static logic.

Proof The core is the key recursion axiom for formulas $[\boldsymbol{E}, e]\varphi$. The following calculation is the heart of our reduction, with agent indices dropped for greater readability. Consider the value $P(\psi)$ of a formula ψ in a (pointed) product model $(\boldsymbol{M}, s) \text{ x } (\boldsymbol{E}, e)$. We abbreviate $P(-)$ in the initial model by $P^{\boldsymbol{M}}$, writing $P^{\boldsymbol{M}x\boldsymbol{E}}$ for values $P(-, -)$ in the product model, and $P^{\boldsymbol{E}}$ for $P(-)$ in the event model. For convenience, we will use the existential dynamic modality $\langle \boldsymbol{E}, e \rangle$. For a start, if we have $\sum_{s'' \in S,\, e'' \in E} P^{\boldsymbol{M}}(s'') \bullet Pre(s'', e'') \bullet P^{\boldsymbol{E}}(e'') > 0$, then we get:

$$P^{MxE}(\psi)$$

$$= \quad \sum_{(s', e') \text{ in } \mathbf{MxE}: \mathbf{MxE}, (s', e') \models \psi} P^{MxE}(s', e')$$

$$= \quad \sum_{s' \in S, e' \in E: \mathbf{M}, s' \models <\mathbf{E}, e'>\psi} P^{MxE}(s', e')$$

$$= \quad \frac{\sum_{s' \in S, e' \in E: \mathbf{M}, s' \models <\mathbf{E}, e'>\psi} P^{M}(s') \bullet Pre(s', e') \bullet P^{E}(e')}{\sum_{s'' \in S, e'' \in E} P^{M}(s'') \bullet Pre(s'', e'') \bullet P^{E}(e'')}$$

The numerator of this last equation can be written as

$$\sum_{\varphi \in \Phi, s' \in S, e' \in E, \mathbf{M}, s' \models \varphi, \mathbf{M}, s' \models <\mathbf{E}, e'>\psi} P^{M}(s') \bullet Pre(\varphi, e') \bullet P^{E}(e')$$

which is equivalent to

$$\sum_{\varphi \in \Phi, e' \in E} P^{M}(\varphi \wedge <\mathbf{E}, e'>\psi) \bullet Pre(\varphi, e') \bullet P^{E}(e')$$

We can analyze the denominator of the equation in a similar way, and rewrite it as

$$\sum_{\varphi \in \Phi, e'' \in E} P^{M}(\varphi) \bullet Pre(\varphi, e'') \bullet P^{E}(e')$$

So we can write the probability $P^{MxE}(\psi)$ in the new model as a term of the following form:

$$P^{MxE}(\psi) = \frac{\sum_{\varphi \in \Phi, e' \in E} P^{M}(\varphi \wedge <\mathbf{E}, e'>\psi) \bullet k_{\varphi, e'}}{\sum_{\varphi \in \Phi, e'' \in E} P^{M}(\varphi) \bullet k_{\varphi, e''}}$$

where, for each φ and f, $k_{\varphi, f}$ is a constant, namely the value $Pre(\varphi, f) \bullet P^{E}(f)$.

We enumerate the finite set of preconditions Φ and the domain of $\mathbf{E}$ as $\varphi_0, \ldots, \varphi_n$ $e_0, \ldots, e_m$. Then we rewrite any dynamic formula $<\mathbf{E}, e> P(\psi) = r$ with P the probability after update to an equivalent equation in which P refers to probabilities in the prior model:

$$\frac{\sum_{1 \leq i \leq n, 1 \leq j \leq m} k_{\varphi i, ej} \bullet P(\varphi_i \wedge <\mathbf{E}, e_j>\psi)}{\sum_{1 \leq i \leq n, 1 \leq j \leq m} k_{\varphi i, ej} \bullet P(\varphi_i)} = r$$

And the latter can be rewritten as a sum of terms:

$$\sum_{1 \leq i \leq n, 1 \leq j \leq m} k_{\varphi i, ej} \bullet P(\varphi_i \wedge <\mathbf{E}, e_j>\psi) + \sum_{1 \leq i \leq n, 1 \leq j \leq m} -r \bullet k_{\varphi i, ej} \bullet P(\varphi_i) = 0$$

Now, to express these observations as one recursion axiom in our formal language, we need sums of terms. Our language with linear inequalities is up to just this job. But then, to restore the harmony of the total system, we must find a reduction for inequalities:

$$[\mathbf{E}, e]\ \alpha_1 \bullet P_i(\psi_1) + \ldots + \alpha_n \bullet P_i(\psi_n) \geq \beta$$

In this formula, we can replace separate terms $P(\psi_k)$ after the dynamic modal operator by their equivalents as just computed. [184] We then obtain an equivalent expression of the form

$$\sum_{1 \leq h \leq k,\ 1 \leq i \leq n,\ 1 \leq j \leq m} \alpha_h \bullet k_{\varphi i,\ ej} \bullet P(\varphi_i \wedge [\mathbf{E}, e_j]\psi_h) + \sum_{1 \leq i \leq n,\ 1 \leq j \leq m} -\beta \bullet k_{\varphi i,\ ej} \bullet P(\varphi_i) \geq 0$$

This is still an inequality χ inside our language. The full axiom then becomes

$$([\mathbf{E}, e]\ \alpha_1 \bullet P_i(\psi_1) + \ldots + \alpha_n \bullet P_i(\psi_n) \geq \beta) \leftrightarrow$$
$$((\sum_{1 \leq i \leq n,\ 1 \leq j \leq m} k_{\varphi i,\ ej} \bullet P(\varphi_i) > 0) \rightarrow \chi) \wedge (((\sum_{1 \leq i \leq n,\ 1 \leq j \leq m} k_{\varphi i,\ ej} \bullet P(\varphi_i) = 0) \rightarrow 0 \leq \beta)$$

This looks technical, but it can easily be computed in specific cases.

The other recursion axioms are as in Chapter 4, with preconditions $Pre_{\mathbf{E},\ e}$ of events e in our setting being the sentences $\bigvee_{\varphi \in \Phi,\ Pre(\varphi,\ e) \geq 0} \varphi$. Our proof concludes with the usual inside-out removal of dynamic modalities, effecting a reduction to the base logic. ■

Our relative style of axiomatization adding dynamics superstructure to a static base logic makes special sense in probabilistic settings, as it factors out the possibly high complexity of the underlying quantitative mathematical reasoning.

8.7 A challenge: weighted learning

Policies and weights Update may be more than our rule so far. Inductive logic (Carnap 1952), learning theory (Kelly 1996), and belief revision theory (Gärdenfors & Rott 1995) also stress *policies* on the part of agents. We have a probability distribution, we observe a new event. The new distribution depends on the *weights* agents assign to past experience versus the latest news. Different weights yield more radical or conservative policies.

Example (adapted from Halpern 2003) The Dark Room.
An object can be light or dark. We start with the equiprobability distribution. Now we see that, with probability $3/4$, the object is dark. What are the new probabilities? ■

Our earlier update rule weighs things here *equally*. We use signals *'Light'*, *'Dark'*, with occurrence probabilities 1 and 0 with the obvious Φ, and observation probabilities $1/4$, $3/4$. The new probability that the object is dark mixes these to a value between $1/2$ and $3/4$.

---

[184] *Caveat*. The denominator of the equation for the posterior probabilities must be greater than 0.

The αβγ formula Suppose agents give different weights to the three factors in our rule, say real values α, β, γ in $[0, 1]$. Here is a generalization:

Definition Weighted Product Update Rule. [185]

$$P^{new}((s, e); (s', e')) := \frac{P(s)(s' \mid \varphi_{s'}) \bullet P(s)(s')^{\alpha} \bullet Pre(s', e')^{\beta} \bullet P(e)(e')^{\gamma}}{\sum_{s'' \in S,\, e'' \in E} P(s)(s'' \mid \varphi_{s''}) \bullet P(s)(s'')^{\alpha} \bullet Pre(s'', e'')^{\beta} \bullet P(e)(e'')^{\gamma}}$$

if the denominator > 0 – and 0, otherwise. ■

Setting all three factors to 1 gives our original update. Setting $\alpha, \beta, \gamma = (0, 0, 1)$ is close to the earlier Jeffrey Update, mixing radicalism and conservatism. [186]

8.7 Conclusion

This chapter has linked dynamic epistemic logic with the probabilistic tradition. We found a product update mechanism based on a principled distinction between prior world probability, occurrence probability, and observation probability. This provides a 'smooth' extension of the discrete update rules of earlier chapters, letting probabilities gently incorporate new information. Moreover, our mechanism has a complete dynamic epistemic probabilistic logic that can handle update with formulas of arbitrary syntactic complexity. Thus dynamic logic and probability are compatible, and there may be interesting flow of ideas across. Of course, the real task is now to extend the bridge-head.

8.8 Further directions and open problems

Plausibility versus probability One obvious issue is how the plausibility models of Chapter 7 relate to a probabilistic approach. The style of thinking is different, in that most plausible worlds may ignore the cumulative probabilistic weight of the less plausible ones. A logical difference is that plausibility logics validate conjunction of beliefs: $(B\varphi \wedge B\psi) \rightarrow B(\varphi \wedge \psi)$,

---

[185] Cf. also Grunwald & Halpern 2003 on related forms of update, including Jeffrey Update.

[186] Technically, we have a pair (Φ, P) of a set of sentences partitioning the space and a probability distribution P over Φ. The Jeffrey Update of a prior P^{old} with the new information is then $P^{new}(s) = P^{old}(s \mid \varphi) \bullet P(\varphi)$. The new signal overrules prior information about the sentences in Φ, just as we had with belief revision policies like $\Uparrow P$, $\uparrow P$ in Chapter 7. For our Dark Room, Jeffrey Update makes the new probability of the object being dark $3/4$, and of its being light $1/4$. Thus we set new values for partition cells, but within these, relative probabilities of worlds remain the same.

while probabilistic approaches do not, since the intersection $\varphi\wedge\psi$ may have lost probability mass below some threshold. This issue has been studied for belief revision with graded modalities (Spohn 1988, Aucher 2004). Can we find systematic transformations?

Surprises Our update rule treats zero denominators as a nuisance. But zero probability events represent a real phenomenon of true *surprises*. These have been studied in Aucher 2005 using infinitesimal numbers. An alternative is the use of conditional probability spaces in Baltag & Smets 2007 to represent surprise events and their epistemic effects.

Dynamifying probabilistic reasoning Our dynamic logic exemplifies our general program of dynamifying existing systems. In probabilistic practice, however, our system may be too baroque, and well-chosen fragments are needed. [187] Probability theory also has challenges beyond our approach. One is the fundamental notion of *expected value*, a weighted utility over possible outcomes. This makes general sense for agency in the 'entanglement' of preference and belief in Chapter 9. To properly incorporate expected value, we need an extension of our dynamic logic with recursion axioms for new expected values after information has come in. [188] Dealing with these steps may eventually also involve temporal extensions of dynamic epistemic logics that can refer to the past (Chapter 11).

Philosophy of science In the philosophy of science, there is a flourishing literature on separate probabilistic update rules for popular scenarios, such as Sleeping Beauty and its ilk. It would be of interest to see if our logics can help systematize the area.

Postulates and Dutch Book arguments Laws of reasoning with probability are often justified by general postulates (cf. our discussion of postulational approaches in Chapters 3, 7). The most famous format are *Dutch Book Arguments* showing how the specific axioms of the probability calculus are the only ones that are fail-safe in multi-agent betting scenarios. Can we do similar analyses for the principles of our dynamic update logics?

8.9 Literature

There is a large literature linking probability, conditionals, and belief revision, for which the reader can consult standard sources. Kooi 2003 first merged probability with public

---

[187] Irma Cornelisse (p.c.) has proposed an interesting subsystem closer to *PAL* with upgrade actions $\Uparrow P, r$ that reset a proposition P toward some new probability r.

[188] The linguistic analysis of questions in van Rooij 2003 and the game-theoretic one of Feinberg 2007 compare expected values in an old model and a new one after new information is received.

announcement logic *PAL*. Van Benthem 2003 extended this to public event models with occurrence probabilities. Van Benthem, Gerbrandy & Kooi 2009 has the system of this chapter. Aucher 2004 gives another analysis of epistemic probabilistic update including surprise events, and Baltag & Smets 2007 one more, using conditional probability spaces (Popper measures). Sack 2009 extends probabilistic *DEL* from finite models to infinite ones, using mathematical notions from probability theory. Grunwald & Halpern 2003 make proposals related to ours, though in another logical framework. Halpern 2003 is probably the major current source on logical approaches to reasoning with uncertainty.