# The Logic of Conditional Doxastic Actions: A theory of dynamic multi-agent belief revision

Alexandru Baltag

Computing Laboratory, Oxford University. Alexandru.Baltag@comlab.ox.ac.uk

#### Sonja Smets

Center for Logic and Philosophy of Science, Vrije Universiteit Brussel. sonsmets@vub.ac.be

ABSTRACT. We present a logic of conditional doxastic actions, obtained by incorporating ideas from belief revision theory into the usual dynamic logic of epistemic actions. We do this by extending to actions the setting of epistemic plausibility models, developed in Baltag and Smets (2006) for representing (static) conditional beliefs. We introduce a natural extension of the notion of update product from Baltag and Moss (2004) to plausibility models.

### 1 Introduction

In this paper, we extend the semantic setting proposed in Baltag and Smets (2006) for "static" multi-agent belief revision to define a notion of *(multi-agent) belief update with an action*, i.e. a dynamic notion of belief revision. This improves on the work in Baltag et al. (1998) and Baltag and Moss (2004), by incorporating ideas from belief revision into dynamic-epistemic logic.

In Baltag and Smets (2006), we proposed two equivalent semantic settings for "static" belief revision, and proved them to be equivalent with each other and with a multi-agent epistemic version of the AGM belief revision theory: conditional doxastic models and epistemic plausibility models. We argued extensively that these settings provided the "right" qualitative semantics for multi-agent belief revision, forming the basis of a conditional doxastic logic (CDL, for short), that captured the main "laws" of hypothetical beliefs. We went beyond static revision, using CDL to explore a restricted notion of "dynamic" belief revision, by modeling and axiomatizing multi-agent belief updates induced by public and private announcements.

In this paper, we go further and describe a full dynamic logic of *conditional doxastic actions*, which subsumes all the other approaches known to us that combine dynamic-epistemic logic with belief revision. We do this by putting together the ideas from Baltag et al. (1998), Baltag and Moss (2004) on the logic of epistemic programs with the ideas from Baltag and Smets (2006). In particular, we adopt the view of "actions" as *having a similar underlying doxastic/epistemic structure* as the "states": thus, we use the same kind of structures (plausibility frames) to model both actions and states. We also adopt the fundamental idea of "update product" from Baltag and Moss (2004), extending it naturally from epistemic Kripke models to plausibility models.

It is important to note that our approach differs from the recent semantical literature on the topic of dynamic belief revision, e.g. Aucher (2003), van Ditmarsch (2005), Segerberg (1998), Segerberg (1999), in the following sense: most Kripke-style models proposed for multiagent belief revision are based on *specific mechanisms that rely on quantitative notions*, such as "degrees of belief", plausibility functions, graded models or probabilistic measures of belief.<sup>1</sup> However, classical AGM belief revision theory, as initiated in Alchourron, Gardenfors and Makinson (1985) and developed e.g. in Gardenfors (1988), is a *qualitative* theory, based on simple postulates concerning a basic operation (revision), of great generality and simplicity. Our approach retains this qualitative flavor of the classical AGM theory.

In this sense, our approach is closer to the one in van Benthem (2005) and van Benthem et al. (2005). Though still based on ("quantitative") models involving plausibility relations, the approach in van Benthem (2005) abstracts away from the details of modeling when considering the associated modal *logic*, which (is *not* based on any "graded belief" operator, as in e.g. Aucher (2003), van Ditmarsch (2005), but) is a simple language of conditional beliefs and update modalities, virtually identical to ours (for public announcements). As a result, the main "reduction axiom" in van Benthem (2005) (which computes, in the style of the Action-Knowledge Axiom in Baltag et al. (1998), the beliefs after a public announcement in terms of the initial conditional beliefs) can be recovered as a special case of our main reduction axiom for conditional beliefs after an action. In a sense, our approach here is simply to go one step further, and abstract away (from the specific details of a particular quantitative implementation of belief revision operators) on the semantic side as well. This leads to a perfect match between the syntax (based on conditional beliefs) and the semantics (in terms of conditional doxastic models), giving our logic a broader, more general scope of application and a greater transparency. In its turn, this greatly facilitates the move to the more general context of *arbitrary* epistemic/doxastic actions.

Our concepts of conditional belief and of CDM can also be seen in the context of the wide logical-philosophical literature on notions of *conditional*, see e.g. Adams (1965), Stalnaker (1968), Ramsey (1931), Lewis (1973), Bennett (2003). One can of course look at our conditional belief operators as non-classical (and non-monotonic!) implications. Our approach can thus be compared with other attempts of using doxastic conditionals to deal with belief revision, see e.g. Gardenfors (1986), Grove (1988), Rott (1989), Fuhrmann and Levi (1994), Ryan and Schobbens (1997), Halpern (2003).

# 2 Preliminaries: Epistemic Plausibility Models and Conditional Doxastic Models

In this section, we review some basic notions and results from Baltag and Smets (2006). **Plausibility Frames.** An *epistemic plausibility frame* is a structure  $\mathbf{S} = (S, \sim_a, \leq_a)_{a \in \mathcal{A}}$ , consisting of a set S, endowed with a family of equivalence relations  $\sim_a$ , called *epistemic indistinguishability relations*, and a family of "well-preorders"  $\leq_a$ , called *plausibility relations*. Here, a "well-preorder" is just a preorder<sup>2</sup> such that every non-empty subset has minimal

<sup>&</sup>lt;sup>1</sup>One could argue that the degrees of belief are given by a plausibility order relation, so by a qualitative, order-theoretic notion, but in fact the way belief revision or update are defined makes an essential use of the "arithmetic" of these (finite or transfinite) degrees, e.g. in Spohn (1988) and Aucher (2003); hence, the quantitative flavor.

 $<sup>^{2}\</sup>mathrm{i.e.}$  a reflexive and transitive relation.

elements.<sup>3</sup> Using the notation  $Min \leq T := \{t \in T : t \leq t' \text{ for all } t' \in T\}$  for the set of minimal elements of T, the last condition says that: for every  $T \subseteq S$ , if  $T \neq \emptyset$  then  $Min < T \neq \emptyset$ .

Plausibility frames for only one agent and without the epistemic relations have been used as models for conditionals and belief revision in Grove (1988), Gardenfors (1986), Gardenfors (1988), Segerberg (1998) etc. Observe that the conditions on the preorder  $\leq_a$  are (equivalent to) Grove's conditions for the (relational version of) his models in Grove (1988). The standard formulation of Grove models (in terms of a "system of spheres", weakening the similar notion in Lewis (1973)) was proved in Grove (1988) to be equivalent to the above relational formulation.<sup>4</sup>

Given a plausibility frame **S**, an **S**-proposition is any subset  $P \subseteq S$ . We say that the state s satisfies the proposition P if  $s \in P$ . Observe that a plausibility frame is just a special case of a Kripke frame. So, as is standard for Kripke frames, we can define an epistemic plausibility model to be an epistemic plausibility frame **S** together with a valuation map  $\| \bullet \| : \Phi \to \mathcal{P}(S)$ , mapping every element of a given set  $\Phi$  of "atomic sentences" into S-propositions.

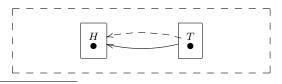
Notation: strict plausibility, doxastic indistinguishability. As with any preorder, the ("non-strict") plausibility relation  $\leq_a$  above has a "strict" (i.e. asymmetric) version  $\leq_a$ , as well as a corresponding equivalence relation  $\simeq_a$ , called "doxastic indistinguishability":

$$s <_a t$$
 iff  $s \leq_a t$  and  $t \not\leq_a s$ 

$$s \simeq_a t$$
 iff  $s \leq_a t$  and  $t \leq_a s$ 

**Interpretation**. The elements of **S** will be interpreted as the *possible states* of a system (or "possible worlds"). The atomic sentences  $p \in \Phi$  represent "ontic" (non-doxastic) facts about the world, that might hold or not in a given state, while the valuation tells us which facts hold at which worlds. The equivalence relations  $\sim_a$  capture the agent's knowledge about the actual state of the system (intuitively based on the agent's (partial) observations of this state): two states s, t are indistinguishable for agent a if  $s \sim_a t$ . In other words, when the actual state of the system is s, then agent a knows only the state's equivalence class  $s(a) := \{t \in S : s \sim_a t\}$ . Finally, the plausibility relations  $\leq_a$  capture the agent's conditional beliefs about (virtual) states of the system : given the information that some possible state of the system is either s or t, agent a will believe the state to be s iff  $s <_a t$ ; will believe the state to be t iff  $t <_a s$ ; otherwise (if  $s \simeq_a t$ ), the agent will consider the two alternatives as equally plausible.

**Example 1**. The father informs the two children (Alice and Bob) that he has put a coin lying face up on the table in front of them. At first, the face is covered (so the children cannot see it). Based on previous experience, (it is common knowledge that) the children believe that the upper face is (very likely to be) Heads: say, they know that the father has a strong preference for Heads. And in fact, they're right: the coin lies Heads up. Next, the father shows the face of the coin to Alice, in the presence of Bob but in such a way that Bob cannot see the face (though of course he can see that Alice sees the face). The plausibility model **S** for this situation is:



<sup>&</sup>lt;sup>3</sup>Observe that the existence of minimal elements implies, by itself, that the relation  $\leq_a$  is both *reflexive* (i.e.  $s \leq_a s$  for all  $s \in S$ ) and *connected* (i.e. either  $s \leq_a t$  or  $t \leq_a s$ , for all  $s, t \in S$ ). Note also that, when the set S is *finite*, a well-preorder is nothing but a connected preorder.

<sup>&</sup>lt;sup>4</sup>A more concrete example of plausibility frames was given in Spohn (1988), in terms of ordinal plausibility maps assigning ordinals d(s) ("the degree of plausibility" of s) to each state  $s \in S$ . In our epistemic multi-agent context, this would endow each agent a with an ordinal plausibility map  $d_a : S \to Ord$ .

Here, we left the father out of the picture (since he only plays the role of God or Nature, not the role of an uncertain agent). The node on the left, labeled with H, represents the actual state of the system (in which the coin lies Heads up), while the node on the right represents the other possible state (in which the coin is Tails up). We use continuous arrows to encode Alice's beliefs and use continuous squares to encode her knowledge, while using dashed arrows and dashed squares for Bob. More precisely: the squares represent the agents' information cells, i.e. the equivalence classes  $s(a) := \{t \in S : s \sim_a t\}$  of indistinguishable states (for each agent a). Observe that Alice's information cells (the continuous squares) are singletons: in every case, she knows the state of the system; Bob's information cell is one big dashed square comprising both states: he doesn't know which state is the real one, so he cannot distinguish between them. The arrows represent the plausibility relations for the two agents; since these are always reflexive, we choose to skip all the loops for convenience. Both arrows point to the node on the left: a priori (i.e. before making any observation of the real state), both agents believe that it is likely that the coin lies Heads up.

**Conditional Doxastic Frames**. A plausibility frame is in fact nothing but a way to encode all the agents' possible *conditional beliefs*. To see this, consider the following equivalent notion, introduced in Baltag and Smets (2006): A *conditional doxastic frame (CD-frame, for short)*  $\mathbf{S} = (S, \{\mathbf{\bullet}_a^P\}_{a \in \mathcal{A}, P \subseteq S})$  consists of a set of states S, together with a family of *conditional (doxastic) appearance* maps, one for each agent a and each possible condition  $P \subseteq S$ . These are required to satisfy the following conditions:

- 1. if  $s \in P$  then  $s_a^P \neq \emptyset$ ;
- 2. if  $P \cap s_a^Q \neq \emptyset$  then  $s_a^P \neq \emptyset$ ;
- 3. if  $t \in s_a^P$  then  $s_a^Q = t_a^Q$ ;
- 4.  $s_a^P \subseteq P;$
- 5.  $s_a^{P \cap Q} = s_a^P \cap Q$ , if  $s_a^P \cap Q \neq \emptyset$ .

A conditional doxastic model (CDM, for short) is a Kripke model whose underlying frame is a CD-frame. The conditional appearance  $s_a^P$  captures the way a state s appears to an agent a, given some additional (plausible, but not necessarily truthful) information P. More precisely: whenever s is the current state of the world, then after receiving new information P, agent a will come to believe that any of the states  $s' \in s_a^P$  might have been the current state of the world (as it was before receiving information P).

Using conditional doxastic appearance, the knowledge s(a) possessed by agent a about state s (i.e. the epistemic appearance of s) can be defined as the union of all conditional doxastic appearances. In other words, something is known iff it is believed in any conditions:  $s(a) := \bigcup_{Q \subseteq S} s_a^Q$ . Using this, we can see that the first condition above in the definition of conditional doxastic frames captures the truthfulness of knowledge. Condition 2 states the success of belief revision, when consistent with knowledge: if something is not known to be false, then it can be consistently entertained as a hypothesis. Condition 3 expresses full introspection of (conditional) beliefs: agents know their own conditional beliefs, so they cannot revise their beliefs about them. Condition 4 says hypotheses are hypothetically believed: when making a hypothesis, that hypothesis is taken to be true. Condition 5 describes minimality of revision: when faced with new information Q, agents keep as much as possible of their previous (conditional) beliefs  $s_a^P$ .

To recover the usual, unconditional beliefs, we put  $s_a := s_a^S$ . In other words: unconditional ("default") beliefs are beliefs conditionalized by trivially true conditions.

For any agent a and any S-proposition P, we can define a *conditional belief operator*  $B_a^P$ :  $\mathcal{P}(S) \to \mathcal{P}(S)$  S-propositions, as the Galois dual of conditional doxastic appearance:

$$B_a^P Q := \{ s \in S : s_a^P \subseteq Q \}$$

We read this as saying that agent a believes Q given P. More precisely, this says that: if the agent would learn P, then (after learning) he would come to believe that Q was the case in the current state (before the learning). The usual (unconditional) belief operator can be obtained as a special case, by conditionalizing with the trivially true proposition  $S: B_aQ := B_a^SQ$ . The knowledge operator can similarly be defined as the Galois dual of epistemic appearance:

$$K_a P := \{ s \in S : s(a) \subseteq P \}.$$

As a consequence of the above postulates, we have the following:

$$K_a P = \bigcap_{Q \subseteq S} B_a^Q P = B_a^{\neg P} \emptyset = B_a^{\neg P} P$$

Equivalence between plausibility models and conditional doxastic models: Any plausibility model gives rise to a CDM, in a canonical way, by putting

$$s_a^P := Min_{\leq_a} \left\{ t \in P : t \sim_a s \right\}$$

where  $Min_{\leq_a} T := \{t \in T : t \leq_a t' \text{ for all } t' \in T\}$  is the set of  $\leq_a$ -minimal elements in T. We call this *the canonical CDM* associated to the plausibility model. The converse is given by a: **Representation Theorem.**<sup>5</sup> Every *CDM* is the canonical *CDM* of some plausibility model.

The advantage of the CDM formulation is that it leads naturally to a complete axiomatization of a *logic of conditional beliefs*, which was introduced in Baltag and Smets (2006) under the name of "Conditional Doxastic Logic"  $(CDL)^6$ : the semantical postulates that define CDM's can be immediately converted into modal axioms governing conditional belief.

Conditional Doxastic Logic (CDL). The syntax of CDL (without common knowledge and common belief operators)<sup>7</sup> is:

$$\varphi := p \mid \neg \varphi \mid \varphi \land \varphi \mid B_a^{\varphi} \varphi$$

while the semantics is given by the obvious compositional clauses for the interpretation map  $|| \bullet ||_{\mathbf{S}} : CDL \to \mathcal{P}(S)$  in a CDM (and so, in particular, in a plausibility model) **S**. In this logic, the *knowledge modality* can be defined as an abbreviation, putting  $K_a \phi := B_a^{\neg \phi} \bot$  (where  $\bot = p \land \neg p$  is an inconsistent sentence), or equivalently  $K_a \phi := B_a^{\neg \phi} \phi$ . This way of defining knowledge in terms of doxastic conditionals can be traced back to (Stalnaker 1968). It is easy

<sup>&</sup>lt;sup>5</sup>This result can be seen as an analogue in our semantic context of Gardenfors' representation theorem in Gardenfors (1986), representing the AGM revision operator in terms of the minimal valuations for some total preorder on valuations.

 $<sup>^{6}</sup>CDL$  is an extension of the well-known logic KL of "knowledge and belief"; see e.g. Meyer and van der Hoek (1995), pg. 94, for a complete proof system for KL.

<sup>&</sup>lt;sup>7</sup>In Baltag and Smets (2006), we present and axiomatize a logic that includes conditional common knowledge and conditional common true belief.

to see that this agrees semantically with the previous definition of the semantic knowledge operator (as the Galois dual of epistemic appearance):  $||K_a\phi||_{\mathbf{S}} = K_a||\phi||_{\mathbf{S}}$ .

**Doxastic Propositions.** A doxastic proposition is a map  $\mathbf{P}$  assigning to each plausibility model (or conditional doxastic model)  $\mathbf{S}$  some S-proposition, i.e. a set of states  $\mathbf{P}_{\mathbf{S}} \subseteq$ S. The interpretation map for the logic CDL can thus be thought of as associating to each sentence  $\varphi$  of CDL a doxastic proposition  $||\varphi||$ . We denote by Prop the family of all doxastic propositions. All the above operators (Boolean operators as well as doxastic and epistemic modalities) on S-propositions induce corresponding operators on doxastic propositions, defined pointwise: e.g. for any doxastic proposition  $\mathbf{P}$ , one can define the proposition  $K_a\mathbf{P}$ , by putting  $(K_a\mathbf{P})_{\mathbf{S}} := K_a\mathbf{P}_{\mathbf{S}}$ , for all models  $\mathbf{S}$ .

**Theorem.** (Baltag and Smets 2006) A complete proof system for CDL can be obtained from any complete axiomatization of propositional logic by adding the following:

Necessitation Rule:

From  $\vdash \varphi$  infer  $\vdash B_a^{\psi} \varphi$ 

Normality:	$\vdash B^{\theta}_{a}(\varphi \to \psi) \to (B^{\theta}_{a}\varphi \to B^{\theta}_{a}\psi)$
Truthfulness of Knowledge:	$\vdash K_a \varphi \to \varphi$
Persistence of Knowledge:	$\vdash K_a \varphi \to B_a^{\psi} \varphi$
Full Introspection:	$\vdash B^{\psi}_{a}\varphi \to K_{a}B^{\psi}_{a}\varphi$
	$\vdash \neg B_a^{\psi} \varphi \to K_a \neg B_a^{\psi} \varphi$
Hypotheses are (hypothetically) accepted:	$\vdash B^{\varphi}_{a}\varphi$
Minimality of revision:	$\vdash \neg B_a^{\varphi} \neg \psi \to (B_a^{\varphi \land \psi} \theta \leftrightarrow B_a^{\varphi} (\psi \to \theta))$

### **3** Action Plausibility Models and Product Update

The belief revision encoded in the models above is of a *static*, purely *hypothetical*, nature. Indeed, the revision operators cannot alter the models in any way: all the possibilities are already there, so both the unconditional and the revised, conditional beliefs *refer to the same world and the same moment in time*. In contrast, a *belief update* in our sense is a dynamic form of belief revision, meant to capture the actual change of beliefs induced by learning (or by other forms of epistemic/doxastic actions).<sup>8</sup> As already noticed before, in e.g. Gerbrandy (1999) and Baltag et al. (1998), the original model does not usually include enough states to capture all the epistemic possibilities that arise in this way. So we now introduce "revisions" that change the original plausibility model.

To do this, we adapt an idea coming from Baltag et al. (1998) and developed in full formal detail in Baltag and Moss (2004). There, the idea was that *epistemic actions should be modeled* in essentially the same way as epistemic states, and this common setting was taken to be given by *epistemic Kripke models*. Since we now enriched our models for states to deal with conditional beliefs, it is natural to follow Baltag and Moss (2004) into extending the similarity between actions and states to this conditional setting, thus obtaining action plausibility models.

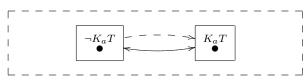
An action plausibility model is just an epistemic plausibility frame  $\Sigma = (\Sigma, \sim_a, \leq_a)_{a \in \mathcal{A}}$ , together with a precondition map pre :  $\Sigma \to Prop$  associating to each element of  $\Sigma$  some doxastic proposition  $pre(\sigma)$ . As in Baltag and Moss (2004), we call the elements of  $\Sigma$  (basic) epistemic actions, and we call  $pre(\sigma)$  the precondition of action  $\sigma$ .

<sup>&</sup>lt;sup>8</sup>But observe the difference between our notion of belief update (originating in dynamic-epistemic logic) and the similar (and vaguer) notion in (Katsuno and Mendelzon 1992).

Interpretation: Beliefs about Changes encode Changes of Beliefs. The name "doxastic actions" might be a bit misleading; the elements of a plausibility model are not intended to represent "real" actions in all their complexity, but only the *doxastic changes* induced by these actions: each of the nodes of the graph represents a *specific kind of change of beliefs (of all the agents)*. As in Baltag and Moss (2004), we only deal here with pure "belief changes", i.e. actions that do not change the "ontic" facts of the world, but only the agents' beliefs<sup>9</sup>. Moreover, we think of these as *deterministic* changes: there is at most one output of applying an action to a state.<sup>10</sup> Intuitively, the precondition defines the *domain of applicability* of  $\sigma$ : this action can be executed on a state *s* iff *s* satisfies its precondition. The plausibility pre-orderings  $\leq_a$  give the agent's beliefs about *changes*, that *encode changes of beliefs*. In this sense, we use such "beliefs about actions" as a way to represent doxastic changes: the information about how the agent changes her beliefs is captured by our action plausibility relations. So we read  $\sigma <_a \sigma'$  as saying that: if agent *a* is given the information that some (virtual) action is *either*  $\sigma$  or  $\sigma'$  (*without being able to know which*), then she believes that  $\sigma$  is the one actually happening.

**Example 2:** Successful Lying. The action of "public successful lying" can be described as follows: given a doxastic proposition  $\mathbf{P}$ , the model consists of two actions  $Lie_a \mathbf{P}$  and  $True_a \mathbf{P}$ , the first being the action in which agent *a* publicly lies that (she knows)  $\mathbf{P}$  (while in fact she doesn't know it), and the second being the action in which *a* makes a truthful public announcement that (she knows)  $\mathbf{P}$ . The preconditions are  $pre(Lie_a \mathbf{P}) = \neg K_a \mathbf{P}$ and  $pre(True_a \mathbf{P}) = K_a \mathbf{P}$ . Agent *a*'s equivalence relation is simply the *identity*: she knows whether she's lying or not. The other agents' equivalence relation is the *total relation*: they cannot know if *a* is lying or not. Let us assume that *a*'s plausibility preorder is also the *total relation*: this would express the fact that agent *a* is *not decided to always lie*; *a priori*, she considers equally plausible that, in any arbitrarily given situation, she will lie or not. But the plausibility relations should reflect the fact that we are modeling a "typically successful lying": by default, in such an action, the hearer trusts the speaker, so he is inclined to believe the lie. Hence, the relation for any hearer  $b \neq a$  should *make it more plausible to him that a is telling the truth rather than lying*:  $True_a \mathbf{P} <_b Lie_a \mathbf{P}$ .

As a *specific example*, consider the scenario in Example 1, and assume now that Alice tells Bob (after seeing that the coin was lying Heads up): "I saw the face, so now I know: The coin is lying Tails up". Assume that Bob trusts Alice completely, so he believes that she is telling the truth. We can model this action using the following action model  $\Sigma$ :



This model has two actions: the one on the left is the real action that is taking place (in which Alice's sentence is a *lie*: in fact, *she doesn't know* the coin is Tails up), while the one on the right is the other possible action (in which Alice is telling the truth: she *does know* the coin is Tails up). We labeled this node with their preconditions,  $\neg K_a T$  for the lying action and  $K_a T$  for the truth-telling action. In each case, Alice *knows* what action she is doing, so her

<sup>&</sup>lt;sup>9</sup>We stress this is a minor restriction, and it is very easy to extend this setting to "ontic" actions. The only reason we stick with this restriction is that it simplifies the definitions, and that it is general enough to apply to all the actions we are interested here, and in particular to all *communication actions*.

 $<sup>^{10}</sup>$ As in Baltag and Moss (2004), we will be able to represent non-deterministic actions as sums (unions) of deterministic ones.

information cells (the continuous squares) are singletons; while Bob is uncertain, so the dashed square includes both actions. As before, we use arrows for plausibility relations, skipping all the loops. As assumed above, Alice is not decided to always lie about this; so, *a priori*, she finds her lying in any such given case to be equally plausible as her telling the truth: this is reflected by the fact that the continuous arrow is bidirectional. In contrast, (Bob's) dashed arrow points only to the node on the right: he really believes Alice!

The Product Update of Two Plausibility Models. We are ready now to define the updated (state) plausibility model, representing the way some action, from an action plausibility model  $\Sigma = (\Sigma, \sim_a, \leq_a, pre)_{a \in \mathcal{A}}$ , will act on an input-state, from an initially given (state) plausibility model  $\mathbf{S} = (S, \sim_a, \leq_a, pre)_{a \in \mathcal{A}}$ . We denote this updated model by  $\mathbf{S} \otimes \Sigma$ , and we call it the update product of the two models. Its states are elements  $(s, \sigma)$  of the Cartesian product  $S \times \Sigma$ . More specifically, the set of states of  $\mathbf{S} \otimes \Sigma$  is

$$S \otimes \Sigma := \{ (s, \sigma) : s \in pre(\sigma)_{\mathbf{S}} \}$$

The valuation is given by the original input-state model: for all  $(s, \sigma) \in S \otimes \Sigma$ , we put  $(s, \sigma) \models p$  iff  $s \models p$ . As epistemic uncertainty relations, we take the *product* of the two epistemic uncertainty relations<sup>11</sup>: for  $(s, \sigma), (s', \sigma') \in S \otimes \Sigma$ ,

$$(s,\sigma) \sim_a (s',\sigma')$$
 iff  $\sigma \sim_a \sigma', s \sim_a s'$ 

Finally, we define the plausibility relation as the *anti-lexicographic preorder relation on pairs*  $(s, \sigma)$ , i.e.:

 $(s,\sigma) \leq_a (s',\sigma')$  iff either  $\sigma <_a \sigma'$  or  $\sigma \simeq_a \sigma', s \leq_a s'$ .

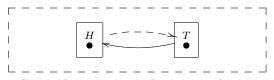
To explain this definition, recall first that we only deal with *pure "belief* Interpretation. changes", not affecting the "facts": this explains our "conservative" valuation. Second, the product construction on the epistemic indistinguishability relation  $\sim_a$  is the same as in Baltag and Moss (2004): if two indistinguishable actions are successfully applied to two indistinguishable input-states, then their output-states are indistinguishable. Third, the anti-lexicographic preorder gives "priority" to the *action* plausibility relation; this is not an arbitrary choice, but is motivated by our above-mentioned interpretation of "actions" as specific types of *belief changes*. The action plausibility relation captures what agents *really believe is going on at the moment*; while the input-state plausibility relations only capture past beliefs. The doxastic action is the one that "changes" the initial doxastic state, and not vice-versa. If the "believed action"  $\alpha$ requires the agent to revise some past beliefs, then so be it: this is the whole point of believing  $\alpha$ , namely to use it to revise or update one's past beliefs. For example, in a successful lying, the action plausibility relation makes the hearer believe that the speaker is telling the truth; so she'll accept this message (unless contradicted by her knowledge), and change her past beliefs appropriately: this is what makes the lying "successful". Giving priority to action plausibility does not in any way mean that the agent's belief in actions is "stronger" than her belief in states; it just captures the fact that, at the time of updating with a given action, the belief about the action is what is actual, what is present, is the current belief about what is going on, while the beliefs about the input-states are in the past.<sup>12</sup> The belief update induced by a given action is nothing but an update with the (presently) believed action.

 $<sup>^{11}</sup>$ Observe that this is precisely the uncertainty relation of the epistemic update product, as defined in Baltag and Moss (2004).

<sup>&</sup>lt;sup>12</sup>Of course, at a later moment, the above-mentioned belief about action (now belonging to the past) might be itself revised. But this is another, future update.

In other words, the anti-lexicographic product update reflects our Motto above: beliefs about changes (as formalized in the action plausibility relations) are nothing but ways to encode changes of belief (i.e. ways to change the original plausibility order on states). This simply expresses our particular interpretation of the (strong) plausibility ordering on actions, and is thus a matter of convention: we decided to introduce the order on actions to encode corresponding changes of order on states. The product update is a consequence of this convention: it just says that a strong plausibility order  $\alpha <_a \beta$  on actions corresponds indeed to a change of ordering, (from whatever the ordering was) between the original input-states s, t, to the order  $(s, \alpha) <_a (t, \beta)$  between output-states; while equally plausible actions  $\alpha \simeq_a \beta$  will leave the initial ordering unchanged:  $(s, \alpha) \leq_a (t, \beta)$  iff  $s \leq_a t$ . So the product update is just a formalization of our interpretation of action plausibility models, and thus it doesn't impose any further limitation to our setting.

**Example 3**: By computing the update product of the plausibility model **S** in Example 1 with the action model  $\Sigma$  in Example 2, we obtain the following plausibility model:



This correctly describes the effect of an action of "successful lying": Bob's plausible ordering is reversed, since he believes Alice, and so now he believes that the coin is lying face up. In contrast, Alice's initial plausibility relation is unaffected (since the two actions were equally plausible, i.e. doxastically equivalent, for her); so, she should keep her *a priori* belief that the coin is Heads up; of course, in this case the last point is not so relevant, since Alice *knows* the state of the coin (as witnessed by the fact that the continuous squares consist of single states). **Proposition**. The update product of a state plausibility model and an action plausibility model is a state plausibility model.

### 4 Product Update in CDM Form

As for the "static" conditional doxastic logic, the axioms of "dynamic" belief revision logic can be easily derived if we first work out the CDM version of the above update product. We do this from scratch, by first introducing a "dynamic" version of the notion of CDM, equivalent to the above concept of action plausibility model:

A conditional doxastic action model (CDAM, for short)  $\Sigma$  is just a conditional doxastic frame  $(\Sigma, \{\bullet_a^{\Pi}\}_{a \in \mathcal{A}, \Pi \subseteq \Sigma})$ , together with a precondition map  $pre : \Sigma \to Prop$  as above. A set of actions  $\Pi \subseteq \Sigma$  can be interpreted as partial information about some real (basic) action  $\sigma \in \Pi$ , or equivalently, as a non-deterministic action (in which one of the actions  $\sigma \in \Pi$  happens, but we are not told which). The conditional appearance  $\sigma_a^{\Pi}$  captures the way action  $\sigma$  appears to agent a, given additional (plausible, but not necessarily truthful) information  $\Pi$  about this action. This means that, in normal circumstances, if after  $\sigma$  happens the agent is told that (one of the actions in)  $\Pi$  has happened, then the agent will believe that in fact (one of the basic actions in)  $\sigma_a^{\Pi}$  has happened. As before, any action plausibility model induces in a canonical way a CDAM, and conversely any CDAM can be represented as the canonical CDAM of some action plausibility model.

**Example: Lying, revisited.** In the successful lying example, if we convert the plausibility model into its canonical CDM, we obtain e.g. that  $(Lie_a \mathbf{P})_b^Q = \{True_a \mathbf{P}\}$  for  $b \neq a$  and

 $Q \not\subseteq \{Lie_a \mathbf{P}\}\)$ . So this lying is indeed generally "successful": no matter what other information is given to b, if it is consistent with a telling the truth, then b believes that a tells the truth. The only case in which the appearance of this action to b is different is when  $Q \subseteq \{Lie_a \mathbf{P}\}\)$ , in which case  $(Lie_a \mathbf{P})_b^Q = Q$ , and in particular,  $(Lie_a \mathbf{P})_b^{\{Lie_a\}} = \{Lie_a \mathbf{P}\}\)$ : so the hearer can discover the lying only if given information that excludes all other possible actions.

Independence of Action's Appearance from Prior Beliefs. The above description assumes that the agent's beliefs about the action are independent of his beliefs about the state: indeed,  $\sigma_a^{\Pi}$  contains no information about, or reference to, the current state's doxastic appearance  $s_a$  to the agent, so it is assumed that this does not influence in any way the appearance of the action. This assumption embodies a certain interpretation of our "appearance" maps: we take an action's appearance to simply denote the action itself, as it appears to the agent. In other words: for the agent, the appearance is the action, pure and simple. When the action  $\sigma$ happens (say, in an unconditional context), it *really* appears to the agent as if (the apparent, un-conditionalized action)  $\sigma_a := \sigma_a^{\Sigma}$  happens. If the agent makes the additional hypothesis that one of the actions in  $\Pi$  happens, then it appears to him that  $\sigma_a^{\Pi}$  is happening. The action's appearance is simply taken here as a brute, new fact: the agent really believes this apparent action is happening (otherwise this would not really be the appearance of this action). This belief cannot be revised at the same time that it is being held: any revision of the action's appearance can only happen in the future. But for the moment, this appearance correctly reflects what the agent thinks to be happening. In contrast, his prior beliefs about the state are just that: *prior* beliefs. They may be subject to revision at this very moment, due to current action (or, more precisely, due to its appearance): indeed, the (apparent) action is the one that *induces* the revision (or update) of the static belief. In a certain sense, the action, as it appears, is the belief update: the apparent action simply encodes the way the agent is compelled to update his prior beliefs. Hence, the action's appearance cannot, by definition, be dependent, or be influenced, by these prior beliefs: the action's appearance is a given, it is what it is, and the prior beliefs are the ones that may be changed by the apparent action, not vice-versa.

Taking the action's appearance as a correct description of the action as seen by the agent, the above independence (of this appearance from prior beliefs) can be understood as a *ratio-nality postulate*: agents should be prepared to revise their prior beliefs when faced with (what appears to them as) truthful new information. Rational agents are not fundamentalists: if given compelling evidence to the contrary (as encoded in the "apparent action"), they will not refuse it due to prior beliefs, but will change these prior beliefs to fit the new evidence. And it does not matter in the least that, at some later point, this "compelling evidence" might turn out to have been a belief (an "apparent action"), not a reality: when this will happen, rational agents might change their minds again. But for the moment, they have to accept the current action as it appears to them, and adjust their previous beliefs appropriately.

An Action's Contextual Appearance. In the context of belief revision, there is a subtle point to be made here: the above independence only refers to the agent's prior *beliefs*, but not to the *agent's knowledge*. No action's appearance can be assumed to be independent of prior knowledge: it might happen that the current state s is such that agent a knows that the believed action  $\sigma_a^{\Pi}$  cannot happen at this state. This is perfectly possible, even in states in which  $\sigma$  does happen, and even if the information  $\Pi$  is correct (i.e.  $\sigma \in \Pi$ ). In such a state, the agent cannot accept the default appearance  $\sigma_a^{\Pi}$ . Prior knowledge may thus influence the action's appearance.

**Example revisited**: "Successful" lying cannot always be successful! Indeed, if the original input-state s is such that an outsider b already knows that P is false, then lying cannot succeed.

In this context, the appearance of the action  $Lie_a \mathbf{P}$  to b is not its default appearance  $True_a \mathbf{P}$ : b cannot believe that a is telling the truth. Instead, the contextual appearance of this action at s is itself:  $(Lie_a \mathbf{P})_b = Lie_a \mathbf{P}$ . The hearer knows the speaker is lying.

So  $\sigma_a^{\Pi}$  should only be thought of as the action's *default appearance* (conditional on  $\Pi$ ) to the agent: in the absence of any other additional information (except for  $\Pi$ ), or whenever the agent's prior knowledge allows this, the agent a will believe that  $\sigma_a^{\Pi}$  has happened.

So how will this action appear in a context in which the default appearance is known to be impossible? We can answer this question by defining a *contextual appearance*  $\sigma_a^{s,\Pi}$  of action  $\sigma$ to agent a at state s, given  $\Pi$ . We can do this by strengthening our conditionalization: at a given state  $s \in S$ , an agent has already some information about the next action, namely that it cannot be inconsistent with his knowledge s(a) of the state. In other words, agent a knows that the action must belong to the set  $\Sigma_{s(a)} := \{\rho \in \Sigma : s(a) \cap pre(\rho) \ \mathbf{g} \neq \emptyset\} = \{\rho \in \Sigma :$  $s \not\models \mathbf{g} \ K_a \neg pre(\rho)\}$ . Putting this information together with the new information  $\Pi$ , we obtain the contextual appearance by *conditionalizing the agent's belief about the action with*  $\Sigma_{s(a)} \cap \Pi$ :

$$\sigma_a^{s,\Pi} := \sigma_a^{\Sigma_{s(a)} \cap \Pi} = \sigma_a^{\{\rho \in \Pi: s(a) \cap pre(\rho) \neq \emptyset\}}$$

This contextual appearance is the one that fully captures the agent's *actual belief about the* action  $\sigma$  in state s, whenever he is given information  $\Pi$ .

An Action's Effect: Deterministic change of state. As announced, we take the basic actions  $\sigma \in \Sigma$  to represent deterministic changes of states. In the following, we will always represent the output-state  $\sigma(s)$  of applying basic action  $\sigma$  to state  $s \in S$  by an ordered pair  $\sigma(s) := (s, \sigma)$ . So, for a given CDM **S** of possible input-states and a given CDAM of possible actions, the set of all possible output-states will be a subset of the Cartesian product  $S \times \Sigma$ . Thus, we could represent post-conditions, i.e. conditions restricting the possible output-states of some (unspecified) action acting on some (unspecified) input-state as subsets  $P \subseteq S \times \Sigma$  of the Cartesian product. Given a basic action  $\sigma \in \Sigma$  and a post-condition  $P \subseteq S \times \Sigma$ , we may denote the set of possible input-states of action  $\sigma$  ensuring post-condition P by:

$$\sigma^{-1}(P) = \{ s \in S : \sigma(s) \in P \} = \{ s \in S : (s, \sigma) \in P \}$$

**Post-conditional Contextual Appearance**. Sometimes, the additional information the agent may be given (or the hypothesis that he may entertain) refers, not directly to the range  $\Pi \subseteq \Sigma$  of possible actions currently happening, but to some *post-condition* P; i.e. the agent might be told that the current action will result in a state  $\sigma(s) = (s, \sigma)$  satisfying some post-condition  $P \subseteq S \times \Sigma$ . He should be able to conditionalize his belief about the current action with this information, in a given context. For this, we define the *contextual appearance*  $\sigma_a^{s,P}$  of action  $\sigma$  at state s, in the hypothesis (that the action will ensure postcondition) P, by putting:

$$\sigma_a^{s,P}:=\sigma_a^{\{\rho\in\Sigma:s(a)\cap\rho^{-1}(P)\neq\emptyset\}}$$

**Example:** Lying, again. Let again  $Lie_a \mathbf{P}$  be the action of successful lying by agent a, and suppose that  $\mathbf{P}$  denotes a *factual* ("ontic") statement (which happens to be false, and thus will remain false after lying). Even if in the original state, the hearer b did *not know* that  $\mathbf{P}$  was false (so that lying *was* successful, and its appearance to b was the default one  $True_a \mathbf{P}$ ), he may be given later that information, as a post-condition  $\neg \mathbf{P}$ . Then, the hearer discovers the lying: the post-conditional contextual appearance of lying (given  $\neg \mathbf{P}$ ) is... lying!

Belief Revision induced by an Action and a Postcondition. We want to calculate now the revision of an agent's beliefs (about an input-state s) induced by an action  $\sigma$  when given some post-condition  $P \subseteq S \times \Sigma$ . We denote this by  $s_a^{\sigma,P}$ . This captures the appearance of the input-state s to agent a, after action  $\sigma$  and after being given the information that P holds at the output-state. As explained already, the agent revises his prior beliefs not in accordance with the actual action, but in accordance to how this action appears to him. As we have seen, the appearance of action  $\sigma$  at state s when given post-condition P is  $\sigma_a^{s,P}$ . So the new information obtained post-factum about the original input-state s is that this state was capable of supporting (one of the actions in)  $\sigma_a^{s,P}$ , and moreover that it yielded an output-state satisfying post-condition P. In other words, the agents learns that the original state was in  $(\sigma_a^{s,P})^{-1}(P)$ . So he has to revise (conditionalize) his prior beliefs about s with this information, obtaining:

$$s_a^{\sigma,P} := s_a^{(\sigma_a^{s,P})^{-1}(P)}$$

**Product Update, in** *CDM* form. We give now a *CDM* equivalent of the above notion of product update: the *product update* of a conditional doxastic model **S** with a conditional doxastic action model  $\Sigma$  is a new conditional doxastic model  $\mathbf{S} \otimes \Sigma$ , whose states are elements  $\sigma(s) := (s, \sigma)$  of a subset  $S \otimes \Sigma$  of the Cartesian product  $S \times \Sigma$ . Note that we prefer here the functional notation  $\sigma(s)$ , instead of  $(s, \sigma)$ . As before, preconditions select the surviving states:

$$S \otimes \Sigma := \{ \sigma(s) : s \in pre(\sigma) | \mathbf{S} \}$$

For any hypothesis  $P \subseteq S \otimes \Sigma$  about the output-state, the conditional appearance (conditioned by P) of an output-state  $\sigma(s)$  to an agent a is given by

$$\sigma(s)_a^P := \sigma_a^{s,P}(s_a^{\sigma,P}) \cap P$$

In words: Agent a's updated belief (about the output-state of a basic action  $\sigma$  applied to an input-state s, when given condition P)  $\sigma(s)_a^P$  can be obtained by applying the action that is believed to happen (i.e. the appearance  $\sigma_a^{s,P}$  of  $\sigma$  to a at s, given post-condition P) to the agent's revised belief about the input-state  $s_a^{\sigma,P}$  (belief revised with the information provided by the apparent action  $\sigma_a^{s,P}$ ), then restricting to the given post-condition P. Finally (as for plausibility models), the valuation on output-states comes from the original states:

$$||p|| \mathbf{S}_{\otimes \Sigma} := \{ \sigma(s) \in S \otimes \Sigma : s \in ||p||_{\mathbf{S}} \}$$

**Proposition**. The two "product update" operations defined above agree: the canonical CDM associated to the (anti-lexicographic) product update of two plausibility models is the product update of their canonical CDM's.

#### 5 The Dynamic Logic

As in Baltag and Moss (2004), we consider a doxastic signature, i.e. a finite (fixed) plausibility frame (or, equivalently, a finite conditional doxastic frame)  $\Sigma$ , together with an ordered list without repetitions ( $\sigma_1, \ldots, \sigma_n$ ) of some of the elements of  $\Sigma$ . Each signature gives rise to a dynamic-doxastic logic  $CDL(\Sigma)$ , as in Baltag and Moss (2004): one defines by double recursion a set of sentences  $\varphi$  and a set of program terms  $\pi$ ; the basic programs are of the form  $\pi = \sigma \vec{\varphi} = \sigma \varphi_1 \ldots \varphi_n$ , where  $\sigma \in \Sigma$  and  $\varphi_i$  are sentences in our logic; program terms are generated from basic programs using non-deterministic sum (choice)  $\pi \cup \pi'$  and sequential composition  $\pi; \pi'$ . Sentences are built using the operators of CDL, and in addition a dynamic modality  $\langle \pi \rangle \varphi$ , taking program terms and sentences into other sentences. As in Baltag and Moss (2004), the conditional doxastic maps on the signature  $\Sigma$  induce in a natural way conditional doxastic maps on basic programs in  $CDL(\Sigma)$ : we put  $(\sigma \vec{\varphi})_a^{\Pi \vec{\varphi}} := \{\sigma' \vec{\varphi} : \sigma' \in \sigma_a^{\Pi}\}$ . The given listing can be used to assign syntactic preconditions for basic programs, by putting:  $pre(\sigma_i \vec{\varphi}) := \varphi_i$ , and  $pre(\sigma \vec{\varphi}) := \top$  (the trivially true sentence) if  $\sigma$  is not in the listing. Thus, the basic programs of the form  $\sigma \vec{\varphi}$  form a (finite) syntactic  $CDAM^{13} \Sigma \vec{\varphi}$ . Every given interpretation  $|| \bullet || : CDL(\Sigma) \to Prop$  of sentences as doxastic propositions will convert this syntactic model into a "real" (semantic) CDAM, called  $\Sigma || \vec{\varphi} ||$ .

To give the semantics, we define by induction two *interpretation maps*, one taking any sentence  $\varphi$  to a doxastic proposition  $||\varphi|| \in Prop$ , the second taking any program term  $\alpha$  to a (possibly non-deterministic) doxastic "program", i.e. a *set* of basic actions in some *CDAM*. The definition is completely similar to the one in Baltag and Moss (2004), so we skip the details here. Suffice to say that the semantics of basic dynamic modalities is given by the inverse map:

$$|| < \sigma \vec{\varphi} > \psi||_{\mathbf{S}} = \left(\sigma ||\vec{\varphi}||_{\mathbf{S}}\right)^{-1} ||\psi||_{\mathbf{S} \otimes \mathbf{\Sigma} ||\vec{\varphi}||}$$

**Notation**. To state our proof system, we encode the notion of *post-conditional contextual* appearance of an action in our syntax. For sentences  $\theta, \psi$  and basic program  $\alpha = \sigma \vec{\varphi}$ , we put:

$$<\alpha_a^{\theta}>\psi:=\bigvee_{\Pi\subseteq\mathbf{\Sigma}\vec{\varphi}}\left(<\alpha_a^{\Pi}>\psi\wedge\bigwedge_{\beta\in\Pi}\neg K_a\neg<\beta>\theta\wedge\bigwedge_{\beta'\notin\Pi}K_a\neg<\beta'>\theta\right)$$

This notation can be justified by observing that it semantically matches the modality corresponding to post-conditional contextual appearance:

$$|| < (\sigma \vec{\varphi})_a^{\theta} > \psi||_{\mathbf{S}} = \{ s \in S : s \in \left( (\sigma || \vec{\varphi} \vec{|} |_{\mathbf{S}})_a^{s, ||\theta||_{\mathbf{S}}} \right)^{-1} ||\psi||_{\mathbf{S} \otimes \mathbf{\Sigma} || \vec{\varphi} ||} \}$$

**Theorem.** A complete proof system for the logic  $CDL(\Sigma)$  is obtained by adding to the above axioms and rules of CDL the following Reduction Axioms:

$$\begin{array}{rcl} <\pi\cup\pi'>\varphi&\leftrightarrow&<\pi>\varphi\vee<\pi'>\varphi\\ <\pi;\pi'>\varphi&\leftrightarrow&<\pi><\pi'>\varphi\\ <\alpha>p&\leftrightarrow⪯(\alpha)\wedge p\\ <\alpha>\neg\varphi&\leftrightarrow⪯(\alpha)\wedge\neg<\alpha>\varphi\\ <\alpha>(\varphi\vee\psi)&\leftrightarrow&<\alpha>\varphi\vee<\alpha>\psi\\ <\alpha>B_a^\theta\varphi&\leftrightarrow⪯(\alpha)\wedge B_a^{<\alpha_a^\theta>\theta}\left[\alpha_a^\theta\right](\theta\rightarrow\varphi) \end{array}$$

where p is any atomic sentence,  $\pi, \pi'$  are programs and  $\alpha$  is a *basic* program in  $L(\Sigma)$ .

Acknowledgments. During the writing of this paper, the second author's research was sponsored by the Flemish Fund for Scientific Research (FWO, Brussel).

<sup>&</sup>lt;sup>13</sup>A syntactic CDAM is just a conditional doxastic frame endowed with a syntactic precondition map, associating sentences to basic action. For justification and examples, in the context of *epistemic* action models, see Baltag and Moss (2004).

## Bibliography

- E. W. Adams. (1965) A logic of conditionals. Inquiry 8: 166-197.
- C.E. Alchourron, P. Gardenfors, D. Makinson (1985). On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *The Journal of Symbolic Logic* **50**, No 2, 510-530.
- G. Aucher (2003). A Combined System for Update Logic and Belief Revision. Master's thesis, University of Amsterdam.
- A. Baltag and L.S. Moss (2004). Logics for Epistemic Programs. Synthese, 139, 165-224.
- A. Baltag, L.S. Moss and S. Solecki (1998). The Logic of Common Knowledge, Public Announcements, and Private Suspicions. In I. Gilboa (ed.), Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge, (TARK'98), 43-56.
- A. Baltag and M. Sadrzadeh. (2005) The Algebra of Multi-Agent Dynamic Belief Revision. To appear.
- A. Baltag and S. Smets (2006). Conditional Doxastic Models: Qualitative Approach A Belief Revision. WOLLIC to Dynamic Accepted for presentation  $\operatorname{at}$ '06. Available athttp://www.vub.ac.be/CLWF/SS/WOLLIC.pdf
- J. Bennett (2003). A philosophical guide to conditionals. Oxford Univ. Press.
- H. van Ditmarsch (2005). Prolegomena to Dynamic Logic for Belief Revision. Synthese, 147, 229-275.
- P. Gardenfors (1986). Belief Revisions and the Ramsey Test for Conditionals. *Philosophical Review*, 95, 81-93.
- P. Gardenfors (1988). Knowledge in Flux: Modelling the Dynamics of Epistemic States. MIT Press, Cambridge MA.
- J. Gerbrandy (1999). Dynamic Epistemic logic. In L.S. Moss et al. (eds), Logic, Language and Information, vol. 2, CSLI Publications, Stanford University.
- J. Y. Halpern (2003). Reasoning about Uncertainty. MIT Press.
- A. Fuhrmann and I. Levi (1994). Undercutting and the Ramsey Test for Conditionals. Synthese, 101,
- A. Grove (1988). Two Modellings for Theory Change. Journal of Philosophical Logic, 17:2, 157-170.
- H. Katsumi and A. O. Mendelzon (1992). On the difference between updating a knowledge base and revising it. Belief Revision, in Cambridge Tracts in Theoretical Computer Science 29, 183-203.
- D. Lewis (1973). Counterfactuals. Blackwell Publishing, Oxford.
- J.-J. Ch. Meyer and W. van der Hoek (1995). Epistemic Logic for AI and Computer Science. Cambridge Univ. Press.
- F. Ramsey (1931). The Foundations of Mathematics and Other Essays. Kegan Paul, London.
- H. Rott (1989). Conditionals and theory change: revisions, expansions, and additions. Synthese, 81, 91-113.
- M. Ryan and P.Y. Schobbens (1997). Counterfactuals and updates as inverse modalities. *Journal of Logic, Language and Information*.
- K. Segerberg (1998). Irrevocable Belief Revision in Dynamic Doxastic Logic. Notre Dame Journal of Formal Logic, 39, No 3, 287-306.
- K. Segerberg (1999). Default Logic as Dynamic Doxastic Logic. Erkenntnis, 50, 333-352.
- W. Spohn (1988). Ordinal conditional functions: A dynamic theory of epistemic states. In W.L. Harper, B. Skyrms (eds.), Causation in Decision, Belief Change and Statistics, vol. 2, 105-134. Reidel, Dordrecht.
- R.C. Stalnaker (1968). A Theory of Conditionals. In N. Rescher (ed.), *Studies in Logical Theory*, Oxford, Blackwell, APQ Monograph No2.
- J. van Benthem, J. van Eijck and B. Kooi (2005). Logics of Communication and Change. Available at http://staff.science.uva.nl/ johan/publications.html.
- J. van Benthem (2005). Dynamic Logic for Belief Change. Working-paper (version 30 november 2005).