Introspective forgetting

Hans van Ditmarsch*
Computer Science, University of Otago, New Zealand & Institut de Recherche en Informatique, Université de Toulouse, France, hans@cs.otago.ac.nz

Andreas Herzig
Institut de Recherche en Informatique, Université de Toulouse, France, herzig@irit.fr

Jérôme Lang
LAMSADE, Université Paris-Dauphine, France, lang@lamsade.dauphine.fr

Pierre Marquis
Centre de Recherche en Informatique de Lens, Université d'Artois, France, marquis@cril.univ-artois.fr

October 27, 2008

Abstract. We model the forgetting of propositional variables in a modal logical context where agents become ignorant and are aware of each others' or their own resulting ignorance. The resulting logic is sound and complete. It can be compared to variable-forgetting as abstraction from information, wherein agents become unaware of certain variables: by employing elementary results for bisimulation, it follows that beliefs not involving the forgotten atom(s) remain true.

Keywords: modal logic, forgetting, abstraction, action logic, belief change

1. There are different ways of forgetting

Becoming unaware In the movie 'Men in Black', Will Smith makes you forget knowledge of extraterrestrials by flashing you with a light in the face. After that, you have forgotten the green ooze flowing out of mock-humans and such: you do not remember that you previously had these experiences. In other words, even though for some specific forgotten fact p, previously Kp or $K\neg p$, the flash victims are not currently aware of their ignorance about p. One solution is to model that the flash victims no longer have p available as a propositional variable in the logical language. This sort of forgetting (amnesia) is dual to awareness, wherein agents become aware of a new propositional variable p.

Becoming ignorant A different sort of forgetting is when you forgot which of two keys fits your office door, because you have been away from town for a while. Is it the bigger or the smaller key? This is

* Hans van Ditmarsch is corresponding author.

about forgetting the value of an atomic proposition p—such as “the bigger key fits the door.” You are embarrassingly aware of your current ignorance: introspection is involved. We have $K(\neg Kp \land \neg K\neg p)$. This sort of forgetting is central to our concerns.

Remembering prior knowledge You may also remember that you *knew* which key it was. You just forgot. Previously Kp or $K\neg p$, and only now $\neg Kp$ and $\neg K\neg p$.

Forgetting values Did it ever happen to you that you met a person whose face you recognize but whose name you no longer remember? Surely! Or that you no longer know the pincode of your bankcard? Hopefully not. But such a thing is very conceivable. This sort of forgetting means that you forgot the value of a proposition, or the assignment of two values from different sets of objects to each other. An atomic proposition about your office keys is a feature with two values only, true and false. The (finitely) multiple-valued feature can be modelled by means of a *number* of atomic propositions. Forgetting of such multiple boolean variables is in our approach similar to forgetting a single boolean variable.

Multi-agent versions of forgetting Will Smith only had to flash a whole group once, not each of its members individually. So, in a multi-agent setting some aspects of collectively ‘becoming unaware’ can be modelled. A different but much more familiar phenomenon is that of an individual becoming unaware in a group: “You forgot my birthday, *again*!”

A group version for becoming ignorant should involve prior common knowledge followed by some sort of collective memory loss. This collective introspective character is not always easy to justify—but societal phenomena come to mind like collective denial of the Holocaust, collective belief among banks that there is an inexhaustive supply of credit, etc. Becoming ignorant for *individuals* in a group is more intuitive, because they interact with their environment: here you are standing in front of your office door again, now in company of four freshmen students, “Ohmigod, I forgot again which is my office key!”

Forgetting modal formulas I may have forgotten whether you knew about a specific review result for our jointly editored journal issue. In other words, previously $K_{me}K_{you}\mathit{accept}$ or $K_{me}K_{you}\neg\mathit{accept}$ but currently $\neg K_{me}K_{you}\mathit{accept}$ and $\neg K_{me}K_{you}\neg\mathit{accept}$. Some meaningful propositions that can be forgotten are themselves modal.

Defaulting on obligations Say I forgot to pick you up at the airport at 4:30 PM. Defaulting on an obligation involves loss of aspects of a plan that one intends to execute. Obligations can be modelled with deontic logics. We will not be concerned with this sort of forgetting.

Forgetting events Forgetting that a certain *event* (action) took place in the past is different again from forgetting a *proposition*. It models increased temporal uncertainty instead of increased epistemic uncertainty. Of course, epistemic uncertainty can follow from temporal uncertainty: if you forgot that you told me that p, you have forgotten the consequences of telling me p, namely that I know p. 'Forgetting of events' amounts to introducing *temporal uncertainty* in the model. This goes beyond the uncertainty about the status of the current world given knowledge of what time it is. We have not looked at this in detail either, partly because in state-of-the-art dynamic epistemics it is problematic to combine temporal modal and epistemic modal uncertainty.

Cognitive modelling of forgetting Forgetting tends to be a gradual process. First you know something, then, slowly, it recedes from memory, i.e. it drops below the level of awareness, and only then you become aware having forgotten it (e.g. when standing in front of your office door, realizing that you forgot which was your office key). This can subsequently even be followed by a process of conscious stepwise recovery of the lost information, by recalling the circumstances under which you still knew. In other words, you finally remember your knowledge. All of these cannot be modelled in a dynamic epistemic logic wherein agents' beliefs are deductively closed.

2. Motivation

A short history of forgetting in AI Forgetting as abstraction (a.k.a. projection, marginalization, ...), not necessarily in an epistemic setting, goes back a long way. Boole described variable abstraction in (Boole, 1854) wherein he called this the 'elimination of middle terms'. Lin and Reiter propose in *'Forget it!'* (Lin and Reiter, 1994) a way to abstract from ground atoms (that can be equated to propositional variables) in a set of first-order beliefs, employing the notion of *similarity of models for a theory except for such a ground atom*. They leave it open whether such forgetting is the result of an agent consciously updating a knowledge base after having learnt about *factual* change, or whether this is simple erosion of her working memory, purely *epistemic* change. Their work was built upon by Lang, Liberatore and Marquis with their in-depth study on the computational costs of transforming theories by variable forgetting (Lang et al., 2003), or rather the costs of determining the independence of parts of a theory from specific variables. Baral and Zhang joined this battleground for more explicit operators for knowledge and belief with their (Baral and Zhang, 2005), wherein the result of an agent forgetting a variable results in her (explicit) ignorance of that variable's value, and in (Zhang and Zhou, 2008), in progress, Zhang and Zhou make an original and interesting backtrack to the ideas of (Lin and Reiter, 1994) by suggesting *bisimulation invariance except for the forgotten variable*, in order to model forgetting. Forgetting has been generalized to logic programs in (Wang et al., 2005; Zhang et al., 2005; Eiter and Wang, 2006) and to description logics in (Zhao et al., 2007). Forgetting of (abstracting from) actions in planning has been investigated in (Erdem and Ferraris, 2007).

Forgetting proposional variables in propositional logic In (Lang et al., 2003), three ways of propositional variable forgetting in propositional logic are distinguished. Given a (propositional) formula φ and a propositional variable p, we write $Fg(\varphi, p)$ for the result (also of type formula) of forgetting variable p in φ. Let $\varphi(\psi/p)$ be the replacement of all (possibly zero) occurrences of p in φ by ψ.

1. $Fg(\varphi, p)$ is the result of replacing p in φ by 'true' in disjunction with replacing p in φ by 'false': $Fg(\varphi, p) \equiv \varphi(\top/p) \vee \varphi(\bot/p)$.

2. $Fg(\varphi, p)$ is the strongest formula not containing p that follows from φ; i.e., $\varphi \models Fg(\varphi, p)$ and p not in $Fg(\varphi, p)$, and for all χ not containing p, if $\varphi \models \chi$ then $Fg(\varphi, p) \models \chi$.

3. $Fg(\varphi, p)$ is the strongest formula that is true in all models (valuations / interpretations V of propositional variables to true and false) that only (may) differ from a φ-model in the valuation of p.

The three formulations are equivalent. For typical examples, $Fg(p \wedge q, p) = q$, and $Fg(p \leftrightarrow q, p) = \top$. The generalization to a set of formulas $\Phi = \{\varphi_1, \varphi_2, ...\}$, and to a finite set of variables $P = \{p_1, \dots, p_n\}$ is obvious. Propositional forgetting is not closed under conjunction, e.g., using the first, constructive, definition: $Fg(p, p) = \top$ and $Fg(\neg p, p) = \top$, whereas $Fg(p \wedge \neg p, p) = \bot$. Note that $Fg(p \wedge \neg p, p)$ is different from $Fg(\{p, \neg p\}, p)$. For the former, we have $\bot$ iff $Fg(p \wedge \neg p, p)$, and for the latter, we have $Fg(\{p, \neg p\}, p)$ iff $\top$.

Forgetting proposional variables in modal formulas Consider again the formula $p \wedge q$, from which q remains after forgetting p. Such a formula is often seen as part of the beliefs or knowledge of an agent. We can express this in the logic as $K(p \wedge q)$ or, equivalently, $Kp \wedge Kq$. What should result when the agent forgets p? This depends. If we interpret forgetting as *becoming unaware*, we would like the result of forgetting p to be Kq. But if we interpret forgetting as *becoming ignorant*, we would like the result of forgetting p to be $Kq \wedge \neg Kp \wedge \neg K\neg p$—'the agent knows q and is uncertain about p', where introspection of knowledge also delivers $Kq \wedge K\neg Kp \wedge K\neg K\neg p$—'the agent knows q and is aware of his uncertainty about p'. To avoid confusion, for forgetting as becoming ignorant we write Fg^i and for forgetting as becoming unaware we write Fg^u.

Let us see what happens if we straightforwardly apply the alternative characterizations of propositional forgetting to modal formulas (implicitly shifting from valuations as models to Kripke structures as models). In all cases this clearly concerns forgetting as becoming unaware. Unfortunately we get counterintuitive results for ignorance formulas, and their equivalence no longer holds. According to construction number 1, wherein we replace the atom p by $\top$ and by $\bot$, in disjunction, $Fg^u(\neg Kp \wedge \neg K\neg p, p)$ is equivalent to the contradiction $\bot$, and this is clearly undesirable.

$Fg^u(\neg Kp \wedge \neg K\neg p, p)$
is by definition
$(\neg K\top \wedge \neg K\neg\top) \vee (\neg K\bot \wedge \neg K\neg\bot)$
iff (assuming the property of belief and knowledge that $K\top \leftrightarrow \top$)
$(\neg\top \wedge \neg\bot) \vee (\neg\bot \wedge \neg\top)$
iff
$\bot$

Clearly, the problem we are having can be localized to the circumstance that for knowledge and belief, $K\top \leftrightarrow \top$ and $K\bot \leftrightarrow \bot$. It is quite common for $\neg Kp$ and $\neg K\neg p$ to be true simultaneously, but replacing this atom by true or false wreaks havoc.

On the other hand, applying the second or third recipe (constructions 2 and 3) we get $\top$ instead. Concerning 2: the only candidate formulas following from $\neg Kp \wedge \neg K\neg p$ but (equivalent to a formula) not containing p are $\top$ and $\bot$. The strongest ψ following from $\neg Kp \wedge \neg K\neg p$ not containing p must therefore be $\top$. Concerning 3: observe that, whatever the original formula φ was, none of Kp, $K\neg p$, $\neg Kp$, $\neg K\neg p$ satisfies the requirement, and clearly also not simple literals p and $\neg p$. This leaves us only with trivial formulas such as $Kp \vee \neg Kp$, that are equivalent to $\top$.

These problems when defining forgetting as becoming unaware can be avoided when restricting oneself to forgetting propositional variables in *positive* epistemic formulas (Herzig et al., 2003), i.e. subformulas expressing ignorance are not allowed, but otherwise it is hard. And forgetting as becoming ignorant is yet a different matter. It is possible to define forgetting of variables p in single-agent modal formulas as an operation $Fg^i(\varphi, p)$ in a syntactic way (work in progress by the authors, (van Ditmarsch et al., 2008b)), but this does not generalize straightforwardly to multi-agent epistemic logic and becomes rather technical anyway. Such problems motivated us to model forgetting as an event in a dynamic epistemic logic.

We will focus throughout our contribution on forgetting as becoming ignorant, that we call introspective forgetting. In Section 5 we will relate this to forgetting as becoming unaware. We remind the reader that for the former we write Fg^i (i for ignorant) and for the latter we write Fg^u (u for unaware), regardless of their context.

Forgetting as a dynamic modal operator Let us now model forgetting an atomic proposition p as an *event* $Fg^i(p)$. We do this in a propositional logic expanded with an epistemic modal operator K and a dynamic modal operator $[Fg^i(p)]$, with obvious multiple-value and multi-agent versions. Formula $[Fg^i(p)]\varphi$ means that after the agent forgets his knowledge about p, φ is true. We call $[Fg^i(p)]$ a *dynamic* modal operator because it is interpreted by a state transformation, more particularly: by changing an information state that is represented by a pointed Kripke model (M, s) into another information state (M', s'). The relation to a (supposed) formula-transforming operation $Fg^i(\psi, p)$ is as follows:

$$\psi \leftrightarrow [Fg^i(p)]Fg^i(\psi, p)$$

In other words: if proposition ψ is true, then after forgetting about p precisely that proposition is now true that is constructed from ψ relative to the forgetting of p—where of course we have only seen such a construction, and an infelicitous one at that, for unawareness forgetting $Fg^u(\psi, p)$.

A *precondition* for event $Fg^i(p)$ seems prior knowledge of the value of p: $Kp \vee K\neg p$. How can you forget something unless you know it in the first place? To make our approach comparable to variable forgetting in the 'abstracting-from-information'-sense, we do not require prior knowledge as a precondition for forgetting. (But it is easy to adapt our results to incorporate that precondition: see Section 3.3.) The obvious *postcondition* for event $Fg^i(p)$ is ignorance of the value of p: $\neg Kp \wedge \neg K\neg p$. It should therefore be valid that

$$[Fg^i(p)](\neg Kp \wedge \neg K\neg p).$$

The further refinement with remembering prior knowledge we defer to further research, see the concluding Section 7.

Forgetting or no-forgetting? Wasn't dynamic epistemic logic supposed to satisfy the principle of 'no forgetting' (a.k.a. 'perfect recall')? This entails that positive knowledge such as Kp and $K\neg p$ is preserved after any event. Or, dually: if you are ignorant about p now, then you must have been ignorant about p before. So how on earth can one model forgetting in this setting? We can, because we cheat. We solve this dilemma by the standard everyday solution of forgetful people: blame others. In this case: blame the world; we *simulate* forgetting by *changing the value of p in the actual or other states, in a way known to be unobservable by the agent*. Thus resulting in her ignorance about p. Our logic therefore facilitates ontic and epistemic change (factual and informative events, belief update and belief revision, ...).

Our solution to make positive knowledge disappear, is different from how belief revision is modelled in dynamic epistemic (doxastic) logic. Prior belief in p that is revised with $\neg p$ and results in belief in $\neg p$ is standardly modelled by considering this a 'soft' or defeasible form of belief, i.e., not knowledge, and implemented by changing a preference relation between states (van Ditmarsch, 2005; Baltag and Smets, 2006).

Having cheated in that way, our logic is equivalent to one without actual change of facts in the one and only way that counts: it makes no difference for believed formulas, i.e., for expressions of the form $K\varphi$ (see Proposition 15).

In the next section we straightforwardly implement a proposal incorporating all the above constraints.

3. Introspective forgetting

We *only* present a version of the logic for a single agent and for forgetting a single variable at a time. All results trivially generalize to multiple agents and forgetting multiple values simultaneously. They are treated separately in Section 6 on page 18.

3.1. Language, structures and semantics

Given is a set P of propositional variables.

Definition 1 (Language and structures) Our language $\mathcal{L}$ is

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi \mid [Fg^i(p), j]\varphi$$

where $p \in P$, $j = 0, 1$, and our structures are *pointed Kripke models* $((S, R, V), s)$, with $R \subseteq (S \times S)$, $V : P \to \mathcal{P}(S)$, and $s \in S$. ⊣

If $P' \subseteq P$, then $\mathcal{L}(P')$ is the language restricted to P'. The diamond versions of our modal operators are defined as $\hat{K}\varphi \equiv \neg K \neg \varphi$ and $\langle Fg^i(p), j\rangle\varphi \equiv \neg[Fg^i(p), j]\neg\varphi$. We also define by abbreviation $[Fg^i(p)]\varphi \equiv ([Fg^i(p), 0]\varphi \wedge [Fg^i(p), 1]\varphi)$. For $[Fg^i(p)]\varphi$ we read "After the agent forgets that p, φ holds." The structures are typically $S5$ to model knowledge and $KD45$ to model belief—but this is not a requirement. We employ a common notion of structural similarity between Kripke models that guarantees their logical equivalence (in other words: they describe the same set of beliefs), namely *bisimilarity*; for a definition see (Blackburn et al., 2001).

The dynamic operator $[Fg^i(p)]$ is relative to the state transformer $Fg^i(p)$ that is an *event model*. The pointed Kripke models are static structures, encoding knowledge and belief, and the event models are dynamic structures, encoding *change of* knowledge and belief. Formally, *multiple-pointed event models* (a.k.a. action models) are structures $(\mathsf{M}, \mathsf{S}') = ((\mathsf{S}, \mathsf{R}, \mathsf{pre}, \mathsf{post}), \mathsf{S}')$, where $\mathsf{S}' \subseteq \mathsf{S}$, where $\mathsf{pre} : \mathsf{S} \to \mathcal{L}$ assigns to each event $\mathsf{s} \in \mathsf{S}$ a *precondition* and where $\mathsf{post} : \mathsf{S} \to (P \to \mathcal{L})$ assigns to each event a *postcondition*, a.k.a. assignment, for each atom of a *finite* subset of all atoms. The remaining atoms do not change value. For such event models see (van Benthem et al., 2006; van Ditmarsch and Kooi, 2008)—we follow notational conventions as in (van Ditmarsch and Kooi, 2008). If $\mathsf{post}(\mathsf{s})(p) = \psi$, then we also write that $p := \psi$ (the valuation of atom p becomes that of formula ψ) in the event of s. Dynamic operators expressing event model execution (*semantics*) can be seen as part of the logical language (*syntax*), similar to how this is done for automata-PDL (Harel et al., 2000).

Forgetting $Fg^i(p)$ is the event model that expresses that the agent cannot distinguish between two assignments having taken place: p becomes true, or p becomes false. It consists of two events, that are both points (this expresses non-determinism). Both events are always executable: their precondition is $\top$. The name of the event where p is set to false is 0, and the name of the event where p is set to true is 1.

Definition 2 (Forgetting) $Fg^i(p)$ is the event model $((\mathsf{S}, \mathsf{R}, \mathsf{pre}, \mathsf{post}), \mathsf{S}')$ where $\mathsf{S} = \{0, 1\}$, $\mathsf{R} = \mathsf{S} \times \mathsf{S}$, $\mathsf{pre}(0) = \top$ and $\mathsf{pre}(1) = \top$, $\mathsf{post}(0)(p) = \bot$ and $\mathsf{post}(1)(p) = \top$ (and $\mathsf{post}(i)(q) = q$ for all $q \neq p$, $j = 0, 1$), and $\mathsf{S}' = \mathsf{S}$. ⊣

Definition 3 (Semantics) Assume an epistemic model $M = (S, R, V)$.

$M, s \models p$	iff	$s \in V(p)$	
$M, s \models \neg\varphi$	iff	$M, s \not\models \varphi$	
$M, s \models \varphi \wedge \psi$	iff	$M, s \models \varphi$ and $M, s \models \psi$	
$M, s \models K\varphi$	iff	for all $t \in S : (s, t) \in R$ implies $M, t \models \varphi$	
$M, s \models [Fg^i(p), j]\varphi$	iff	$M \otimes Fg^i(p), (s, j) \models \varphi$	for $j = 0, 1$

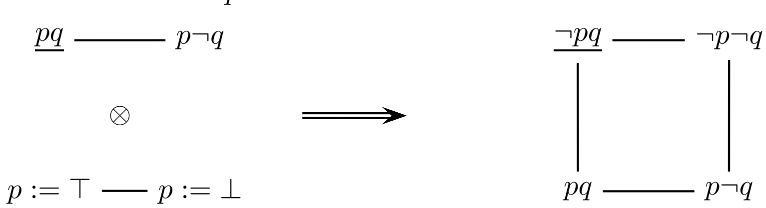
where $M \otimes Fg^i(p) = (S', R', V')$ such that $S' = S \times \{0, 1\}$, $((s, i), (t, j)) \in R'$ iff $(s, t) \in R$ and $i, j \in \{0, 1\}$, $V'(p) = \{(s, 1) \mid s \in S\}$ and $V'(q) = V(q) \times \mathsf{S}$ for $q \neq p$. The set of validities is called FG. $\dashv$

In fact, $M \otimes Fg^i(p)$ is the restricted modal product of M and event model $Fg^i(p)$ according to (Baltag et al., 1998; van Ditmarsch and Kooi, 2008), which in this case amounts to taking two copies of the model M, making p true everywhere in the first, making p false everywhere in the second, and making corresponding states indistinguishable for the agent.

Another reason to write $[Fg^i(p)]\varphi$ for $[Fg^i(p), 0]\varphi \wedge [Fg^i(p), 1]\varphi$—apart from preferring the former over the latter as the conceptual but not actual primitive of the logical language—is that we can indeed see $Fg^i(p)$ as an event model without a chosen point, i.e., as a non-deterministic event with two possible outcomes.

Example We visualize $S5$ models by linking states that are indistinguishable for an agent. Reflexivity and transitivity are assumed. In these visualizations we abuse the language by writing valuations instead of states and postconditions instead of events, and we write $\top$ for an event with empty postcondition. The actual state is underlined.

Suppose the agent knows p but does not know q (and where in fact q is true), and where the agent forgets that p. The execution of event model $Fg^i(p)$ (and where in fact p becomes false) is pictured as follows. In the resulting Kripke model, the agent no longer knows p, and remains uncertain about q.



3.2. Axiomatization

To obtain a complete axiomatization **FG** for the logic FG we apply the reduction axioms for event models, as specified in (Baltag et al., 1998) and (van Ditmarsch and Kooi, 2008). The more intuitive version of these axioms is found and explained at length in Proposition 6, below.

Definition 4 (Axiomatization FG) Only axioms involving Fg are shown.

$$\begin{array}{lcll}
[Fg^i(p),0]p & \leftrightarrow & \bot & \\
[Fg^i(p),1]p & \leftrightarrow & \top & \\
[Fg^i(p),j]q & \leftrightarrow & q & \text{for } q \neq p \\
[Fg^i(p),j]\neg\varphi & \leftrightarrow & \neg[Fg^i(p),j]\varphi & \\
[Fg^i(p),j](\varphi \wedge \psi) & \leftrightarrow & [Fg^i(p),j]\varphi \wedge [Fg^i(p),j]\psi & \\
[Fg^i(p),j]K\varphi & \leftrightarrow & K[Fg^i(p),j]\varphi & \dashv
\end{array}$$

Theorem 5 Axiomatization **FG** is sound and complete. ⊣

Proof. Completeness follows automatically, as this logic is merely event model logic for one particular event (Baltag et al., 1998; van Ditmarsch and Kooi, 2008). It may be instructive for the reader to see why the above axioms are instances or simple consequences of axioms in this event logic. For example, our basic cases are applications of a general axiom (for any event, not just forgetting) of form $[\mathsf{M},\mathsf{s}]p \leftrightarrow (\mathsf{pre}(\mathsf{s}) \rightarrow \mathsf{post}(\mathsf{s})(p))$. We get:

$[Fg^i(p),0]p$ iff $(\mathsf{pre}(0) \rightarrow \mathsf{post}(0)(p))$ iff $(\top \rightarrow \bot)$ iff $\bot$

$[Fg^i(p),1]p$ iff $(\mathsf{pre}(1) \rightarrow \mathsf{post}(1)(p))$ iff $(\top \rightarrow \top)$ iff $\top$

$[Fg^i(p),j]q$ iff $(\mathsf{pre}(j) \rightarrow \mathsf{post}(j)(q))$ iff $(\top \rightarrow q)$ iff q

The case for conjunction after an event is basic. Both negation and knowledge commute with (the deterministic event models representing) forgetting, because these events can always be executed.

3.3. Semantic results

Proposition 6 (Principles of forgetting)

$$\begin{array}{lcll}
[Fg^i(p)]p & \leftrightarrow & \bot & \\
[Fg^i(p)]q & \leftrightarrow & q & \text{for } q \neq p \\
[Fg^i(p)](\varphi \wedge \psi) & \leftrightarrow & [Fg^i(p)]\varphi \wedge [Fg^i(p)]\psi & \\
[Fg^i(p)]K\varphi & \leftrightarrow & K[Fg^i(p)]\varphi & \\
[Fg^i(p)][Fg^i(p)]\varphi & \leftrightarrow & [Fg^i(p)]\varphi & \dashv
\end{array}$$

Proof. All proofs are trivial. The case $[Fg^i(p)]p$ for the forgotten atom expresses that you cannot guarantee that p is true after forgetting it *by way of varying its value.* Variable p became either true (after event $[Fg^i(p), 1]$) or false (after event $[Fg^i(p), 0]$) but the agent does not know which of the two actually took place. Formally, we have that $[Fg^i(p)]p$ iff $[Fg^i(p), 0]p \wedge [Fg^i(p), 1]p$ iff (see previous) $(\top \rightarrow \bot) \wedge (\top \rightarrow \top)$ which equals $\bot$. The reader might object that in forgetting, facts do not 'really' change value. We can accommodate that: see Section 4 for such a forgetting operator. But in all that matters this operator is indistinguishable from $Fg^i(p)$. Also, from our point of view, saying that facts should not 'really' change value amounts to saying that forgetting should not 'really' be possible for rational agents (as our logic validates the principle of no-forgetting)... Still, it happens.

There is no case for negation, because $[Fg^i(p)]\neg\varphi$ is *not* equivalent to $\neg[Fg^i(p)]\varphi$. The left-hand side reduces as follows: $[Fg^i(p)]\neg\varphi$ iff $[Fg^i(p), 0]\neg\varphi \wedge [Fg^i(p), 1]\neg\varphi$ iff (using the axioms) $\neg[Fg^i(p), 0]\varphi \wedge \neg[Fg^i(p), 1]\varphi$ iff $\neg[Fg^i(p), 0]\varphi \wedge \neg[Fg^i(p), 1]\varphi$. But the right-hand side reduces differently: $\neg[Fg^i(p)]\varphi$ iff $\neg([Fg^i(p), 0]\varphi \wedge [Fg^i(p), 1]\varphi)$ iff $\neg[Fg^i(p), 0]\varphi \vee \neg[Fg^i(p), 1]\varphi)$; we end up with a disjunction instead of a conjunction![1]

The epistemic operator commutes with the forgetting operator. Thus the consequences of forgetting are known before it takes place ('no miracles'). We emphasize that the negated epistemic operator (for 'possible that') does *not* commute with forgetting ($[Fg^i(p)]\neg K\varphi$ is *not* equivalent to $\neg K[Fg^i(p)]\varphi$); therefore, K cannot be eliminated.

The idempotency of forgetting is not surprising either. After having become ignorant about p, repeating that procedure has no effect. This shows directly from the semantics. (It is also possible to have this as a derived theorem schema, given the axiomatization. This needs induction over the formula φ true after forgetting. The proof is rather involved.) When executing forgetting in an epistemic state, every state in its domain is split into one wherein p is true and an indistinguishable one wherein p is false. The next execution of $Fg^i(p)$ simply repeats that procedure, and we now end up with four indistinguishable states for every initial state, in two of which p is false and in two of which p is true.

[1] We tried to have the non-deterministic $[Fg^i(p)]\varphi$ as our only primitive in the language, and to determine an axiomatization based on that. It is then very possible to define the pointed (deterministic) versions of the forgetting event as notational abbreviations of the not-pointed primitive, namely as $[Fg^i(p), 1]\varphi \equiv [Fg^i(p)](p \rightarrow \varphi)$ and $[Fg^i(p), 0]\varphi \equiv [Fg^i(p)](\neg p \rightarrow \varphi)$. From this then follows the axiom for non-determinism $[Fg^i(p)]\varphi \leftrightarrow [Fg^i(p), 0]\varphi \wedge [Fg^i(p), 1]\varphi$. In the axiomatization we then need an axiom "$[Fg(p)]\neg\varphi$ iff $\neg[Fg(p)](\neg p \rightarrow \varphi) \wedge \neg[Fg(p)](p \rightarrow \varphi)$". The formula on the right has no lower complexity then the formula on the left—at least not in any meaningful way that we could use to continue reductions; and note that there is no reduction rule for implication!

Formally, the result is bisimilar to just executing forgetting once. Let M be the model we start out with, then $\mathfrak{R} : (s,j) \leftrightarrows (s,i,j)$, for $i,j = 0,1$, is a total bisimulation between $M \otimes Fg^i(p)$ and $M \otimes Fg^i(p) \otimes Fg^i(p)$.

Some more derived principles are:

Corollary 7

$$\begin{array}{lcl} [Fg^i(p)]\neg p & \leftrightarrow & \bot \\ [Fg^i(p)]Kp & \leftrightarrow & \bot \\ [Fg^i(p)]K\neg p & \leftrightarrow & \bot \end{array} \qquad \dashv$$

It is important to point out the differences between $Fg^i(\varphi, p)$ and $[Fg^i(p)]\varphi$ again. In the former, φ is a precondition, a formula true before forgetting. In the latter, φ is a postcondition, a formula true after forgetting. Let $Fg^i(\varphi, p)$ be some such adequate notion for syntactic progression of φ (i.e., how φ changes as a result of forgetting p defined a function on φ and p), for introspective forgetting. Then we should have that

$$\varphi \leftrightarrow [Fg^i(p)]Fg^i(\varphi, p)$$

is valid. This relation is *so* tight, that the difference (in the language!) is mere syntactic sugar. This suggests that the two semantic definitions of propositional forgetting à la Lang et al. (Lang et al., 2003) could be thought of as the projections to a non-epistemic setting of the dynamic modal operations defined here; employing theoretical results for the complexities of dynamic logics might throw an interesting light on their optimality results.

As long as R is serial, so that $K\bot \leftrightarrow \bot$ and $K\top \leftrightarrow \top$, ignorance is indeed obtained:

Proposition 8 On the class of serial Kripke models, $[Fg^i(p)](\neg Kp \wedge \neg K\neg p)$ is valid. $\dashv$

Proof. Validity is trivial. Thus we have derivability. It is instructive to see the derivation in the axiomatization **FG**. First, we have that

$[Fg^i(p)](\neg Kp \wedge \neg K\neg p)$ iff $[Fg^i(p)]\neg Kp \wedge [Fg^i(p)]\neg K\neg p$.

For the left conjunct of that:

$[Fg^i(p)]\neg Kp$ iff $\neg[Fg^i(p), 0]Kp \wedge \neg[Fg^i(p), 1]Kp$.

Finally, for the left conjunct of *that*:

$\neg[Fg^i(p), 0]Kp$ iff $\neg K[Fg^i(p)]p$ iff (see above) $\neg K\bot$ iff $\neg\bot$ iff $\top$.

This completes the proof one of four similar cases (conjuncts).

Corollary 9 $[Fg^i(p)]K(\neg Kp \wedge \neg K\neg p)$ is valid and derivable. $\dashv$

Proof. Automatically, as $[Fg^i(p)]K(\neg Kp \wedge \neg K\neg p)$ iff $K[Fg^i(p)](\neg Kp \wedge \neg K\neg p)$. (Knowledge commutes with forgetting.)

3.4. Precondition of prior knowledge

If we were to require a precondition of knowledge about p prior to forgetting p, i.e. precondition $Kp \vee K\neg p$, the forgetting operation becomes partially executable, namely subject to that condition, and the derived principles would become as follows. The last expresses that one cannot forget something twice (at least not without, in between, learning about p again).

Proposition 10 (Principles of forgetting with prior knowledge)

$$\begin{array}{lll} [Fg^i(p)]p & \leftrightarrow & \neg(Kp \vee K\neg p) \\ [Fg^i(p)]q & \leftrightarrow & (Kp \vee K\neg p) \rightarrow q \text{ for } q \neq p \\ [Fg^i(p)](\varphi \wedge \psi) & \leftrightarrow & [Fg^i(p)]\varphi \wedge [Fg^i(p)]\psi \\ [Fg^i(p)]K\varphi & \leftrightarrow & (Kp \vee K\neg p) \rightarrow K[Fg^i(p)]\varphi \\ [Fg^i(p)][Fg^i(p)]\bot & & \end{array}$$

⊣

4. Forgetting without changing the real world

A number of different perspectives on forgetting propositional variables (such as release, elimination, bisimulation quantification, symmetric contraction, and value swapping or switching) all amount to the same: although resulting in different structures, these cannot be distinguished from each other in the language by the agent, i.e., they all represent the same set of believed formulas.

An unfortunate side effect of our modelling of forgetting *seems* that the actual value of p gets lost in the process of forgetting, such as in the example on page 9. This is undesirable if we *only* want to model that the agents forget the value of p but that otherwise nothing changes: in particular, the actual value of p should not change. Two ways to overcome that deficiency are (purely) *epistemic forgetting* and *forgetting by swapping values.*

Definition 11 (Epistemic forgetting) Epistemic forgetting is the pointed event model $(Fg^{i\varepsilon}(p), n)$ where $Fg^{i\varepsilon}(p)$ is like $Fg^i(p)$ except that there is one more event n in the model, indistinguishable from the other two, with empty postcondition (and with precondition $\top$). ⊣

The point n with precondition $\top$ represents the event that 'nothing happens'. As it is the point of the event model, it ensures that the actual value of p does not change. We can visualize this event as follows—the actual event with precondition $\top$ is underlined.

$$p := \top \text{ ——— } \underline{\top} \text{ ——— } p := \bot$$

Definition 12 (Axioms for epistemic forgetting) The axioms for $(Fg^{i\varepsilon}(p), n)$ are as for $Fg^{i}(p)$ except that

$$\begin{array}{lcl} [Fg^{i\varepsilon}(p), n]p & \leftrightarrow & p \\ [Fg^{i\varepsilon}(p), n]\neg\varphi & \leftrightarrow & \neg[Fg^{i\varepsilon}(p), n]\varphi \\ [Fg^{i\varepsilon}(p), n]K\varphi & \leftrightarrow & K[Fg^{i\varepsilon}(p)]\varphi \end{array}$$

where $[Fg^{i\varepsilon}(p)]\varphi$ is defined (as usual for event models) as $[Fg^{i\varepsilon}(p), n]\varphi \wedge [Fg^{i\varepsilon}(p), 0]\varphi \wedge [Fg^{i\varepsilon}(p), 1]\varphi$. $\dashv$

The soundness of the first is demonstrated by

$[Fg^{i\varepsilon}(p), n]p$ iff $(\mathsf{pre}(n) \rightarrow \mathsf{post}(n)(p))$ iff $(\top \rightarrow p)$ iff p

Idempotency of forgetting, in Proposition 6, above, obviously also holds for epistemic forgetting $(Fg^{i\varepsilon}(p), n)$:

$$[Fg^{i\varepsilon}(p), n][Fg^{i\varepsilon}(p), n]\varphi \text{ iff } [Fg^{i\varepsilon}(p), n]\varphi$$

We have the results that all factual information is indeed preserved for epistemic forgetting, and that for purely epistemic formulas (formulas of type $K\varphi$) the two different definitions of forgetting are indistinguishable.

Proposition 13 (Preservation of factual information)
ψ iff $[Fg^{i\varepsilon}(p), n]\psi$, for boolean ψ. $\dashv$

Proof. Simple induction on booleans. The case for p is the axiom above. Negation: $[Fg^{i\varepsilon}(p), n]\neg\varphi$ iff (see above) $\neg[Fg^{i\varepsilon}(p), n]\varphi$ iff (IH) $\neg\varphi$. Conjunction: $[Fg^{i\varepsilon}(p), n](\varphi \wedge \psi)$ iff $([Fg^{i\varepsilon}(p), n]\varphi$ and $[Fg^{i\varepsilon}(p), n]\psi)$ iff (IH) (φ and ψ) iff $\varphi \wedge \psi$.

Lemma 14
$[Fg^{i\varepsilon}(p)]\psi$ iff $[Fg^{i}(p)]\psi$. $\dashv$

Proof. In the formulation of this lemma we are vague. The language with $[Fg^{i\varepsilon}(p), n]\varphi$ as inductive construct is different from the language with $[Fg^{i}(p)]\varphi$ as inductive construct. More strictly the result is that $[Fg^{i\varepsilon}(p), n]K\psi \leftrightarrow [Fg^{i}(p)]K\mathsf{trs}(\psi)$ is valid, subject to the translation with inductive clause $\mathsf{trs}([Fg^{i\varepsilon}(p), n]\varphi) = [Fg^{i}(p)]\mathsf{trs}(\varphi)$ and otherwise trivial clauses. The proof *can* be done with induction on ψ (where, for a crucial base case, note that $[Fg^{i\varepsilon}(p)]p$ and $[Fg^{i}(p)]p$ are both *false*), but it seems far more elegant to point out that for any epistemic model M, product $M \otimes Fg^{i\varepsilon}(p)$ is bisimilar to $M \otimes Fg^{i}(p)$, via the following (total) bisimulation $\mathfrak{R} : M \otimes Fg^{i\varepsilon}(p) \rightarrow M \otimes Fg^{i}(p)$—let $M = (S, R, V)$, and $s \in S$:

$$\begin{array}{ll} \mathfrak{R}: & (s,0) \mapsto (s,0) \\ & (s,1) \mapsto (s,1) \\ & (s,n) \mapsto (s,1) \;\text{ if } s \in V(p) \\ & (s,1) \mapsto (s,1) \;\text{ if } s \notin V(p) \end{array}$$

As bisimulation implies logical equivalence, the Lemma then follows immediately.

Proposition 15 (Epistemic propositions are preserved)
$[Fg^{i\varepsilon}(p), n]K\psi$ iff $[Fg^{i}(p)]K\psi$. ⊣

Proof. We have that $[Fg^{i\varepsilon}(p), n]K\psi$ iff $K[Fg^{i\varepsilon}(p)]\psi$ and that $[Fg^{i}(p)]K\psi$ iff $K[Fg^{i}(p)]\psi$. We now have that $(K[Fg^{i\varepsilon}(p)]\psi$ iff $K[Fg^{i}(p)]\psi)$ if ('iff', as long as we assume seriality) $([Fg^{i\varepsilon}(p)]\psi$ iff $[Fg^{i}(p)]\psi)$. Now apply Lemma 14.

Yet another way to model forgetting is by making every state in the model indistinguishable from one wherein the value of p has been swapped / switched: if true, it became false, and if false it became true.

Definition 16 (Forgetting by swapping values) Forgetting by swapping is the pointed event model $(Fg^{i\updownarrow}(p), n)$ that is like $Fg^{i}(p)$ except that in one event, the actual event n, nothing happens, whereas in the other event sw the assignment $p := \neg p$ is executed.

$$p := \neg p \;\text{———}\; \underline{\top}$$

Again we can adjust the axiomatization, we obtain the results that actual facts do not change value, and that propositions under the scope of the epistemic operator are preserved. These results can be summarized as:

Proposition 17 (Forgetting by swapping values)

- $[Fg^{i\updownarrow}(p), n]p \leftrightarrow p$ and $[Fg^{i\updownarrow}(p), sw]\neg p \leftrightarrow \neg p$.
- $\psi \leftrightarrow [Fg^{i\updownarrow}(p), n]\psi$ is valid for boolean ψ.
- $[Fg^{i\updownarrow}(p)]\psi$ iff $[Fg^{i\varepsilon}(p)]\psi$ iff $[Fg^{i}(p)]\psi$.

– $[Fg^{i\updownarrow}(p), n]K\psi$ iff $[Fg^{i\varepsilon}(p), n]K\psi$ iff $[Fg^{i}(p)]K\psi$. ⊣

The question then remains which of the three to prefer. We now know that from an agent's perspective the three are indistinguishable, and from a modeller's perspective $Fg^{i\updownarrow}(p)$ and $Fg^{i\varepsilon}(p)$ are to be preferred over $Fg^{i}(p)$, as they preserve factual knowledge. (This means that in a truly multi-agent setting, where other agents make observations similar to the modeller, individual forgetting in a group cannot be adequately modelled by $Fg^{i}(p)$ but only by the other two! See Section 7.) From a computational perspective, the two-point event models $Fg^{i\updownarrow}(p)$ and $Fg^{i}(p)$ seem preferable over the three-point $Fg^{i\varepsilon}(p)$, but on the other hand $Fg^{i}(p)$ might be preferable over the other two, as it does not involve a check on the current value of p in every state.

5. Forgetting as becoming unaware—revisited

Scrambling the valuation of the forgotten atom Instead of making p randomly (but indistinguishably!) true or false in every state of the Kripke model, the more proper way of 'releasing the value of p' in a modal logical context is to make p randomly true in a subset of the domain of the model. This is the proper generalization of the third perspective (number 3, on page 4) on variable forgetting for propositional logic. One can then make all those results indistinguishable from one another for the agent. Unlike the former, where two copies of the model M suffice, we now need $2^{|M|}$ copies.

Consider again the structure pq——$p\neg q$ encoding that the agent knows p but is ignorant about q. In proper Lin and Reiter (Lin and Reiter, 1994) fashion, the models agreeing with pq——$p\neg q$ on anything except maybe p are the following four:

pq——$p\neg q$ pq——$\neg p\neg q$ $\neg pq$——$p\neg q$ $\neg pq$——$\neg p\neg q$

These four still have ignorance about q in common, but only two of them satisfy ignorance about p. The only meaningful formulas about p true in all models, such as $\neg Kq \wedge Kp \vee \neg Kp$, that are equivalent to p-less formulas, in this case, to $\neg Kq$.

This encodes *unawareness* of p, but not *ignorance* about p. Consider the even more abstract perspective of *bisimulation quantification*: apart from the four above, pq——$p\neg q$ is also similar except for the value of p to other structures, e.g. to pq——$p\neg q$——$\neg pq$,three indistinguishable states, which satisfies that $K(p \vee q)$ is true. Because if we abstract from the value of p, q——$\neg q$ is bisimilar to q——$\neg q$——q. In other words, not just the valuation of p may vary 'at random' but also the epistemic uncertainty about its value. This prepares the ground for the next paragraph.

Bisimulation quantification Becoming unaware of an atom p can be modelled as universal bisimulation quantification over p (Visser, 1996; French, 2006) namely as

Definition 18 (Forgetting as bisimulation quantification)

$$[Fg^{u\forall}(p)]\varphi \equiv \forall p\varphi \qquad \dashv$$

where $(M,s) \models \forall p\varphi$ iff for all (M',s') such that $(M',s') \leftrightarrows_{P-p} (M,s)$: $(M',s') \models \varphi$. The notation $(M',s') \leftrightarrows_{P-p} (M,s)$ means that epistemic state (M',s') is bisimilar to epistemic state (M,s) with respect to the set of all atoms *except* p.[2]

Applied to forgetting, this is the original proposal in (Zhang and Zhou, 2008), that was later independently but very similarly suggested again in (van Ditmarsch and French, 2009). According to the proposal of (Zhang and Zhou, 2008), we now have that: $\Phi \rightarrow [Fg^{u\forall}(p)]\varphi$ is a theorem iff (in their notation) $KForget(\Phi,p) \models \varphi$ is a theorem. In (Zhang and Zhou, 2008) it is also observed that by constraining the Φ-models to those also satisfying $\neg Kp \wedge \neg K\neg p$ (some sort of belief expansion following the belief contraction induced by going to all $(P-p)$-bisimilar models), we can then enforce ignorance about p, as in (Baral and Zhang, 2005), thus relating introspective forgetting to unawareness forgetting. This is possible, and correct, but different from what we will do now to relate the two.

5.1. The relation between unawareness and ignorance

Given some epistemic model M, the model wherein p is introspectively forgotten according to $Fg^{i}(p)$ (or any other of the introspective forgetting operators that we distinguished) is bisimilar to M except for p.

$$M \leftrightarrows_{P-p} M \otimes Fg^{i}(p)$$

The total bisimulation is defined as $\mathfrak{R} : s \mapsto (s,1)$ if $s \in V(p)$ and otherwise $\mathfrak{R} : s \mapsto (s,0)$. This is a useful observation as, without any further effort, we immediately have that

Theorem 19 If $\psi \in \mathcal{L}(P-p)$ then $\psi \leftrightarrow [Fg^{u\forall}(p)]\psi$ is valid. $\dashv$

[2] A relation with variable forgetting $Fg(\varphi,p)$ for propositional logic as in (Lang et al., 2003) is that $K\varphi \rightarrow [Fg^{u\forall}(p)]K\psi$ is valid for propositional φ,ψ such that $Fg(\varphi,p) \models \psi$. This is because $\varphi \rightarrow [Fg^{u\forall}(p)]Fg(\varphi,p)$ implies $K\varphi \rightarrow K[Fg^{u\forall}(p)]Fg(\varphi,p)$ which is equivalent to $K\varphi \rightarrow [Fg^{u\forall}(p)]KFg(\varphi,p)$. From that and $Fg(\varphi,p) \models \psi$ follows $K\varphi \rightarrow [Fg^{u\forall}(p)]K\psi$.

It also follows that for formulas not containing the forgotten variable, unawareness forgetting and introspective forgetting amount to the same.

Corollary 20 If $\psi \in \mathcal{L}(P - p)$ then (ψ iff) $[Fg^{u\forall}(p)]\psi$ iff $[Fg^{i}(p)]\psi$. $\dashv$

Alternatively—although not practically—we can define an introspective version of the forgetting operation by bisimulation quantification: make corresponding points in all $P-p$-bisimilar models indistinguishable for the agent.

Given (M, s), let $\mathfrak{M} = \{(M', s') \mid (M, s) \leftrightarroweq_{P-p} (M', s')\}$. Given $\mathfrak{R} : (M, s) \leftrightarroweq_{P-p} (M', s')$ and $\mathfrak{R}' : (M, s) \leftrightarroweq_{P-p} (M'', s'')$, add pairs (s', s'') to the relation R on $\mathfrak{M}$ whenever there is a $s \in S$ such that $(s, s') \in \mathfrak{R}$ and $(s, s'') \in \mathfrak{R}'$. Clearly, $\mathfrak{M} \models \neg Kp \wedge \neg K \neg p$. Now define

Definition 21 $(M, s) \models [Fg^{i\forall}(p)]\varphi$ iff $(\mathfrak{M}, s) \models \varphi$, with $\mathfrak{M}$ as above. $\dashv$

Then $[Fg^{i\forall}(p)](\neg Kp \wedge \neg K \neg p)$ is valid. We even have that

$$M \leftrightarroweq_{P-p} \mathfrak{M} \leftrightarroweq_{P-p} M \otimes Fg^{i}(p).$$

So, we can after all reclaim introspective forgetting using a bisimulation quantification. Of course $Fg^{i}(p)$ is preferable for computational reasons to $Fg^{i\forall}(p)$: it is constructive, and requires for a given model M merely two copies; whereas $Fg^{i\forall}(p)$ is non-constructive but declarative; it is a clearly over-the-top 'construction' and it requires zillions of copies of M: $\mathfrak{M}$ is infinite!

6. More variables and more agents

Multiple variables The forgetting of multiple variables can be modelled by a simple adjustment. For n propositional variables, we get an event model $Fg^{i}(p_1, \ldots, p_n)$ with a domain consisting of 2^n events, one for each combination of assignments of different variables to true and false. All prior results still follow (including bisimulation quantification for n variables).

Proposition 22 $[Fg^{i}(p, q)]\varphi \leftrightarrow [Fg^{i}(p)][Fg^{i}(q)]\varphi$ is valid. $\dashv$

Combining learning and forgetting One might wish to combine forgetting with other dynamic operations such as learning (by public announcements). We simply add an inductive construct $[\varphi]\psi$ to the language, which stands for 'after announcement of φ, ψ holds' (see (van

Ditmarsch et al., 2007b)). The resulting logic is again equally expressive as epistemic logic: just add the rewrite rules involving announcements. One can formulate principles such as, for forgetting with prior knowledge about p as precondition: $\varphi \leftrightarrow [Fg^i(p)]([p]\varphi \vee [\neg p]\varphi)$, where $[p]$ and $[\neg p]$ are the dynamic operators corresponding to 'revealing the truth about p', i.e., the announcement of p and the announcement of $\neg p$, respectively; a.k.a. the sensing action of p. In other words: if you forget about p and afterwards learn about p again, nothing has changed. (The principle does not hold without the precondition of prior knowledge, as $Fg^i(p)$ is also executable when the agent is ignorant about p.) The principle is reminiscent of the recovery principle in AGM belief revision.

Multiple agents Given a parameter set of agents A, we adjust the language to $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_a\varphi \mid [Fg^i_B(p)]\varphi$, where $a \in A$ and $B \subseteq A$, and we adjust the accessibility relation $R \subseteq (S \times S)$ to an accessibility function $R : A \rightarrow \mathcal{P}(S \times S)$—accessibility relation R_a (for $R(a)$) interprets operator K_a (knowledge of agent a). The case $Fg^i_A(p)$ where $B = A$ models forgetting as group ignorance, and the case $Fg_a(p)$ where $B = \{a\}$ models the forgetting of an individual a in the presence of A (see page 2). Both are most succinctly modelled by a version of 'swapping values' forgetting (Definition 16) namely as event models visualized as

$$p := \neg p \overset{A}{\text{———}} \underline{\top} \qquad\qquad p := \neg p \overset{a}{\text{———}} \underline{\top}$$

The visualization on the right means that all agents *except* a can distinguish between the two alternatives: access for a is the universal relation on the domain, and access for all agents in $A - a$ is the identity. Again, all former results generalize, both versions are axiomatizable very similarly to the previous, etc. In contrast, the more obvious multi-agent version of $Fg^i(p)$ (with assignments to true and to false only) does *not* model individual forgetting in a group: this would express that the other agents learn that p is true or learn that p is false, clearly undesirable. In this case we really need either $(Fg^{i\updownarrow}(p), n)$ or $(Fg^{i\varepsilon}(p), n)$.

7. Further research

Remembering prior knowledge For the agent to recall prior knowledge we have to be able to refer to past events. Let $Fg^i(p)^-$ be the converse of $Fg^i(p)$ (e.g. in the sense of (Aucher and Herzig, 2007; Yap, 2006; Sack, 2007)). Awareness of present ignorance and prior knowledge about p can now be formalized as

$$K(\neg Kp \wedge \neg K\neg p \wedge \langle Fg^i(p)^-\rangle(Kp \vee K\neg p))$$

We now need a structure allowing us to interpret such converse events. This is not possible in pointed Kripke models, but it can be elegantly done employing what is known as the 'forest' produced by the initial Kripke model and all possible sequences of all $Fg^i(p)$ events (for all atoms), see (Parikh and Ramanujam, 1985; van Benthem et al., 2007; Pacuit, 2007; van Ditmarsch et al., 2007a; Yap, 2006; Sack, 2007). We now add assignments to the language, as in the underlying proposal, and additionally add theories for event models using converse actions (Aucher and Herzig, 2007; van Benthem et al., 2007). Thus we get a complete axiomatization, though not a reduction result to epistemic formulas (converse events cannot be eliminated from the logical language by equivalences). It seems that a reduction result can *still* not be obtained if we require prior knowledge about p as preconditions in the event model for forgetting.

Forgetting modal formulas How to model the forgetting of modal formulas is a different piece of cake altogether; in this case we have made no progress yet. We recall the introductory example wherein "I may have forgotten whether you knew about a specific review result for our jointly editored journal issue": previously $K_{me}K_{you}\mathit{accept}$ or $K_{me}K_{you}\neg\mathit{accept}$ but currently $\neg K_{me}K_{you}\mathit{accept}$ and $\neg K_{me}K_{you}\neg\mathit{accept}$. It should be noted that even in a propositional logical setting, the forgetting of non-variables is entirely unclear.

Forgetting of events This amounts to introducing *temporal uncertainty* in the model, e.g., when given a history of five prior events, the agent has forgotten that event number three took place (including, therefore, all its epistemic postconditions). This can be done by introducing histories of events to structures, or moving to a temporal epistemic perspective using 'forests', as above, see (van Benthem et al., 2007). It is clear how this has to be done, and the results should prove interesting.

Complexities It would be worthwhile to generalize the results in (Lang et al., 2003) to the modal case.

Acknowledgements

Previous versions of this work were presented at the ESSLLI 2008 workshop Logic and Intelligent Interaction, in Hamburg, Germany (see `http://ai.stanford.edu/~epacuit/LaII/`), and at the Advances in Artificial Intelligence conference (AI08), in Auckland, New Zealand (see (van Ditmarsch et al., 2008a)). The underlying version contains: a more extensive overview comparing it to known results in propositional variable forgetting, more proof details, and more details and results on multiple-variable and multi-agent forgetting. The authors would like to thank the four anonymous AI08 referees, and also the participants of the ESSLLI 2008 Hamburg workshop Logic and Intelligent Interaction, in particular Audrey Yap who observed that temporal uncertainty may also result in epistemic uncertainty that amounts to forgetting. Hans van Ditmarsch acknowledges support of the Netherlands Institute of Advanced Study where he was Lorentz Fellow in 2008.

References

Aucher, G. and A. Herzig: 2007, 'From DEL to EDL : Exploring the Power of Converse Events'. In: K. Mellouli (ed.): *ECSQARU*, Vol. 4724 of *Lecture Notes in Computer Science*. pp. 199–209.

Baltag, A., L. Moss, and S. Solecki: 1998, 'The Logic of Public Announcements, Common Knowledge, and Private Suspicions'. In: I. Gilboa (ed.): *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK 98)*. pp. 43–56.

Baltag, A. and S. Smets: 2006, 'Dynamic Belief Revision over Multi-Agent Plausibility Models'. Proceedings of LOFT 2006 (7th Conference on Logic and the Foundations of Game and Decision Theory).

Baral, C. and Y. Zhang: 2005, 'Knowledge updates: semantics and complexity issues'. *Artificial Intelligence* **164**(1-2), 209–243.

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

Boole, G.: 1854, *An investigation of the laws of thought*. Basingstoke: Macmillan Publishers.

Eiter, T. and K. Wang: 2006, 'Forgetting and Conflict Resolving in Disjunctive Logic Programming'. In: *Proceedings of AAAI*.

Erdem, E. and P. Ferraris: 2007, 'Forgetting Actions in Domain Descriptions'. In: *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*. pp. 409–414.

French, T.: 2006, 'Bisimulation quantifiers for modal logic'. Ph.D. thesis, University of Western Australia.

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

Herzig, A., J. Lang, and P. Marquis: 2003, 'Action representation and partially observable planning using epistemic logic'. In: *Proceedings of IJCAI*.

Lang, J., P. Liberatore, and P. Marquis: 2003, 'Propositional Independence: Formula-Variable Independence and Forgetting'. *J. Artif. Intell. Res. (JAIR)* **18**, 391–443.

Lin, F. and R. Reiter: 1994, 'Forget It!'. In: *Proceedings of AAAI Fall Symposium on Relevance, New Orleans*. pp. 154–159.

Pacuit, E.: 2007, 'Some comments on history-based structures'. To appear in Journal of Applied Logic.

Parikh, R. and R. Ramanujam: 1985, 'Distributed Processing and the Logic of Knowledge'. In: *Logic of Programs*, Vol. 193 of *Lecture Notes in Computer Science*. pp. 256–268. A newer version appeared in *Journal of Logic, Language and Information*, vol. 12, 2003, pp. 453–467.

Sack, Y.: 2007, 'Adding Temporal Logic to Dynamic Epistemic Logic'. Ph.D. thesis, Indiana University, Bloomington, USA.

van Benthem, J., J. Gerbrandy, and E. Pacuit: 2007, 'Merging frameworks for interaction: DEL and ETL'. In: D. Samet (ed.): *Proceedings of TARK 2007*. pp. 72–81.

van Benthem, J., J. van Eijck, and B. Kooi: 2006, 'Logics of Communication and Change'. *Information and Computation* **204(11)**, 1620–1662.

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

van Ditmarsch, H. and T. French: 2009, 'Simulation and information'. In: J. Broersen and J.-J. Meyer (eds.): *Proceedings of KRAMAS, September 2008, Sydney, Australia*. To appear. Also available as electronic proceedings of LOFT 2008, Amsterdam.

van Ditmarsch, H., A. Herzig, J. Lang, and P. Marquis: 2008a, 'Introspective forgetting'. In: W. Wobcke and M. Zhang (eds.): *AI 2008: Advances in Artificial Intelligence*.

van Ditmarsch, H., A. Herzig, J. Lang, and P. Marquis: 2008b, 'Only knowing and progression of belief'. Manuscript with work in progress.

van Ditmarsch, H. and B. Kooi: 2008, 'Semantic Results for Ontic and Epistemic Change'. In: G. Bonanno, W. van der Hoek, and M. Wooldridge (eds.): *Logic and the Foundations of Game and Decision Theory (LOFT 7)*, Texts in Logic and Games 3. Amsterdam: Amsterdam University Press, pp. 87–117.

van Ditmarsch, H., J. Ruan, and W. van der Hoek: 2007a, 'Model checking dynamic epistemics in branching time'. In: *(Informal) Proceedings of FAMAS 2007, Durham UK*.

van Ditmarsch, H., W. van der Hoek, and B. Kooi: 2007b, *Dynamic Epistemic Logic*, Vol. 337 of *Synthese Library*. Springer.

Visser, A.: 1996, 'Bisimulations, Model Descriptions and Propositional Quantifiers'. Logic Group Preprint Series 161, Department of Philosophy, Utrecht University.

Wang, K., A. Sattar, and K. Su: 2005, 'A Theory of Forgetting in Logic Programming'. In: *AAAI*. pp. 682–688.

Yap, A.: 2006, 'Product Update and Looking Backward'. Technical report, University of Amsterdam. ILLC Research Report PP-2006-39.

Zhang, Y., N. Foo, and K. Wang: 2005, 'Solving Logic Program Conflict through Strong and Weak Forgettings'. In: *Proceedings of IJCAI*. pp. 627–634.

Zhang, Y. and Y. Zhou: 2008, 'Knowledge Forgetting: Properties and Applications'. Work in progress under submission.

Zhao, Y., K. Wang, R. Topor, J. Pan, and F. Giunchiglia: 2007, 'Semantic Cooperation and Knowledge Reuse by Using Autonomous Ontologies'. In: *ISWC/ASWC*. pp. 666–679.