View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Utrecht University Repository

Belief updates in Multiple Agent Systems

Jan-Willem Roorda

UU-CS-1999-25

Belief Updates in Multiple Agent Systems

Jan-Willem Roorda*

Abstract

We give a model for a multi-agent system which describes how the knowledge and beliefs of agents should change when they receive new information. The formal tool we use for this description is a combination of modal and dynamic logic. Two core notions in our model of belief update are the *expansion of the knowledge and beliefs of an agent*, and *the processing of new information by an agent*. An expansion has been defined as the change in the knowledge and beliefs of an agent when it decides to believe an incoming formula while holding on to its current propositional beliefs. To prevent our agents from forming inconsistent beliefs they do not expand with every piece of information they receive. Instead of that, our agents remember their original beliefs (the beliefs they had before receiving any information) and every piece of information they receive. After every receipt of information they decide which (consistent) subset of the received information should be incorporated into their original beliefs. The decision which subset should be used is based on the trust the agents have in the sources of the received information. This procedure is called the processing of new information. We show that our model of belief update behaves in an intuitive way and that it is not sensitive to the critique on comparable models.

1 Introduction

In this paper we describe a model for the dynamic doxastic and epistemic behaviour of rational agents. Our main goal is to give a model for a multi-agent system which describes how the knowledge and beliefs of agents should change when they receive new information.

There are two reasons why a formal description of such behaviour of rational agents is useful. First of all, real world implementations of agent systems will most probably be very complex. Since tests never can guarantee the correct working of such systems, it is desirable to prove their correctness in a formal way. In order to do so, a formal description of their behaviour is essential. Secondly, agents are equipped with features, like knowledge and belief, of which

*e-mail: jwroorda@math.uu.nl

no consensus about their exact meaning exists. In order to unambiguously define their characteristics, a formal specification is needed.

The formal tool that we use to describe the behaviour of agents is a combination of modal and dynamic logic. We assume that the reader is familiar with both these logics and furthermore, that he knows how they can be used to describe a multi-agent system. If not, the book of [Goldblatt, 1992] is suggested as an introduction to modal and dynamic logic. A good introduction to how these logics can be used to describe multi-agent systems is found in [Hoek *et al.*, 1999].

A well known model of belief update in a model context is given by [Hoek *et al.*, 1999] and [Linder *et al.*, 1995]. This model is part of the KARO-framework for the specification of the behaviour of rational agents. Critique on this model, given by others, has served as inspiration for this paper. We give an alternative model that is not sensitive to this critique.

As stated before, our main goal is to describe how in a multi-agent system, the knowledge and beliefs of agents should change when they receive new information. In order to attain this goal we first need to explain how we model knowledge and beliefs of agents. We do so in chapter two.

After we have described how we model knowledge and belief, we give a precise description of what we mean by updating the beliefs of an agent. Referring to the definition of [Halpern and Friedman, 1998], we call this the *ontology* of our model.

Two core notions in our ontology are *the expansion of the knowledge and beliefs of an agent by a propositional formula* ϕ and *the processing of new information by an agent*. We define an expansion as the change in the knowledge and the beliefs of an agent when it decides to believe an incoming formula while holding one to its current propositional beliefs. In chapter four we describe formally how the knowledge and beliefs of an agent should change upon expanding.

It is important to stress that we believe that an agent not always should perform a belief expansion when receiving a piece of new information. In general a newly received piece of information should compete with previously received pieces of information to be expanded in the beliefs of the agent. We will describe the processing of new information in chapter five.

In chapter six we restate critique that has been given by others on the KARO-model, we show that our model is not sensitive to it.

We end this article with our conclusions and suggestions for further research.

2 Models of knowledge and belief

To model the knowledge and beliefs of the agents in our multi-agent system we use Kripke models. In this section we give the syntax and semantics of our formal system. Furthermore we give some useful definitions and lemmas.

2.1 Syntax

Definition 2.1 (Language)
Given a finite set $\mathcal{A} = \{1, ..., n\}$ of agents and a set $\mathcal{P}$ of propositional symbols we define the language $\mathcal{L}$ as follows.

$\mathcal{L}$ is the smallest set containing $\mathcal{P}$ such that

- $\phi, \psi \in \mathcal{L} \Rightarrow \neg\phi, \phi \wedge \psi \in \mathcal{L}$
- $i \in \mathcal{A}, \phi \in \mathcal{L} \Rightarrow \mathsf{K}_i\phi, \mathsf{B}_i\phi \in \mathcal{L}$

$\mathcal{L}_0$ is the propositional subset of $\mathcal{L}$.

2.2 Semantics

Definition 2.2 (Kripke models)
The class $\mathcal{M}$ of Kripke models contains tuples $M = (S, \pi, \mathrm{R}, \mathsf{B})$.

where

- *S is a set of possible states,*
- *$\pi : S \to \mathcal{P} \to \{0, 1\}$ is a function that assigns truth values to the propositional symbols in the possible worlds,*
- *$\mathrm{R} : \mathcal{A} \to S \times S$ is a function that assigns to each agent its epistemic accessibility relation,*
- *$\mathrm{B} : \mathcal{A} \to S \times S$ is a function that assigns to each agent its doxastic accessibility relation.*

To achieve the right properties of knowledge and belief we restrict the class of models $\mathcal{M}$ to $\mathcal{M}^*$.

Definition 2.3 (Knowledge and belief models)
The class $\mathcal{M}^*$ of knowledge and belief models is the greatest subset of the class of Kripke models $\mathcal{M}$ such that for every $M = (S, \pi, \mathrm{R}, \mathsf{B}) \in \mathcal{M}^*$ holds

- *$\mathrm{R} : \mathcal{A} \to S \times S$ is an equivalence relation,*
- *$\mathrm{B} : \mathcal{A} \to S \times S$ is a serial, transitive and euclidean relation,*
- *for every $i \in \mathcal{A}$ holds $\mathrm{B}_i \subseteq \mathrm{R}_i$*
- *for every $i \in \mathcal{A}$ and $u, s, t \in S$ holds $\mathrm{R}_i(s, t) \Rightarrow (\mathrm{B}_i(s, u) \Leftrightarrow \mathrm{B}_i(t, u))$.*

We use the same system for our models of knowledge and belief as [Linder *et al.*, 1995], whose system in turn is based on [Kraus and Lehmann, 1988].

We make a distinction between the formulas that an agent *knows* and the formulas that an agent *believes*. We require that known facts are true, (an agent can however believe formulas that are false), and that our agents have positive and negative introspection for both knowledge and belief. Furthermore, we require that the agents know their own beliefs. Finally, we require that every known formula is believed.

We achieve these goals by, first of all, requiring that the epistemic accessibility relations are equivalence relations; we use the standard S5-model for knowledge. Furthermore, we demand that the doxastic accessibility relations are serial, transitive and euclidean relations relations, so that we end up the standard KD-45 model for belief. Finally, to achieve the wanted relation between knowledge and beliefs, we require that the doxastic relations are included in the epistemic relations, and that for every $i \in \mathcal{A}$ and $u, s, t \in S$ holds $\mathrm{R}_i(s,t) \Rightarrow (\mathrm{B}_i(s,u) \Leftrightarrow \mathrm{B}_i(t,u))$

Definition 2.4 (Truth in Kripke models)
Let $M = (S, \pi, \mathrm{R}, \mathsf{B})$ be a Kripke model and $s \in S$. We define:

$$\begin{array}{lcl}
(M,s) \models p \in \mathcal{P} & \Leftrightarrow & \pi(s)(p) = 1 \\
(M,s) \models \phi \wedge \psi & \Leftrightarrow & (M,s) \models \phi \text{ and } (M,s) \models \psi \\
(M,s) \models \neg\phi & \Leftrightarrow & (M,s) \not\models \neg\phi \\
(M,s) \models \mathsf{K}_i\phi & \Leftrightarrow & \forall t \in S[\mathrm{R}_i(s,t) \Rightarrow (M,t) \models \phi)] \\
(M,s) \models \mathsf{B}_i\phi & \Leftrightarrow & \forall t \in S[\mathrm{B}_i(s,t) \Rightarrow (M,t) \models \phi)]
\end{array}$$

2.3 Elementary equivalence and bisimulation

Definition 2.5 (Depth of a formula)
The depth of a formula $\phi \in \mathcal{L}$, notation $\mathrm{depth}(\phi)$, is defined as follows:

$$\begin{array}{lcl}
\mathrm{depth}(p) & = & 0 \\
\mathrm{depth}(\neg\phi) & = & \mathrm{depth}(\phi) \\
\mathrm{depth}(\phi \wedge \psi) & = & \max(\mathrm{depth}(\phi), \mathrm{depth}(\psi)) \\
\mathrm{depth}(\mathsf{K}_i\phi) & = & 1 + \mathrm{depth}(\phi) \\
\mathrm{depth}(\mathsf{B}_i\phi) & = & 1 + \mathrm{depth}(\psi)
\end{array}$$

Definition 2.6 (Elementary equivalence)
Two Kripke worlds, say (M,s) and (M',s'), are called elementary equivalent, notation $(M,s) \equiv (M',s')$ if, and only if, for every $\phi \in \mathcal{L}$ $(M,s) \models \phi \Leftrightarrow (M',s') \models \phi$

Two Kripke models, say (M,s) and (M',s'), are called n-elementary equivalent, notation $(M,s) \equiv_n (M',s')$ if, and only if, for every $\phi \in \mathcal{L}$ with $\mathrm{depth}(\phi) \leq n$ $(M,s) \models \phi \Leftrightarrow (M',s') \models \phi$

Definition 2.7 (Bounded bisimulation)
Let (M,s) and (M',s') be two Kripke worlds, say $M = (S, \pi, \mathrm{R}, \mathrm{B})$ and $M' = (S', \pi', \mathrm{R}', \mathrm{B}')$. Let n be a natural number. We call (M,s) and (M',s') 0-bisimular, notation $(M,s) \simeq_0 (M',s')$, if, and only if, $\pi(s) = \pi'(s')$. We call (M,s) and (M',s') $n+1$-bisimular, notation $(M,s) \simeq_{n+1} (M',s')$, if, and only if

- $\pi(s) = \pi'(s')$
- $\forall i \forall t[\mathrm{R}_i(s,t) \Rightarrow \exists t'(\mathrm{R}'_i(s',t') \wedge (M,t) \simeq_n (M',t'))]$ *(forward choice for* R*)*
- $\forall i \forall t'[\mathrm{R}'_i(s',t') \Rightarrow \exists t(\mathrm{R}_i(s,t) \wedge (M,t) \simeq_n (M',t'))]$ *(backward choice for* R*)*

- $\forall i \forall t[\mathrm{B}_i(s,t) \Rightarrow \exists t'(\mathrm{B}'_i(s',t') \wedge (M,t) \simeq_n (M',t'))]$ *(forward choice for* B*)*
- $\forall i \forall t'[\mathrm{B}'_i(s',t') \Rightarrow \exists t(\mathrm{B}_i(s,t) \wedge (M,t) \simeq_n (M',t'))]$ *(backward choice for* B*)*

Proposition 2.1 $(M,s) \equiv_n (M',s') \Leftrightarrow (M,s) \simeq_n (M',s')$

Proof: Like in [Gerbrandy, 1998].

3 Ontology

In their article "Belief Revision: A Critique" [Halpern and Friedman, 1998] describe the *ontology* of a belief change process as a careful description what it means for something to be believed [or known] by an agent, and what the status is of incoming information. In this section we will give a description of the ontology of our model of belief update.

3.1 Believing and knowing

Our model of knowledge and belief is based on the notion of possible worlds. We say that an agent believes, respectively knows, the formula ϕ if ϕ is valid in all the worlds it considers doxasticaly, respectively epistemicaly, possible.

The difference between the the notions of knowing and believing is that we require that known formulas are true, whilst an agent can have false beliefs. The reason that we equip our agents with beliefs, is that they are needed for (common sense) reasoning. If our agents would only be equipped with knowledge, it would be impossible for them to reason with information from sources that are not hundred percent reliable. And because in real world applications it is not often the case that agents are confronted with absolutely reliable information, it would pose severe restrictions on their reasoning capabilities.

One of the reasons that we equip our agents with knowledge is that it can act as a certain lower bound for beliefs, which can be used to filter incoming information. For instance, when we know that in the world the agent is going to live in certain formulas are, and will remain, true, we can incorporate them in the knowledge of our agent. When the agent receives incoming information that contradicts these formulas, it will not use them to update his beliefs, because it knows that the incoming information cannot be true.

3.2 Incoming information

Our agents live in a system in which they receive information from different sources. They will use the information they receive from these sources to try to attain a more precise picture of the world they live in.

Two core notions in our ontology are *the expansion of the knowledge and beliefs of an agent by a propositional formula* ϕ and *the processing of new information* *by an agent.* We define expansion as the change in the knowledge and beliefs of an agent when it decides to believe an incoming formula while holding on to its current propositional beliefs.

In chapter four we will describe in an exact way how the beliefs and the knowledge set of an agent should change upon expanding. But for the remainder to make sense, it is needed to give some characteristics of expanding.

So, let us concentrate for a moment on what should happen with the knowledge and beliefs of an agent, when it decides to believe an incoming propositional formula ϕ while holding on to its current propositional beliefs. Because the agent decides to believe the formula ϕ it should be the case that that agent no longer considers worlds doxasticaly possible in which ϕ is not true. However, because incoming information need not to be true, and our notions of knowledge incorporate that an agent should only have true knowledge, these worlds should remain epistemicaly possible. A restriction on the doxasticaly possible worlds is not enough, because we would like that in every world that the agent considers (doxasticaly or epistemicaly) possible, it believes that ϕ is true. This ensures that we retain positive introspection. We will not go further into these questions now, for the remainder of this section it is enough to know that after an expansion with ϕ the agent believes that ϕ, but not necessarily knows that ϕ.

An essential assumption in our ontology is that when an agent receives new information, it has the choice whether it will use this information to expand his beliefs. The reason for this is the following. Our agents live in a world in which they receive information from a wide variety of sources. The information that these sources provide need not be true. So it can be the case that an agent receives contradicting information. If an agent should expand his knowledge and beliefs by every piece of information it receives, it could end up with inconsistent beliefs. Of course this is unwanted.

To prevent an agent from forming inconsistent beliefs, we let the agent decide which pieces of received information it will use to expand its beliefs. This decision will be based on the trust the agent has in the sources of the received information. An essential aspect of this decision procedure is that if agent chooses to use a certain piece of information, he is able to review this choice later on. For instance, suppose an agent starts its life with knowledge K and beliefs B. Now, if an agent receives a piece of information, say ϕ, it can choose to expand his knowledge and beliefs with ϕ, which yields, say $K \star \phi$ and $B \star \phi$. But, if it later on receives an other piece of information, say ψ, it can decide that the update of ϕ was not wanted after all, and use only ψ to update his *original* knowledge K and *original* beliefs B, which yields $K \star \psi$ and $B \star \psi$.

Another core assumption in our ontology is that an agent starts its life without proper beliefs. By proper believing a formula we mean that the agent believes, but does not know, the formula. The reason that we use this assumption, is that we consider precisely those propositional formulas that an agents knows as unquestionable. Every other propositional formula that an agent believes should have some kind of justification. In our model this justification consists of the fact the formula is believed because of an expansion with a certain piece of incoming information. If repeated expansion leads to inconsistent beliefs, the agent should decide by means of a selection function which of the performed expansion actions should be recovered.

4 The belief expansion function

4.1 Notation

Let (M, s) be a Kripke world, ϕ a formula of propositional logic, and a a label of an agent. By $(M, s)[\phi, a]$ we denote the Kripke world in which the beliefs of agent a are updated with the formula ϕ. We can think of $[\cdot, \cdot]$ as a function of type: $\mathcal{M} \times \mathcal{S} \times \mathcal{L}_0 \times \mathcal{A} \rightarrow \mathcal{M} \times \mathcal{S}$.

4.2 Desired characteristics

In this section we describe what characteristics a belief expansion function $[\cdot, \cdot]$ should have. We describe these characteristics in terms of the formulas an expanded Kripke world makes true.

Suppose that we update the knowledge and beliefs of agent a in the Kripke world (M, s) with the propositional formula ϕ.

First of all, the function should change *knowledge* and *beliefs*, not reality, so the truth of formulas consisting of propositional variables should not be changed.

Furthermore, the function should only change the knowledge and beliefs of agent a, and not those of other agents. So the truth of formulas $\mathsf{B}_b\alpha$ and $\mathsf{K}_b\alpha$ should not be changed for $a \neq b$.

The set of worlds which agent a considers doxasticaly possible should be changed. Agent a should consider worlds doxasticaly possible only if ϕ is true in them. Furthermore, the beliefs of agent a should be updated in the doxasticaly possible worlds too. So, it should be the case that in the updated Kripke world the formula $\mathsf{B}_a\alpha$ is true if, and only if, for all states t holds, [if [t is a doxasticaly possible world in the original Kripke world (M, s) and (M, t) models ϕ] then [(M, t) expanded with ϕ models α]].

Updating the beliefs of agent a should not change the worlds it considers epistemicaly possible. But of course the agent should know that its beliefs are updated in the worlds it considers epistemicaly possible. So the updated model should satisfy the formula $\mathsf{K}_a\alpha$ if, and only if, for all states t that agent a considers epistemicaly possible, (M, t) expanded with ϕ models α.

Finally, the function should behave naturally towards $\wedge$ and $\neg$.

We will call a function which satisfies these criteria *proper*.

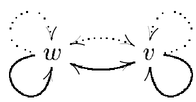
All this is summarized below.

Definition 4.1 (Proper belief expansion functions)
A belief expansion function $[\cdot,\cdot] : \mathcal{M} \times \mathcal{S} \times \mathcal{L}_0 \times \mathcal{A} \to \mathcal{M} \times \mathcal{S}$ *is called* proper *if*

$$
\begin{array}{lll}
(M,s)[\phi,a] \models p & \Leftrightarrow & (M,s) \models p \\
(M,s)[\phi,a] \models \alpha \wedge \beta & \Leftrightarrow & (M,s)[\phi,a] \models \alpha \text{ and } (M,s)[\phi,a] \models \beta \\
(M,s)[\phi,a] \models \neg\alpha & \Leftrightarrow & (M,s)[\phi,a] \not\models \alpha \\
(M,s)[\phi,a] \models \mathsf{B}_b\alpha & \Leftrightarrow & (M,s) \models \mathsf{B}_b\alpha \\
(M,s)[\phi,a] \models \mathsf{B}_a\alpha & \Leftrightarrow & \forall t((\mathrm{B}_a(s,t) \wedge (M,t) \models \phi) \Rightarrow (M,t)[\phi,a] \models \alpha) \\
(M,s)[\phi,a] \models \mathsf{K}_b\alpha & \Leftrightarrow & (M,s) \models \mathsf{K}_b\alpha \\
(M,s)[\phi,a] \models \mathsf{K}_a\alpha & \Leftrightarrow & \forall t((\mathrm{R}_a(s,t) \Rightarrow (M,t)[\phi,a] \models \alpha)
\end{array}
$$

4.3 An example

Let (M,w) be the Kripke world described by the following picture. To keep things simple, the world is inhabited by only one agent. In the picture the doxastic accessibility relation of this agent is reflected by the continuous arrows, and its epistemic accessibility relation by the dotted arrows.



$$w \models p, v \models \neg p$$

Suppose we update the beliefs of the agent in world w with p. According to our definition the new Kripke world $(M,w)[p,a]$ models the following formulas:

- $\mathsf{B}p$
- $\mathsf{BB}p$
- $\neg\mathsf{K}p$
- $\mathsf{KB}p$

So we have that after the update agent a believes that p, it believes that it believes that p. It does not know that p, but it knows that it believes that p.

4.4 Properties of proper belief expansion functions

In this section we give some properties of proper belief expansion functions, they serve to strengthen the opinion that proper belief expansion functions behave in an intuitive way.

Proposition 4.1 (Properties of proper belief expansion functions)

For all Kripke models $(M,s) = ((S,\pi,\mathrm{R},\mathrm{B}),s)$, propositional formulas ϕ,ψ, agents $a,b \in \mathcal{A}$ and proper belief expansion functions $[\cdot,\cdot]$ the following holds:

$(M, s)[\phi, a] \models \mathsf{B}_a\phi$
$(M, s)[\phi, a] \models \mathsf{B}_a\mathsf{B}_a\phi$
$(M, s)[\phi, a] \models \mathsf{K}_a\mathsf{B}_a\phi$
$(M, s)[\phi, a] \models \mathsf{K}_a\psi \quad \Leftrightarrow \quad (M, s) \models \mathsf{K}_a\psi$

The first clause states that after the expansion of the beliefs of agent a with ϕ it believes that ϕ. The second and third clause express that after the expansion the agent knows and believes that it believes ϕ. The fourth clause states that the knowledge of propositional formulas is not changed by expansions.

Proof:
We only prove the first two clauses. By definition of proper expansion functions the first clause states $\forall t((\mathrm{B}_a(s, t) \wedge (M, t) \models \phi) \Rightarrow (M, t)[\phi, a] \models \phi)$. Because ϕ is propositional the righthand side is equivalent to $(M, t) \models \phi$. By using the definition of proper belief expansion functions again, we see that the second clause states that $\forall t((\mathrm{B}_a(s, t) \wedge (M, t) \models \phi) \Rightarrow (M, t)[\phi, a] \models B_a\phi)$. But the righthand side is just the first clause.

4.5 Explicit definition of the proper belief expansion function

A nice property of our definition of proper belief expansion functions is that it defines a unique function. This is proved below.

Proposition 4.2 *There is, up to elementary equivalence, at most one proper belief expansion function*

Proof:
Let $[\cdot, \cdot]$ and $[\cdot, \cdot]'$ be two proper belief expansion functions. We have to prove that for all Kripke worlds (M, s), for all propositional formulas ϕ, for all agents a that $(M, s)[\phi, a] \equiv (M, s)[\phi, a]'$.

Suppose a propositional formula ψ is given. We will prove by induction on the complexity of ψ that for all Kripke worlds (M, s) $(M, s)[\phi, a] \models \psi \Leftrightarrow (M, s)[\phi, a]' \models \psi$. The basic case is trivial. Suppose that the proposition holds for formulas α and β. The induction steps for $\psi = \alpha \wedge \beta$, $\psi = \neg\alpha$, $\psi = \mathsf{B}_b\alpha$ and $\psi = \mathsf{K}_b\alpha$ are trivial. Suppose $\psi = \mathsf{B}_a\alpha$, then $\forall t((\mathrm{B}_a(s, t) \wedge (M, t) \models \phi) \Rightarrow (M, t)[\phi, a] \models \alpha)$. The i.h. states that $(M, t)[\phi, a] \models \alpha \Leftrightarrow (M, t)[\phi, a]' \models \alpha$. So $(M, s)[\phi, a]' \models \mathsf{B}_a\alpha$. Suppose $\psi = \mathsf{K}_b\alpha$, then $\forall t(\mathrm{R}_a(s, t) \Rightarrow (M, t)[\phi, a] \models \alpha)$. The i.h. states that $(M, t)[\phi, a] \models \alpha \Leftrightarrow (M, t)[\phi, a]' \models \alpha$. So $(M, s)[\phi, a]' \models \mathsf{K}_a\alpha$.

We define *the* proper belief expansion function below.

Definition 4.2 (The proper belief expansion function)
Define $[\cdot,\cdot]_p$ *as follows:*

$(M, s)[\phi, a]_p = (M', s')$

where

$$
\begin{array}{lcl}
M & = & (S, \pi, \mathrm{R}, \mathrm{B}) \\
M' & = & (S', \pi', \mathrm{R}', \mathrm{B}') \\
\mathrm{new} : S \to S \times \{0\} & = & \lambda s.(s, 0) \\
\mathrm{old} : S \to S \times \{1\} & = & \lambda s.(s, 1) \\
S' & = & \mathrm{old}(S) \bigcup \mathrm{new}(S) \\
s' & = & \mathrm{new}(s) \\
\pi' & = & \{(\mathrm{old}(u), p, \pi(u, p)) | u \in S, p \in \mathcal{P}\} \cup \\
 & & \{(\mathrm{new}(u), p, \pi(u, p)) | u \in S, p \in \mathcal{P}\}
\end{array}
$$

and the B_i *and* R_i *are such that:*

$$
\begin{array}{lcll}
\mathrm{B}'_i(\mathrm{old}(u), \mathrm{old}(v)) & \Leftrightarrow & \mathrm{B}_i(u, v) & i \in \mathcal{A} \\
\mathrm{B}'_a(\mathrm{new}(u), \mathrm{new}(v)) & \Leftrightarrow & \mathrm{B}_a(u, v) \wedge (M, v) \models \phi & \\
\mathrm{B}'_b(\mathrm{new}(u), \mathrm{old}(v)) & \Leftrightarrow & \mathrm{B}_b(u, v) & (a \neq b) \\
\mathrm{B}'_a(\mathrm{new}(u), \mathrm{old}(v)) & \Leftrightarrow & \bot & \\
\mathrm{B}'_i(\mathrm{old}(u), \mathrm{new}(v)) & \Leftrightarrow & \bot & i \in \mathcal{A} \\
\mathrm{B}'_b(\mathrm{new}(u), \mathrm{new}(v)) & \Leftrightarrow & \bot &
\end{array}
$$

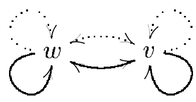
$$
\begin{array}{lcll}
\mathrm{R}'_i(\mathrm{old}(u), \mathrm{old}(v)) & \Leftrightarrow & \mathrm{R}_i(u, v) & i \in \mathcal{A} \\
\mathrm{R}'_a(\mathrm{new}(u), \mathrm{new}(v)) & \Leftrightarrow & \mathrm{R}_a(u, v) & \\
\mathrm{R}'_b(\mathrm{new}(u), \mathrm{old}(v)) & \Leftrightarrow & \mathrm{R}_b(u, v) & (a \neq b) \\
\mathrm{R}'_a(\mathrm{new}(u), \mathrm{old}(v)) & \Leftrightarrow & \bot & \\
\mathrm{R}'_i(\mathrm{old}(u), \mathrm{new}(v)) & \Leftrightarrow & \bot & i \in \mathcal{A} \\
\mathrm{R}'_b(\mathrm{new}(u), \mathrm{new}(v)) & \Leftrightarrow & \bot &
\end{array}
$$

The intuition behind this definition is, when expanding with ϕ, we make two copies of the original Kripke world. The one containing the states $\mathrm{old}(S)$ is an exact copy of the original Kripke world. Because, for all agents not equal to a, all from s' epistemicaly and doxasticaly accessible worlds, lie in this "old" copy, they see no difference between the original and the expanded Kripke world. Agent a, on the other hand, is able to see the "new" copy, where worlds not satisfying ϕ are no longer doxasticaly reachable.

This definition is based on the ideas in [Baltag *et al.*]. We extended their definition for the concept of knowledge.

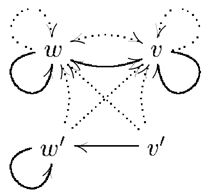
4.6 An example

We show for a specific Kripke world the result of an expansion via $[\cdot,\cdot]_p$. To make things easier we treat an example with only doxastic relations. Let (M, w) be the Kripke world described by the picture below. In the picture the doxastic accessibility relation of agent a is reflected by the continuous arrows, the doxastic accessibility relation of agent b by the dotted arrows.



$w \models p, v \models \neg p$

The expansion of the beliefs of agent a with p in state w yields the following Kripke world (M', w').



$w, w' \models p; v, v' \models \neg p$

4.7 Properties of $[\cdot,\cdot]_p$

Proposition 4.3 $[\cdot,\cdot]_p$ *is proper*

Proof:
Recalling definition 4.1 we should prove that for given propositional variable $p \in \mathcal{P}$, agents a and b such that $b \neq a$, formulas $\phi \in \mathcal{L}_0$ and formulas $\alpha, \beta \in \mathcal{L}$:

$$
\begin{array}{lcl}
(M,s)[\phi,a]_p \models p & \Leftrightarrow & (M,s) \models p \\
(M,s)[\phi,a]_p \models \alpha \wedge \beta & \Leftrightarrow & (M,s)[\phi,a]_p \models \alpha \text{ and } (M,s)[\phi,a]_p \models \beta \\
(M,s)[\phi,a]_p \models \neg\alpha & \Leftrightarrow & (M,s)[\phi,a]_p \not\models \alpha \\
(M,s)[\phi,a]_p \models \mathsf{B}_b\alpha & \Leftrightarrow & (M,s) \models \mathsf{B}_b\alpha \\
(M,s)[\phi,a]_p \models \mathsf{B}_a\alpha & \Leftrightarrow & \forall t((\mathrm{B}_a(s,t) \wedge (M,t) \models \phi) \Rightarrow (M,t)[\phi,a]_p \models \alpha) \\
(M,s)[\phi,a]_p \models \mathsf{K}_b\alpha & \Leftrightarrow & (M,s) \models \mathsf{K}_b\alpha \\
(M,s)[\phi,a]_p \models \mathsf{K}_a\alpha & \Leftrightarrow & \forall t((\mathrm{R}_a(s,t) \Rightarrow (M,t)[\phi,a]_p \models \alpha)
\end{array}
$$

The first item follows directly from the definitions of s' and π'. The second and third item follow directly from the semantical definitions of $\wedge$ and $\neg$ in modal logic.

The fourth claim is justified as follows. "$\Rightarrow$" Suppose $(M,s)[\phi,a]_p \models \mathsf{B}_b\alpha$. Let $(M,s)[\phi,a]_p = (M',s')$, as in the definition, then $(M',s') \models \mathsf{B}_b\alpha$. Hence $\forall t \in S'(\mathrm{B}'_b(s',t) \Rightarrow (M',t) \models \alpha)$. And thus by the definition of s' $\forall t \in S'(\mathrm{B}'_b(\mathrm{new}(s),t) \Rightarrow (M',t) \models \alpha)$. By observing that $\mathrm{B}'_b(\mathrm{new}(s),x)$ only is realized by $x \in \mathrm{old}(S)$ it follows that $\forall t \in S(\mathrm{B}'_b(\mathrm{new}(s),\mathrm{old}(t)) \Rightarrow (M',\mathrm{old}(t)) \models \alpha)$. Again by the definition of B'_b $\forall t \in S(\mathrm{B}_b(s,t) \Rightarrow (M',\mathrm{old}(t)) \models \alpha$. Observe that

$(M', \text{old}(t)) \equiv (M, t)$. So $\forall t \in S(\text{B}_b(s,t) \Rightarrow (M,t) \models \alpha$. So $(M,s) \models \mathsf{B}_b\alpha$. All the proof steps are reversible, so this is also a proof for "$\Leftarrow$".

The fifth claim is justified as follows. "$\Rightarrow$" Suppose $(M,s)[\phi,a]_p \models \mathsf{B}_a\alpha$. Then, by definition, $(M',s') \models \mathsf{B}_a\alpha$. Hence $\forall t \in S'(\text{B}'_a(s',t) \Rightarrow (M',t) \models \alpha)$. And thus by the definition of s' $\forall t \in S'(\text{B}'_a(\text{new}(s),t) \Rightarrow (M',t) \models \alpha)$. By observing that $\mathsf{B}'_a(\text{new}(s),x)$ only is realized by $x \in \text{new}(S)$ it follows that $\forall t \in S(\text{B}'_a(\text{new}(s),\text{new}(t)) \Rightarrow (M',\text{new}(t)) \models \alpha)$. Using the definition of B'_a we find that $\forall t \in S((\text{B}_a(s,t) \wedge (M,t) \models \phi) \Rightarrow (M',\text{new}(t)) \models \alpha)$. And thus $\forall t \in S((\text{B}_a(s,t) \wedge (M,t) \models \phi) \Rightarrow (M,t)[\phi,a]_p \models \alpha)$. This yields $(M,s) \models \mathsf{B}_a\alpha$. All proof steps are reversible, so this is also a proof for "$\Leftarrow$".

The sixth claim can be proven in the same as the fourth claim has been proved.

Finally, we prove the last claim. "$\Rightarrow$" Suppose $(M,s)[\phi,a]_p \models \mathsf{K}_a\alpha$. Then, by definition, $(M',s') \models \mathsf{K}_a\alpha$. Hence $\forall t \in S'(\text{R}'_a(s',t) \Rightarrow (M',t) \models \alpha)$. And thus by the definition of s' $\forall t(\text{R}'_a(\text{new}(s),t) \Rightarrow (M',t) \models \alpha)$. By observing that $\text{R}'_a(\text{new}(s),x)$ only is realized by $x \in \text{new}(S)$ it follows that $\forall t \in S(\text{R}'_a(\text{new}(s),\text{new}(t)) \Rightarrow (M',\text{new}(t)) \models \alpha)$. Using the definition of R'_a, we find that $\forall t \in S(\text{R}_a(s,t) \Rightarrow (M',\text{new}(t)) \models \alpha)$. And thus $\forall t \in S(\text{R}_a(s,t) \Rightarrow (M,t)[\phi,a]_p \models \alpha)$. This yields $(M,s) \models \mathsf{K}_a\alpha$. Again, all proof steps are reversible, so this is also a proof for "$\Leftarrow$".

Proposition 4.4 $[\cdot,\cdot]_p$ *is welldefined on the set of Kripke worlds modulo elementary equivalence*

Proof:
We will prove by induction on n that for every $n \in \mathbb{N}, (M,s),(M',s') \in \mathcal{M} \times \mathcal{S}, \phi \in \mathcal{L}_0$ and $a \in \mathcal{A}$ holds $[(M,s) \equiv_n (M',s') \Rightarrow (M,s)[\phi,a]_p \equiv_n (M',s')[\phi,a]_p]$.

The case $n = 0$ is trivial.

The induction hypothesis is that for every $(M,s),(M',s'),\phi,a$ holds $[(M,s) \equiv_n (M',s') \Rightarrow (M,s)[\phi,a]_p \equiv_n (M',s')[\phi,a]_p]$. (*)

We have to prove that for every $(M,s),(M',s'),\phi,a$ holds $[(M,s) \equiv_{n+1} (M',s') \Rightarrow (M,s)[\phi,a]_p \equiv_{n+1} (M',s')[\phi,a]_p]$.

Suppose $(M,s),(M',s'),\phi,a$ are given, s.t. $(M,s) \equiv_{n+1} (M',s')$, together with a formula ψ with $\text{depth}(\psi) \leq n+1$. We have to prove that $(M,s)[\phi,a]_p \models \psi \Leftrightarrow (M',s')[\phi,a]_p \models \psi]$. We prove this by induction on the structure of ψ.

The basic case is trivial.

The induction hypotheses is that $(M,s)[\phi,a]_p \models \alpha \Leftrightarrow (M',s')[\phi,a]_p \models \alpha$ and $(M,s)[\phi,a]_p \models \beta \Leftrightarrow (M',s')[\phi,a]_p \models \beta$.

We have to prove that for $\psi = \alpha\wedge\beta, \neg\alpha, \mathsf{B}_b\alpha, \mathsf{K}_b\alpha, \mathsf{B}_a\alpha, \mathsf{K}_a\alpha$ holds $(M,s)[\phi,a]_p \models \psi \Leftrightarrow (M',s')[\phi,a]_p \models \psi$

The cases for $\psi = \alpha \wedge \beta, \neg\alpha, \mathsf{B}_b\alpha, \mathsf{K}_b\alpha$ are trivial.

We will prove the last two cases. Because $\text{depth}(\psi) \leq n+1$, we have $\text{depth}(\alpha) \leq n$.

Suppose $(M,s)[\phi,a]_p \models \mathsf{B}_a\alpha$. Then $\forall t \in S(\mathrm{B}_a(s,t) \wedge (M,t) \models \phi \Rightarrow (M,t)[\phi,a]_p \models \alpha)$ (**). We have to prove that $(M',s')[\phi,a]_p \models \mathsf{B}_a\alpha$, i.e. $\forall t' \in S'(\mathrm{B}'_a(s',t') \wedge (M',t') \models \phi \Rightarrow (M',t')[\phi,a]_p \models \alpha)$.

Suppose that for some $t' \in S'$ $\mathrm{B}'_a(s',t') \wedge (M',t') \models \phi$ holds. We know $(M,s) \equiv_{n+1} (M',s')$, so by proposition 1.1 $(M,s) \simeq_{n+1} (M',s')$. So, by backward choice for B_a, there is a $t \in S$ such that $\mathrm{B}_a(s,t)$ and $(M,t) \simeq_n (M',t')$. So, for this t we have $(M,t) \equiv_n (M',t')$ (***). And for this t, we have $(M,t) \models \phi$. So by (**) we have $(M,t)[\phi,a]_p \models \alpha$. By (*) and (***) we have $(M,t)[\phi,a]_p \equiv_n (M',t')[\phi,a]_p$. So $(M',t')[\phi,a]_p \models \alpha$. So indeed $(M',s')[\phi,a]_p \models \mathsf{B}_a\alpha$.

Suppose $(M,s)[\phi,a]_p \models \mathsf{K}_a\alpha$. Then $\forall t \in S(\mathrm{R}_a(s,t) \Rightarrow (M,t)[\phi,a]_p \models \alpha)$ (**2). We have to prove that $(M',s')[\phi,a]_p \models \mathsf{K}_a\alpha$, i.e. $\forall t' \in S'(\mathrm{R}'_a(s',t') \Rightarrow (M',t')[\phi,a]_p \models \alpha)$.

Suppose that for some $t' \in S'$ $(\mathrm{R}'_a(s',t')$ holds. We know $(M,s) \equiv_{n+1} (M',s')$, so by proposition 1.1 $(M,s) \simeq_{n+1} (M',s')$. So, by backward choice for R_a, there is a $t \in S$ such that $\mathrm{R}_a(s,t)$ and $(M,t) \simeq_n (M',t')$. So, for this t we have $(M,t) \equiv_n (M',t')$ (***2). By (**2) we have $(M,t)[\phi,a]_p \models \alpha$. By (*) and (***2) we have $(M,t)[\phi,a]_p \equiv_n (M',t')[\phi,a]_p$. So $(M',t')[\phi,a]_p \models \alpha$. So indeed $(M',s')[\phi,a]_p \models \mathsf{K}_a\alpha$.

The above result is nice, but actually we want more. We want the result of an expansion of a knowledge and belief world to result in a knowledge and belief world. (Or at least to result in a world that is elementary equivalent to a knowledge and belief world.) I.e. we want that for every $(M,s) \in \mathcal{M}^* \times \mathcal{S}$, $\phi \in \mathcal{L}_0$, $a \in \mathcal{A}$: there exist $(M',s') \in \mathcal{M}^* \times \mathcal{S}$ such that $(M,s)[\phi,a] \equiv (M',s')$. Unfortunately, this is not the case, for example, suppose $(M,s) \in \mathcal{M}^* \times \mathcal{S}$, such that $(M,s) \models \mathsf{K}_b\neg\mathsf{B}_a\phi$, then $(M,s)[\phi,a] \models \mathsf{K}_b\neg\mathsf{B}_a\phi$ and $(M,s)[\phi,a] \models \mathsf{B}_a\phi$, i.e. agent b has false knowledge! A solution to this problem would be to restrict the knowledge and belief models to models in which agents do not have knowledge about knowledge and beliefs of other agents that is subject to change by expansions. Unfortunately, we have not succeeded in characterizing this class of models in terms of restrictions on the accessibility relations.

5 The process new information action

As described in the previous sections, our agents do not expand with every piece of information they receive. Instead of that, they remember their original beliefs (the beliefs they had before receiving any information) and every piece of information they receive. After every receipt of information they decide which (consistent) subset of the received information should be expanded into their original beliefs. The decision which subset should be used is based on the trust the agents have in the sources of the received information.

We call this procedure the process new information (p_n_i) action. An agent can use the forget information (f_i) action to undo the effects of a process new information action.

To model this actions we shift from a static modal logic to a dynamic modal logic. We describe its syntax and semantics in this section.

5.1 Dynamic Modal Logic

Definition 5.1 (Language)
Given a finite sets $\mathcal{A} = \{1, ..., n\}$ of agents and $\mathcal{SC} = \{I, II, III, ...\}$ of sources, and sets $\mathcal{P}$ of propositional symbols and At *of atomic actions, we define the language $\mathcal{L}$ and the class of actions* Ac *by mutual induction as follows.*

$\mathcal{L}$ is the smallest set containing $\mathcal{P}$ such that

- $\phi, \psi \in \mathcal{L} \Rightarrow \neg\phi, \phi \wedge \psi \in \mathcal{L}$
- $i \in \mathcal{A}, \alpha \in \mathsf{Ac}, \phi \in \mathcal{L} \Rightarrow \mathsf{K}_i\phi, \mathsf{B}_i\phi, \langle \mathsf{do}_i(\alpha)\rangle\phi, \mathsf{A}_i(\alpha) \in \mathcal{L}$

Ac *is the smallest set containing* At *such that*

- $\phi \in \mathcal{L} \Rightarrow \mathsf{confirm}\ \phi \in \mathsf{Ac}$
- $\alpha_1, \alpha_2 \in \mathcal{L} \Rightarrow (\alpha_1; \alpha_2) \in \mathsf{Ac}$
- $\phi \in \mathcal{L}, \alpha_1, \alpha_2 \in \mathsf{Ac} \Rightarrow (\mathsf{if}\ \phi\ \mathsf{then}\ \alpha_1\ \mathsf{else}\ \alpha_2\ \mathsf{fi}\ ) \in \mathsf{Ac}$
- $\phi \in \mathcal{L}, \alpha \in \mathsf{Ac} \Rightarrow (\mathsf{while}\ \phi\ \mathsf{do}\ \alpha\ \mathsf{od}) \in \mathsf{Ac}$
- $\phi \in \mathcal{L}_0, t \in \mathcal{SC} \Rightarrow \mathsf{p\_n\_i}(\phi, t) \in \mathsf{Ac}$
- $\phi \in \mathcal{L}_0, t \in \mathcal{SC} \Rightarrow \mathsf{f\_i}(\phi, t) \in \mathsf{Ac}$

Definition 5.2 (Kripke models)
The class $\mathcal{M}$ of Kripke models contains tuples $M = (S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel})$.

where

- $S, \pi, \mathrm{R}, \mathrm{B}$ *are as in definition 2.3,*
- $\mathrm{r} : \mathcal{A} \times \mathsf{At} \rightarrow \mathcal{M} \times \mathcal{S} \rightarrow \mathcal{M} \times \mathcal{S}$ *is such that $\mathrm{r}(i, a)(M, s)$ yields the Kripke world transition in (M, s) caused by the event $\mathsf{do}_i(a)$,*
- $\mathrm{c} : \mathcal{A} \times \mathsf{At} \rightarrow \mathcal{M} \times \mathcal{S} \rightarrow \{0, 1\}$ *is the capability function such that $c(i, a)(M, s)$ indicates whether the agent i is capable of performing the action a in s,*
- $M_o = ((S_o, \pi_o, \mathrm{R}_o, \mathrm{B}_o), s_o)$ *is called the original Kripke world, it models the knowledge and beliefs of the agents when they start their lives. The original Kripke world is a Kripke world as in definition 2.3, to ensure that the agents start their lives without proper beliefs, we demand that $\mathrm{R}_{o(i)} = \mathrm{B}_{o(i)}$ for all i,*
- H *is the received information function, of type: $\mathcal{A} \rightarrow [\mathcal{L}_0 \times \mathcal{SC}]$. The received information function gives for every agent a list of the formulas it received in the past, together with their sources,*
- Sel *is the information selection function of type $\mathcal{M} \times \mathcal{S} \rightarrow \mathcal{A} \rightarrow [\mathcal{L}_0 \times \mathcal{SC}] \rightarrow [\mathcal{L}_o]$, it selects for an agent which information it will use to form its knowledge and beliefs.*

We assume that the original knowledge and beliefs do not compromise proper beliefs. By proper believing a formula we mean that the agent believes, but does not know, the formula. The reason that we use this assumption, is that we consider precisely those propositional formulas that an agents knows as unquestionable. Every other propositional formula that an agent beliefs should have some kind of justification, in our model this justification will consist of the fact the formula is believed because of an expansion with a certain piece of incoming information. By this we ensure that our agents are always able to make the judgement whether a proper believed formula should be given up in favor of an incoming formula.

Definition 5.3 (Truth in Kripke models)
Let $M = (S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel})$ *be a Kripke model and* $s \in S$. *We define:*

$$\begin{array}{lcl}
(M,s) \models p \in \mathcal{P} & \Leftrightarrow & \pi(p) = 1 \\
(M,s) \models \phi \wedge \psi & \Leftrightarrow & (M,s) \models \phi \text{ and } (M,s) \models \psi \\
(M,s) \models \neg\phi & \Leftrightarrow & \text{not}(M,s) \models \neg\phi \\
(M,s) \models \mathsf{K}_i\phi & \Leftrightarrow & \forall t \in S[(s,t) \in \mathrm{R}_i \Rightarrow (M,t) \models \phi)] \\
(M,s) \models \mathsf{B}_i\phi & \Leftrightarrow & \forall t \in S[(s,t) \in \mathrm{B}_i \Rightarrow (M,t) \models \phi)] \\
(M,s) \models [\mathsf{do}_i(\alpha)]\phi & \Leftrightarrow & \forall (M',s')((M',s') \in \mathrm{r}(i,\alpha)(M,s) \Rightarrow (M',s') \models \phi) \\
(M,s) \models \mathsf{A}_i\alpha & \Leftrightarrow & \mathrm{c}(i,\alpha) = 1
\end{array}$$

where r and c are extended as follows:

$$\begin{array}{lcl}
\mathrm{r}(i, \mathsf{confirm}\ \phi)(M,s) & = & \{(M,s)\} \text{ if } (M,s) \models \phi \text{ and } \emptyset \text{ otherwise} \\
\mathrm{r}(i, \alpha_1; \alpha_2)(M,s) & = & (\mathrm{r}(i,\alpha_2) \circ \mathrm{r}(i,\alpha_1))(M,s) \\
\mathrm{r}(i, \mathsf{if}\ \phi\ \mathsf{then}\ \alpha_1\ \mathsf{else}\ \alpha_2\ \mathsf{fi}\ )(M,s) & = & \mathrm{r}(i,\alpha_1)(M,s) \text{ if}(M,s) \models \phi \\
 & & \text{and } \mathrm{r}(i,\alpha_2)(M,s) \text{ otherwise} \\
\mathrm{r}(i, \mathsf{while}\ \phi\ \mathsf{do}\ \alpha_1\ \mathsf{od})(M,s) & = & \{(M',s') | \exists k \in \mathbb{N}, \exists M_0, ..., M_k \in \mathcal{M}, \\
 & & \exists s_0, ..., s_k \in \mathcal{S}((M_0,s_0) = (M,s), \\
 & & (M_k, s_k) = (M', s') \\
 & & \& \forall j < k (M_{j+1}, s_{j+1}) \in \\
 & & \mathrm{r}(i, \mathsf{confirm}\ \phi; \alpha_1)(M_j, s_j)) \\
 & & \& (M', s') \models \neg\phi)\}
\end{array}$$

$$\begin{array}{lcl}
\mathrm{c}(i, \mathsf{confirm}\ \phi)(M,s) & = & 1, \text{ if } (M,s) \models \phi \text{ and } 0 \text{ otherwise} \\
\mathrm{c}(i, \alpha_1, \alpha_2)(M,s) & = & \mathrm{c}(i,\alpha_1) \text{ and } \mathrm{c}(i,\alpha_2)(\mathrm{r}(i,\alpha_1)(M,s)) \\
\mathrm{c}(i, \mathsf{if}\ \phi\ \mathsf{then}\ \alpha_1\ \mathsf{else}\ \alpha_2\ \mathsf{fi}\ )(M,s) & = & \mathrm{c}(i, \mathsf{confirm}\ \phi; \alpha_1) \text{ or } \mathrm{c}(i, \mathsf{confirm}\ \neg\phi; \alpha_2) \\
\mathrm{c}(i, \mathsf{while}\ \phi\ \mathsf{do}\ \alpha_1\ \mathsf{od})(M,s) & = & 1 \text{ if } \exists k \in \mathbb{N}(\mathrm{c}(i, (\mathsf{confirm}\ \phi; \alpha_1)^k; \\
 & & \mathsf{confirm}\ \neg\phi)(M,s) = 1) \\
 & & \text{and } 0 \text{ otherwise}
\end{array}$$

In the next section we complete the definition of the semantics by defining the process new information action.

5.2 The process new information action

Before we can define the semantics of the process new information action, we need to make a few technical definitions.

Definition 5.4 (Belief expansion on the new Kripke worlds)
We define the working of the belief expansion function on our new Kripke worlds as follows. Let $(M, s) = ((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s)$ *be a Kripke world.*

$(M, s)[\phi, a] = ((S', \pi', \mathrm{R}', \mathrm{B}', (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s')$

where $S', \pi', \mathrm{R}', \mathrm{B}'$ *and* s' *are such that* $((S, \pi, \mathrm{R}, \mathrm{B}), s)[\phi, a] = ((S', \pi', \mathrm{R}', \mathrm{B}'), s')$.

Definition 5.5 (Retrieving the for expansion selected information)
Given (M, s), H *and* Sel, *define:*

$\mathrm{Filter}(M, s)(\mathrm{Sel})(\mathrm{H}) = [(\phi, i) | \phi \in \mathrm{Sel}(M, s)(i)(\mathrm{H}(i)), i \in \mathcal{A}]$

Filter applied to a Kripke world, a selection function and a received information function yields a list of (agent, formula) tuples. The meaning of this list is the following: if (ϕ, i) is an element of this list then agent i has decided to expand its knowledge and beliefs by ϕ. This is purely a technical definition. In general the output of this function will be used as input for the belief expansion function, see the following definition.

Definition 5.6 (The belief expansion function on lists)
Let $U : [\mathcal{L}_0 \times \mathcal{A}]$, *define:*

$$\begin{array}{lcl} (M, s)[\,] & = & (M, s), \\ (M, s)[U] & = & ((M, s)[\phi, t])[U'], \text{ if } U = (\phi, t) : U' \end{array}$$

Definition 5.7 (The process new information action)
Let $M = (S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel})$ *be a Kripke model. Let* $s \in S, \phi \in \mathcal{L}_0$, $i \in A$ *and* $t \in \mathcal{SC}$. *We define:*

$\mathrm{r}(i, \mathsf{p\_n\_i}(\phi, t))(M, s) =$
$\{((S_o, \pi_o, \mathrm{R}_o, \mathrm{B}_o, (M_o, s_o), \mathrm{H}', \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s_o)[\mathrm{Filter}(M, s)(\mathrm{Sel})(\mathrm{H}')]\}$.

where H' *is such that* $\mathrm{H}'(i) = (\phi, t) : \mathrm{H}(i)$ *and* $\mathrm{H}'(j) = \mathrm{H}(j)$ *for* $j \neq i$.

furthermore:
$\mathrm{c}(i, \mathsf{p\_n\_i}(\phi, t))(M, s) = 1$ *for all* (M, s).

The intuition behind this definition is the following. After agent i has performed the $\mathsf{p\_n\_i}(\phi, t)$ action, the information (ϕ, t) will be added to its list of received information $\mathrm{H}(i)$, this yields $\mathrm{H}'(i)$. By means of the selection function Sel, it will decide which elements of the list of received information $\mathrm{H}'(i)$ it will use to form his knowledge and beliefs; this yields $\mathrm{Sel}(M, s)(i)(\mathrm{H}'(i))$. The model which reflects the new beliefs of agent i is the original model (which reflects its original knowledge and beliefs) expanded with the information in $\mathrm{Sel}(M, s)(i)(\mathrm{H}'(i))$. Because the new model also must reflect the knowledge and beliefs of the other agents (which have not changed) we also expand with the elements from $\mathrm{Sel}(j)(M, s)(\mathrm{H}(j))$ for $j \neq i$.

For this intuition to make sense, it must be case that the knowledge and beliefs of the agents are the result of expansions of their original knowledge and beliefs with the selected information from their received information list. This leads to the following definition.

Definition 5.8 (Proper Kripke worlds)
A Kripke world $((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s)$ *is called proper iff.*
$((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s) =$
$((S_o, \pi_o, \mathrm{R}_o, \mathrm{B}_o, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s_o)[\mathrm{Filter}(M, s)(\mathrm{Sel})(\mathrm{H})]$

In our definition of the process new information action we have made a rather arbitrary choice about the order in which we expand the Kripke world with the (agent,formula) tuples. The next proposition shows that this choice does not influence the meaning of the resulting Kripke world.

Proposition 5.1 $(M, s)[\phi, i][\psi, j] \equiv (M, s)[\psi, j][\phi, i]$

Proof:
The case for $i \neq j$ is trivial. The case for $i = j$ can be proved, by showing, by induction on the structure of α, that $(M, s)[\phi, i][\psi, i] \models \alpha \Leftrightarrow (M, s)[\phi \wedge \psi, i] \models \alpha$.

Definition 5.9 (The forget information action)
Let $M = (S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel})$ *be a Kripke model. Let* $s \in S, \phi \in \mathcal{L}_0$, $i \in A$ *and* $t \in \mathcal{SC}$ *s.t.* $(\phi, t) \in \mathrm{H}(i)$. *We define:*

$\mathrm{r}(i, \mathsf{f\_i}(\phi, t))(M, s) =$
$\{((S_o, \pi_o, \mathrm{R}_o, \mathsf{B}_o, (M_o, s_o), \mathrm{H}', \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s_o)[\mathrm{Filter}(M, s)(\mathrm{Sel})(\mathrm{H}')]\}$.

where H' *is such that* $\mathrm{H}'(i) = \mathrm{H}(i) - (\phi, t)$ *and* $\mathrm{H}'(j) = \mathrm{H}(j)$ *for* $j \neq i$.

furthermore:
$\mathrm{c}(i, \mathsf{f\_i}(\phi, t))(M, s) = 1$ *if, and only if,* $(\phi, t) \in \mathrm{H}(i)$.

The intuition behind this action is similar to the intuition behind the p_n_i-action. Only in this case a piece of information is deleted from the list of received information $\mathrm{H}(i)$. In applications this can be useful, for instance when an agent discovers that a particular piece of information did not make any sense at all, for instance due to a sensor-malfunction.

5.3 Characteristics of the information selection function

One of the reasons for introducing the selection function in our system, was to prevent our agents forming inconsistent beliefs. We will call a selection function which does this job, a proper selection function.

Definition 5.10 (Proper selection functions)
A information selection function $\mathrm{Sel} : \mathcal{M} \times \mathcal{S} \to \mathcal{A} \to [\mathcal{L}_0 \times \mathcal{SC}] \to [\mathcal{L}_0]$ *is called proper if for every proper Kripke world* $(M, s) = ((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s)$ *and agent* i *holds*

$(M_o, s_o) \not\models \mathsf{B}_i\bot \Rightarrow (M, s) \not\models \mathsf{B}_i\bot$.

Demanding that a selection function is proper is needed, but not enough to let our agents work in an intuitive way. To illustrate this, the trivial selection function given by $\mathrm{Sel}_t(M, s)(i)(\mathrm{H}(i)) = [\,]$ is proper, but by using this selection function our agents will never update their knowledge and beliefs. This leads to the definition of eager selection functions. We call a selection function which lets the agents use every piece of information as long as that does not lead to a contradiction an eager selection function.

Definition 5.11 (Eager selection functions)
A proper selection function $\mathrm{Sel} : \mathcal{M} \times \mathcal{S} \to \mathcal{A} \to [\mathcal{L}_0 \times \mathcal{SC}] \to [\mathcal{L}_0]$ *is called eager, if for every Kripke world* $(M, s) = ((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s)$ *and agent* i *holds*

$((S_o, \pi_o, \mathrm{R}_o, \mathrm{B}_o, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s_o)[\mathrm{Filter}(M, s)(\mathrm{Sel}_n)(H)] \not\models \mathsf{B}_i\bot$

$\Rightarrow$

$\mathrm{Sel}(M, s)(i)(H(i)) = \mathrm{Sel}_n(M, s)(i)(\mathrm{H}(i))$

where $\mathrm{Sel}_n(M, s)(i)[(\phi_0, t_0), ..., (\phi_k, t_k)] = [\phi_0, ..., \phi_k]$ *(the naive selection function)*

Because the definition of eager selection functions does not say anything about which information should be selected when an agent receives contradicting information, demanding that a selection function is eager still does not ensure that the selection function works in an intuitive way. But because further restrictions would imply arbitrary choices to a certain extent, and because we can prove a lot of nice properties of our system by only demanding that the selection function is eager, we do not suggest further restrictions.

5.4 Properties of our system

In this section we give properties of our system of belief update, they serve to strengthen the opinion that our system behaves in an intuitive way.

Proposition 5.2 (Properties of our system of belief update)

For every proper Kripke model $(M, s) = ((S, \pi, \mathrm{R}, \mathrm{B}, (M_o, s_o), \mathrm{H}, \mathrm{r}, \mathrm{c}, \mathrm{Sel}), s)$, that contains an eager selection function, $\phi \in \mathcal{L}_0$, $\psi \in \mathcal{L}$, $t \in \mathcal{SC}$ the following holds:

$$
\begin{array}{lll}
(M, s) \models \neg\mathsf{B}_i\neg\phi & \to & [\mathsf{do}_i(\mathsf{p\_n\_i}(\phi, t))]\mathsf{B}_i\phi \\
(M, s) \models \mathsf{B}_i\phi \wedge \neg\mathsf{B}_i\bot & \to & [\mathsf{do}_i(\mathsf{p\_n\_i}(\neg\phi, t))]\neg\mathsf{B}_i\bot \\
(M, s) \models \mathsf{B}_i\psi & \to & [\mathsf{do}_i((\mathsf{p\_n\_i}(\phi, t); (\mathsf{f\_i}(\phi, t))]\mathsf{B}_i\psi
\end{array}
$$

The first clause states that if agent a does not believe ϕ, before receiving ϕ, it will believe ϕ afterwards. The second clause states that if agent a does believe ϕ, receiving $\neg\phi$ will not lead to inconsistent beliefs. The third clause expresses that forgetting after processing information does not lead to a change in the beliefs of an agent.

For a selection function, that compromises that the trust of agent i in source I, exceeds the trust of the agent in source II, we have:

$(M,s) \models \neg \mathsf{B}_i \neg \phi \wedge \neg \mathsf{B}_i \phi \to [\mathsf{do}_i((\mathsf{p\_n\_i}(\phi, I); (\mathsf{p\_n\_i}(\neg \phi, II))]B_i \phi$

$(M,s) \models \mathsf{K}_j(\neg \mathsf{B}_i \neg \phi \wedge \neg \mathsf{B}_i \phi \to [\mathsf{do}_i((\mathsf{p\_n\_i}(\phi, I); (\mathsf{p\_n\_i}(\neg \phi, II))]\mathsf{B}_i \phi)$

The first clause states that agents give priority to information from more trustworthy sources. The second clause is an example of the fact that agents are aware of each other selection functions. This can or cannot be desired for a specific implementation of a multi-agent system. If this property is unwanted the notion of Kripke models should be changed such that different worlds can contain different selection functions. In this way one can model uncertainty about other agents' selection functions.

6 Critique on the KARO-model

For our model for belief update update we have been inspired by critique on the KARO-model. In this section we treat the main points of critique and show that our model is not sensitive to them.

6.1 Unwanted side effects to the beliefs of other agents

[Gerbrandy, 1998] shows that in the KARO-model an update of the beliefs of agent a leads to changes in the beliefs of other agents. This is unwanted because we assume that other agents are not aware of the update of agent a.

It follows directly from our definition of proper belief expansion functions that in our model the knowledge and beliefs of other agents remain fixed when an agent updates its beliefs.

6.2 Priority to incoming information

[Halpern and Friedman, 1998] claim that the principle of incoming information, which states that observing ϕ leads to believing ϕ, is to be doubted. They argue that this principle leads to counter intuitive results if the incoming information is received from sources that are not completely reliable. In that case the beliefs of an agent which receives contradicting information, are completely dependent on the order in which the information is received. Halpern and Friedman argue that this is wrong, they state that in stead of only relying on the order of the incoming information, incoming information should compete for acceptance against what is already believed. In order to describe a model in which this it the case, they state that "we need to describe how to combine a given strength of belief in the observation with the strengths of beliefs in the original epistemic state".

This is exactly what we have done in our model of belief update. If an agent (proper) believes something, it is the result of a belief expansion with information from a certain source. If an agent receives incoming information that contradicts its current (proper) believes, it will decide by means of its selection function which of the sources is more trustworthy.

More or less the same point of critique is made by [Dragoni *et al.*, 1994]. They argue that "in a multi-agent environment, where information comes from a variety of [..] sources with different degrees of reliability [..] the chronological sequence of the informative acts has nothing to do with their credibility or importance".

7 Conclusion

In this paper we have given a model of belief update which describes how the knowledge and beliefs of agents should change when they receive new information. The formal tool that we have used for this description is a combination of modal and dynamic logic.

Two core notions in our ontology of belief update are the *expansion of the knowledge and beliefs of an agent with a propositional formula*, and *the processing of new information by an agent*. An expansion has been defined as the change in the knowledge and beliefs of an agent when it decides to believe an incoming formula while holding on to its current propositional beliefs.

To prevent our agents from forming inconsistent beliefs they do not expand their knowledge and beliefs with every piece of information they receive. In stead of that they remember their original beliefs (the beliefs they had before receiving any information) and every piece of information they receive. After every receipt of information they decide which (consistent) subset of the received information should be expanded into their original beliefs. The decision which subset should be used is based on the trust the agents have in the sources of the received information.

We have shown that our model of belief update behaves in an intuitive way and that it is not sensitive to the critique on the KARO-model.

We have limited ourselves to the updates of the knowledge and beliefs of a single agent with a propositional formula. A suggestion for further research would be to extend the model to updates of knowledge and beliefs of groups of agents with modal formulas.

The knowledge and belief models that we defined in 2.3 are not closed under expansion. This is easy to see; there is a knowledge and belief model that entails that agent a knows that agent b does not believe ϕ. Expansion of the knowledge and beliefs of agent b with ϕ leads to false knowledge of a. We have not been able to define restrictions on the accessibility relations of our knowledge and belief models, that yield the greatest subset that is closed under expansion. This could be a topic of further research.

References

[Baltag *et al.*] Alexandru Baltag, Lawrence S. Moss and Slawomir Solecki (to appear), *The Logic of Public Announcements and Common Knowledge*, http://www.math.indiana.edu/home/moss/articles.html.

[Dragoni *et al.*, 1994] Aldo Franco Dragoni and Paolo Giorgini, *Belief Revision Through the Belief Function Formalism in a Multi-Agent Environment*, volume 1193 of *Lecture Notes in Computer Science (subseries LNAI)*, Springer Verlag, 1994.

[Gerbrandy, 1998] Jelle Gerbrandy, *Bisimulations on Planet Kripke*, Ph.D. Thesis, Universiteit van Amsterdam, 1998.

[Goldblatt, 1992] R. Goldblatt, *Logics of Time and Computation*, volume 7 of *CSLI Lecture Notes*, Standford, 1992.

[Halpern and Friedman, 1998] Jospeh Y. Halpern and Nir Friedman, *Belief Revision: A Critique*, International Conference on Knowledge Representation and Reasoning, 421-431, 1996. http://simon.cs.cornell.edu/Info/People/halpern/, 1998.

[Hoek *et al.*, 1999] W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer, *An Integrated Modal Approach to Rational Agents* in *Foundations of Rational Agency*, M.Woolridge and A.Rao (eds.), 133-167, 1999

[Kraus and Lehmann, 1988] S.Kraus, D. Lehman, Knowledge, belief and time, *Theoretical Computer Science*, 58:155-174, 1988.

[Linder *et al.*, 1995] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer, *Actions That Make You Change Your Mind* in *Knowledge and Belief in Philosophy and Artificial Intelligence*, Armin Laux, Heinrich Wansing (eds.), 104-146, 1995