

Seeing is Believing And so are Hearing and Jumping

B. van Linder W. van der Hoek J.-J. Ch. Meyer
Utrecht University
Department of Computer Science
P.O. Box 80.089
3508 TB Utrecht
The Netherlands
Email: bernd@cs.ruu.nl

Abstract

In this paper a formal framework is proposed in which various informative actions are combined, corresponding to the different ways in which rational agents can acquire information. In order to solve the various conflicts that could possibly occur when acquiring information from different sources, we propose a classification of the information that an agent possesses according to credibility. Based on this classification, we formalize what it means for agents to have seen or heard something, or to believe something by default. We present a formalization of observations, communication actions, and the attempted jumps to conclusions that constitutes default reasoning. To implement these informative actions we use a general belief revision action which satisfies the AGM postulates; dependent on the credibility of the incoming information this revision action acts on one or more parts of the classified belief sets of the agents. The abilities of agents formalize both the limited capacities of agents to acquire information, and the preference of one kind of information acquisition to another. The framework presented in this paper might be seen as an attempt to formalize the behavior of intelligent information agents which will facilitate life in the information age.

*Content Areas: Formalizing Rational Agents, Belief Revision,
Reasoning about Action, Knowledge Representation.*

1. Introduction

The formalization of rational agents is a topic of continuing interest in Artificial Intelligence. Research on this subject has held the limelight ever since the pioneering work of Moore [36, 37] in which knowledge and actions are considered. Over the years important contributions have been made on both *informational* aspects like knowledge and belief [14, 35], and *motivational*¹ aspects like commitments and obligations [7]. Recent developments include the work

¹The notions *informational* and *motivational* are both due to Shoham & Cousins [42].

on agent-oriented programming [41, 43], the Belief-Desire-Intention architecture [38, 39, 40], logics for the specification and verification of multi-agent systems [44, 45], logics for agents with bounded rationality [21, 22], and cognitive robotics [26, 27].

In our research [18, 19, 29, 30, 31, 32, 33] we defined a *theorist* logic for rational agents, i.e., a logic that is used to *specify*, and to *reason about*, (various aspects of) the behavior of rational agents. We consider both informational and motivational attitudes, and furthermore pay attention to various aspects of actions. In the basic architecture the *knowledge*, *belief* and *abilities* of agents, as well as the *opportunities* for and the *results* of their actions are formalized. In this framework it can, for instance, be modelled that an agent knows that some action is *correct* to achieve some state of affairs since it knows that performing the action will lead to this state, and that it knows that an action is *feasible* in the sense that the agent knows of its ability to perform the action. In subsequent research we extended our framework with nondeterministic actions [19], epistemic tests [32], communicative actions [30], actions that model default reasoning [31], actions that model belief revision [29], and a formalization of motivational attitudes [33]. Dunin-Keplicz & Radzikowska [8] use our framework to reason about typical, as opposed to certain, effects of actions.

The main contribution of this paper is a deeper investigation into *informative* actions, which correspond to the various ways in which agents can acquire information. We propose a formalization of three different informative actions, viz. *observations* (these are basically the aforementioned epistemic tests), actions modelling *communication*, and actions that model the jumping to conclusions which is typical for *default reasoning*. We are in particular interested in the various ways in which these different informative actions may interact, and how possible conflicts that may result from this interaction, can be solved. In order to resolve these possible information conflicts, we propose a classification of the information that an agent possesses according to credibility. Based on this classification, we formalize what it means for agents to have seen or heard something, or to believe something by default. Using the various informational attitudes that we thus introduced, we define the aforementioned informative actions. In the definitions of these actions an important part is played by a general belief revision action which satisfies the AGM postulates. The ability of agents to perform the three informative actions deals both with the limited capacities of agents to gather information and the preference of one (credible) source of information over another (less credible) one. The various definitions result in a framework in which agents can acquire information from various sources, solve conflicts between various information items, and attach degrees of credibility to their beliefs dependent on the way these beliefs were acquired. As such, the framework presented in this paper could be seen as an attempt to formalize the behavior of *intelligent information agents* [2, 28] which will help us sift our way through the information age.

Organization of the paper The rest of the paper is organized as follows. In Sect. 2 we (re)introduce some of our ideas on knowledge, belief, abilities,

opportunities, and results; furthermore the formal definitions of our framework are given. Those already familiar with our basic framework can safely skip this section. In Sect. 3 we present a classification by credibility of the information of an agent. In Sect. 4 we combine three informative actions into the basic framework. In 4.1 we formalize observations, in 4.2 communicative actions are formalized, and in 4.3 we formalize default jumps. In Sect. 5 we define the ability of agents with respect to the informative actions that we consider. In Sect. 6 we round off.

2. Knowledge, belief, abilities, opportunities, and results

Here we restrict ourselves to informational attitudes and action aspects. At the informational level we consider both *knowledge and belief*. Formalizing these notions has been a subject of continuing research both in analytical philosophy and in AI [14, 16]. In representing knowledge and belief we follow, both from a syntactical and a semantic point of view, the approach common in epistemic and doxastic logic: the formula $\mathbf{K}_i\varphi$ denotes the fact that agent i knows φ , and $\mathbf{B}_i\varphi$ that agent i believes φ . For the semantics we use Kripke-style possible worlds models.

At the action level we consider *results, abilities* and *opportunities*. In defining the result of an action, we follow ideas of Von Wright [46], in which the state of affairs brought about by execution of the action is defined to be its result. An important aspect of any investigation of action is the relation that exists between ability and opportunity. In order to successfully complete an action, both the opportunity and the ability to perform the action are necessary. Although these notions are interconnected, they are surely not identical [24]: the ability of agents comprises mental and physical powers, moral capacities, and human and physical possibility, whereas the opportunity to perform actions is best described by the notion of circumstantial possibility. A nice example that illustrates the difference between ability and opportunity is that of a lion in a zoo [10]: although the lion will (ideally) never have the opportunity to eat a zebra, it certainly has the ability to do so. We propose that in order to make our formalization of rational agents, like for instance robots, as accurate and realistic as possible, abilities and opportunities need also be distinguished in AI environments. The abilities of the agents are formalized via the \mathbf{A}_i operator; the formula $\mathbf{A}_i\alpha$ denotes the fact that agent i has the ability to do α . When using the definitions of opportunities and results as given above, the framework of (propositional) dynamic logic provides an excellent means to formalize these notions. Using events $\text{do}_i(\alpha)$ to refer to the performance of the action α by the agent i , we consider the formulae $\langle \text{do}_i(\alpha) \rangle \varphi$ and $[\text{do}_i(\alpha)]\varphi$. In our deterministic framework, $\langle \text{do}_i(\alpha) \rangle \varphi$ is the stronger of these formulae; it represents the fact that the agent i has the opportunity to do α and that doing α results in φ being true. The formula $[\text{do}_i(\alpha)]\varphi$ is noncommittal about the opportunity of the agent to do α but states that should the opportunity arise, only states of affairs satisfying φ would result. Besides the possibility to formalize both opportunities and results when using dynamic logic, another advantage lies in the compatibility of epistemic, doxastic and dynamic logic from a semantic point

of view: a Kripke style semantics can be used to provide meaning to epistemic, doxastic and dynamic notions.

2.1. DEFINITION. Let a finite set $\mathcal{A} = \{1, \dots, n\}$ of agents, and some denumerable sets Π of propositional symbols and At of atomic actions be given. The language \mathcal{L} and the class of actions Ac are defined by mutual induction as follows.

- \mathcal{L} is the smallest superset of Π such that
 - if $\varphi, \psi \in \mathcal{L}$ then $\neg\varphi, \varphi \vee \psi \in \mathcal{L}$
 - if $i \in \mathcal{A}$, $\alpha \in Ac$ and $\varphi \in \mathcal{L}$ then $\mathbf{K}_i\varphi, \mathbf{B}_i\varphi, \langle \text{do}_i(\alpha) \rangle\varphi, \mathbf{A}_i\alpha \in \mathcal{L}$
- Ac is the smallest superset of At such that
 - if $\varphi \in \mathcal{L}$ then $\text{confirm } \varphi \in Ac$
 - if $\alpha_1 \in Ac$ and $\alpha_2 \in Ac$ then $\alpha_1; \alpha_2 \in Ac$
 - if $\varphi \in \mathcal{L}$ and $\alpha_1, \alpha_2 \in Ac$ then $\text{if } \varphi \text{ then } \alpha_1 \text{ else } \alpha_2 \text{ fi} \in Ac$
 - if $\varphi \in \mathcal{L}$ and $\alpha_1 \in Ac$ then $\text{while } \varphi \text{ do } \alpha_1 \text{ od} \in Ac$

The purely propositional fragment of \mathcal{L} is denoted by \mathcal{L}_0 .

The constructs $\wedge, \rightarrow, \leftrightarrow, \mathbf{tt}, \mathbf{ff}, \mathbf{M}_i$ and $[\text{do}_i(\alpha)]\varphi$ are defined in the usual way. Other constructs are introduced by definitional abbreviation: \mathbf{skip} is $\text{confirm } \mathbf{tt}$, $\text{if } \varphi \text{ then } \alpha_1 \text{ fi}$ is $\text{if } \varphi \text{ then } \alpha_1 \text{ else } \mathbf{skip} \text{ fi}$, α^0 is \mathbf{skip} , and α^{n+1} is $\alpha; \alpha^n$.

2.2. REMARK. The confirm action behaves essentially like the test actions in dynamic logic [13, 15]. As such this action differs substantially from tests as they are looked upon by humans: these genuine tests are usually assumed to contribute to the information of the agent that performs the test [32], whereas by performing $\text{confirm } \varphi$ it is just *confirmed* (verified, checked) that φ holds. The meaning of the other actions in Ac is respectively: the atomic action, sequential composition, conditional composition, and repetitive composition; \mathbf{skip} denotes the void action.

In the following definitions it is assumed that some set $\mathbf{bool} = \{\mathbf{0}, \mathbf{1}\}$ of truth values is given.

2.3. DEFINITION. The class \mathbf{M} of Kripke models consists of all tuples $\mathcal{M} = \langle \mathcal{S}, \pi, \mathbf{R}, \mathbf{B}, \mathbf{r}, \mathbf{c} \rangle$ such that

- \mathcal{S} is a set of possible worlds, or states.
- $\pi : \Pi \times \mathcal{S} \rightarrow \mathbf{bool}$ is a total function that assigns a truth value to propositional symbols in possible worlds.
- $\mathbf{R} : \mathcal{A} \rightarrow \wp(\mathcal{S} \times \mathcal{S})$ is a function that yields the epistemic accessibility relations for a given agent. This function is such that $\mathbf{R}(i)$ is an equivalence relation for all i . For reasons of practical convenience we define $[s]_{\mathbf{R}(i)}$ to be $\{s' \in \mathcal{S} \mid (s, s') \in \mathbf{R}(i)\}$.
- $\mathbf{B} : \mathcal{A} \times \mathcal{S} \rightarrow \wp(\mathcal{S})$ is a function that yields the set of doxastic alternatives for a given agent in a given state. To model the kind of belief that we like to model it is demanded that for all agents i and for all possible worlds s and s' it holds that:

- $B(i, s) = B(i, s')$ if $s' \in [s]_{R(i)}$
- $B(i, s) \subseteq [s]_{R(i)}$
- $B(i, s) \neq \emptyset$
- $\mathbf{r} : \mathcal{A} \times At \rightarrow \mathcal{S} \rightarrow \wp(\mathcal{S})$ is such that $\mathbf{r}(i, a)(s)$ yields the (possibly empty) state transition in s caused by the event $\text{do}_i(a)$. This function is such that for all atomic actions a it holds that $|\mathbf{r}(i, a)(s)| \leq 1$ for all i and s , i.e., these events are *deterministic*.
- $\mathbf{c} : \mathcal{A} \times At \rightarrow \mathcal{S} \rightarrow \mathbf{bool}$ is the capability function such that $\mathbf{c}(i, a)(s)$ indicates whether the agent i is capable of performing the action a in s .

2.4. DEFINITION. Let $\mathcal{M} = \langle \mathcal{S}, \pi, R, B, \mathbf{r}, \mathbf{c} \rangle$ be some Kripke model from **M**. For propositional symbols, negated formulae, and disjunctions, $\mathcal{M}, s \models \varphi$ is inductively defined as usual. For the other clauses $\mathcal{M}, s \models \varphi$ is defined as follows:

$$\begin{array}{ll}
\mathcal{M}, s \models \mathbf{K}_i \varphi & \Leftrightarrow \forall s' \in \mathcal{S} [(s, s') \in R(i) \Rightarrow \mathcal{M}, s' \models \varphi] \\
\mathcal{M}, s \models \mathbf{B}_i \varphi & \Leftrightarrow \forall s' \in \mathcal{S} [s' \in B(i, s) \Rightarrow \mathcal{M}, s' \models \varphi] \\
\mathcal{M}, s \models \langle \text{do}_i(\alpha) \rangle \varphi & \Leftrightarrow \exists \mathcal{M}', s' [\mathcal{M}', s' \in \mathbf{r}(i, \alpha)(\mathcal{M}, s) \ \& \ \mathcal{M}', s' \models \varphi] \\
\mathcal{M}, s \models \mathbf{A}_i \alpha & \Leftrightarrow \mathbf{c}(i, \alpha)(\mathcal{M}, s) = \mathbf{1}
\end{array}$$

where \mathbf{r} and \mathbf{c} are extended as follows:

$$\begin{array}{ll}
\mathbf{r}(i, a)(\mathcal{M}, s) & = \mathcal{M}, \mathbf{r}(i, a)(s) \\
\mathbf{r}(i, \text{confirm } \varphi)(\mathcal{M}, s) & = \{(\mathcal{M}, s)\} \text{ if } \mathcal{M}, s \models \varphi \text{ and } \emptyset \text{ otherwise} \\
\mathbf{r}(i, \alpha_1; \alpha_2)(\mathcal{M}, s) & = \mathbf{r}(i, \alpha_2)(\mathbf{r}(i, \alpha_1)(\mathcal{M}, s)) \\
\mathbf{r}(i, \text{if } \varphi \text{ then } \alpha_1 & = \mathbf{r}(i, \alpha_1)(\mathcal{M}, s) \text{ if } \mathcal{M}, s \models \varphi \text{ and} \\
\quad \text{else } \alpha_2 \text{ fi})(\mathcal{M}, s) & \mathbf{r}(i, \alpha_2)(\mathcal{M}, s) \text{ otherwise} \\
\mathbf{r}(i, \text{while } \varphi \text{ do } \alpha_1 \text{ od})(\mathcal{M}, s) & = \{(\mathcal{M}', s') \mid \exists k \in \mathbb{N} \exists \mathcal{M}_0, s_0 \dots \exists \mathcal{M}_k, s_k \\
& [\mathcal{M}_0, s_0 = \mathcal{M}, s \ \& \ \mathcal{M}_k, s_k = \mathcal{M}', s' \ \& \ \forall j < k \\
& [\mathcal{M}_{j+1}, s_{j+1} = \mathbf{r}(i, \text{confirm } \varphi; \alpha_1)(\mathcal{M}_j, s_j) \\
& \ \& \ \mathcal{M}', s' \models \neg \varphi]\}
\end{array}$$

$$\text{where } \mathbf{r}(i, \alpha)(\emptyset) = \emptyset$$

and

$$\begin{array}{ll}
\mathbf{c}(i, a)(\mathcal{M}, s) & = \mathbf{c}(i, a)(s) \\
\mathbf{c}(i, \text{confirm } \varphi)(\mathcal{M}, s) & = \mathbf{1} \text{ if } \mathcal{M}, s \models \varphi \text{ and } \mathbf{0} \text{ otherwise} \\
\mathbf{c}(i, \alpha_1; \alpha_2)(\mathcal{M}, s) & = \mathbf{c}(i, \alpha_1)(\mathcal{M}, s) \ \& \ \mathbf{c}(i, \alpha_2)(\mathbf{r}(i, \alpha_1)(\mathcal{M}, s)) \\
\mathbf{c}(i, \text{if } \varphi \text{ then } \alpha_1 & = \mathbf{c}(i, \text{confirm } \varphi; \alpha_1)(\mathcal{M}, s) \text{ or} \\
\quad \text{else } \alpha_2 \text{ fi})(\mathcal{M}, s) & \mathbf{c}(i, \text{confirm } \neg \varphi; \alpha_2)(\mathcal{M}, s) \\
\mathbf{c}(i, \text{while } \varphi \text{ do } \alpha_1 \text{ od})(\mathcal{M}, s) & = \mathbf{1} \text{ if } \exists k \in \mathbb{N} [\mathbf{c}(i, (\text{confirm } \varphi; \alpha_1)^k; \\
& \quad \text{confirm } \neg \varphi)(\mathcal{M}, s) = \mathbf{1}] \\
& \text{and } \mathbf{0} \text{ otherwise}
\end{array}$$

$$\text{where } \mathbf{c}(i, \alpha)(\emptyset) = \mathbf{1}$$

Validity on and satisfiability in the class **M** of models are defined as usual.

2.5. REMARKS. Definition 2.4 gives rise to the following remarks.

- When defining the R and B functions as in Definition 2.3, we end up with a notion of knowledge that satisfies an S5 axiomatization, and a notion of belief that satisfies a KD45 axiomatization (both according to the Chellas qualification [6]). This means in particular that knowledge is veridical,

that belief is non-absurd, and that agents have positive and negative introspection on both their knowledge and their belief. Knowledge and belief are related to each other as in the system of Kraus & Lehmann [25].

- The notion of actions as considered in Definition 2.4 generalizes that of state-transformers as it is typical for dynamic logic [15], and allows also for actions that transform pairs (Model, State). The reason for this generalization lies in the fact that we also account for non-standard actions like ‘to observe’ [32], ‘to commit’ [33] and ‘to inform’ [30] in our framework, and these non-standard actions transform models rather than states.
- With regard to the abilities of agents, the motivation for the choices made in Definition 2.4 is the following. The definition of $c(i, \text{confirm } \varphi)$ expresses that an agent is able to get confirmation for a formula φ if and only if φ holds. Note that the definitions of $r(i, \text{confirm } \varphi)$ and $c(i, \text{confirm } \varphi)$ imply that in circumstances such that φ holds, agents have both the opportunity and the ability to confirm φ . An agent is capable of performing a sequential composition $\alpha_1; \alpha_2$ if and only if it is capable of performing α_1 (now), and it is capable of executing α_2 after it has performed α_1 . An agent is capable of performing a conditional composition, if either it is able to get confirmation for the condition and thereafter perform the then-part, or it is able to confirm the negation of the condition and perform the else-part afterwards. An agent is capable of performing a repetitive composition $\text{while } \varphi \text{ do } \alpha_1 \text{ od}$ if and only if it is able to perform the action $(\text{confirm } \varphi; \alpha_1)^k; \text{confirm } \neg\varphi$ for some natural number k .

3. A classification of information

When modelling real-life situations it will hardly ever be the case that agents have complete information about the world and their place in it. In particular when planning it may be necessary for agents to try and acquire additional information from whatever source possible. Here we consider three possible sources of information for an agent, an *endogenous* one and two *exogenous* sources. The first exogenous source consists of the *observations* that an agent makes about the current world [32]; the second exogenous source of information for a rational agent is *communication* with other agents [30]. The endogenous source of information that we consider is the possibility to *adopt assumptions by default* [31]. Other authors came up with more or less the same sources of information for rational agents. Castelfranchi [4, 5] also considers three sources, viz. one endogenous source comprising reasoning and introspection (due to their perfect reasoning and introspection capacities this source is already implicitly present in the agents that we formalize), and perception and communication as exogenous sources. Dunin-Keplicz & Treur [9] consider exactly the same sources of information as we do.

In any situation in which information is acquired from different sources, the possibility of conflicts exists. That is, it is possible that information acquired from one source contradicts information acquired from another source. To solve these information conflicts, it is important to define some strategies that pre-

scribe what to do in the case of conflicts between new, incoming information and old, previously acquired information. Castelfranchi [4] proposes that strategies which resolve information conflicts should take two main aspects into consideration. The first of these could be captured by the notion of *economical* — or *conservative* — aspects. That is, as long as there is no very good reason to favor new beliefs over old ones, an agent should stick to its old beliefs thereby avoiding the need for (costly) belief revisions. Having said so, the second point that should be taken into consideration is that more credible information should always be favored over less credible information. So, in cases where the incoming information is strictly more credible than the already present information, a revision in favor of the new information should be performed, whereas in the case of peer conflicts economical aspects prevail, and therefore no revision takes place.

In order to compare various items of information, some degree of credibility should be attached to them. In particular, it is necessary to have the possibility to determine the credibility of each of the beliefs that an agent may have in a certain state. In order to be able to do so we propose to structure the information that an agent possesses into four sets, one inside the other. The innermost of these sets contains the *knowledge* of the agent. Knowledge can be seen as the information that an agent is born or built with. The set directly encompassing the knowledge set contains the *observational* beliefs of the agent. These are the beliefs that an agent has on the ground of its knowledge and the observations it has made. The third set contains the *communicational* beliefs of an agent. These are the combined beliefs that it either knows or acquired through observations and/or communication. The outermost set contains the *default beliefs*. These are the beliefs for which application of a default may have been a necessary condition. The credibility of a belief formula is determined by the smallest belief set that it is a member of, i.e., formulae that are known have the highest credibility whereas formulae that are believed by default have the lowest one. In order to syntactically account for this structuring of beliefs, we tag the belief modality as follows.

3.1. DEFINITION. The language \mathcal{L} as given in Definition 2.1 is extended as follows:

- if $\varphi \in \mathcal{L}$ then $\mathbf{B}_i^o\varphi, \mathbf{B}_i^c\varphi, \mathbf{B}_i^d\varphi \in \mathcal{L}$

The intuitive interpretation of $\mathbf{B}_i^o\varphi$ is that φ belongs to the observational beliefs of agent i , $\mathbf{B}_i^c\varphi$ states that φ belongs to the communicational beliefs of i , and \mathbf{B}_i^d refers to the default beliefs of agent i .

3.2. CONVENTION. Both for reasons of practical convenience and to keep our notation uniform, we sometimes use \mathbf{B}_i^k , for *known* beliefs, as a general way of writing \mathbf{K}_i .

For the semantics of the various tagged belief operators we propose the use of *belief clusters* [34], which are basically sets of worlds, that are situated within each other. That is, the knowledge cluster is the largest one, corresponding to the idea that knowledge is the most credible kind of belief. The observational

belief cluster is situated within the knowledge cluster; this cluster contains the communicational belief cluster, which on its turn contains the default belief cluster. Hence the credibility of a belief is inversely proportional with the size of the cluster on which the belief is interpreted. Note furthermore that from the point of view of belief revision [11, 12], *expansions* of the various belief sets correspond to *shrinkings* of the associated belief clusters and *contractions* of the belief sets correspond to *extensions* of the clusters (cf. [29]).

3.3. DEFINITION. The Kripke models given in Definition 2.3 are modified as follows: the function B is replaced by three functions $B^o : \mathcal{A} \rightarrow \wp(\mathcal{S} \times \mathcal{S})$, $B^c : \mathcal{A} \times \mathcal{S} \rightarrow \wp(\mathcal{S})$ and $B^d : \mathcal{A} \times \mathcal{S} \rightarrow \wp(\mathcal{S})$. Defining the set $[s]_{B^o(i)}$ analogously to $[s]_{R(i)}$, these functions are such that for all $i \in \mathcal{A}$, $s, s' \in \mathcal{S}$:

- $B^o(i)$ is an equivalence relation
- $B^d(i, s) \neq \emptyset$
- $B^d(i, s) \subseteq B^c(i, s) \subseteq [s]_{B^o(i)} \subseteq [s]_{R(i)}$
- if $s' \in [s]_{R(i)}$ then $[s']_{B^o(i)} = [s]_{B^o(i)}$, $B^c(i, s') = B^c(i, s)$ and $B^d(i, s') = B^d(i, s)$

The tagged belief formulae are interpreted as expected:

3.4. DEFINITION. For a Kripke model \mathcal{M} as given in Definition 3.3 and a state $s \in \mathcal{M}$, \models is extended as follows:

$$\begin{aligned} \mathcal{M}, s \models \mathbf{B}_i^o \varphi &\Leftrightarrow \forall s' \in \mathcal{S} [(s, s') \in B^o(i) \Rightarrow \mathcal{M}, s' \models \varphi] \\ \mathcal{M}, s \models \mathbf{B}_i^c \varphi &\Leftrightarrow \forall s' \in B^c(i, s) [\mathcal{M}, s' \models \varphi] \\ \mathcal{M}, s \models \mathbf{B}_i^d \varphi &\Leftrightarrow \forall s' \in B^d(i, s) [\mathcal{M}, s' \models \varphi] \end{aligned}$$

When defining knowledge and the various notions of belief as in Definition 3.4 we indeed obtain the credibility ordering we are aiming at as an ordinary material implication. According to this ordering, knowing some formula is the strongest informational attitude, and believing it by default is the weakest one.

3.5. PROPOSITION. *Define the ordering $>$ on informational operators by $\mathbf{B}_i^k > \mathbf{B}_i^o > \mathbf{B}_i^c > \mathbf{B}_i^d$, and let \succeq be the reflexive, transitive closure of $>$. Then for all $\mathbf{X}, \mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$, and for all formulae φ :*

- $\mathbf{X} \succeq \mathbf{Y} \Rightarrow \models \mathbf{X}\varphi \rightarrow \mathbf{Y}\varphi$

The various notions of belief that we introduced validate different axiomatizations: knowledge and observational belief satisfy an S5-axiomatization, whereas communicational and default belief validate an KD45-axiomatization.

3.6. PROPOSITION. *Let $\mathbf{X} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o\}$ and $\mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$. For all formulae φ we have:*

- | | |
|--|---|
| 1. $\models \mathbf{Y}(\varphi \rightarrow \psi) \rightarrow (\mathbf{Y}\varphi \rightarrow \mathbf{Y}\psi)$ | K |
| 2. $\models \neg(\mathbf{Y}\varphi \wedge \mathbf{Y}\neg\varphi)$ | D |
| 3. $\models \mathbf{X}\varphi \rightarrow \varphi$ | T |
| 4. $\models \mathbf{Y}\varphi \rightarrow \mathbf{Y}\mathbf{Y}\varphi$ | 4 |

5. $\models \neg \mathbf{Y}\varphi \rightarrow \mathbf{Y}\neg\mathbf{Y}\varphi$ 5
6. $\models \varphi \Rightarrow \models \mathbf{Y}\varphi$ N

Although our system is a generalization of the more common systems for knowledge and belief [17, 25], it is also the case in our system that in certain situations multiple informational operators prefixing a formula collapse to one operator.

3.7. DEFINITION. A formula φ is *i-doxastic sequenced* if there is a formula ψ , and operators $\mathbf{X}_1, \dots, \mathbf{X}_m \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d, \neg\mathbf{B}_i^k, \neg\mathbf{B}_i^o, \neg\mathbf{B}_i^c, \neg\mathbf{B}_i^d\}$ and $m > 0$ such that $\varphi = \mathbf{X}_1 \dots \mathbf{X}_m \psi$.

3.8. PROPOSITION. Let φ be an *i-doxastic sequenced* formula, and let $\mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$. Then:

- $\models \mathbf{Y}\varphi \leftrightarrow \varphi$

3.9. OBSERVATION. Given the four modal operators $\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d$ it is possible to model nine different informational attitudes of a given agent with regard to a given formula. If we define for $x \in \{k, o, c, d\}$, $\mathbf{Agn}_i^x \varphi =^{\text{def}} \neg\mathbf{B}_i^x \varphi \wedge \neg\mathbf{B}_i^x \neg\varphi$, representing the fact that agent i is *agnostic* with regard to φ on the level x , the attitudes with respect to a formula φ are the following:

- | | |
|---|---|
| 1. $\mathbf{B}_i^k \varphi$ | ‘ <i>i</i> knows φ ’ |
| 2. $\mathbf{B}_i^k \neg\varphi$ | ‘ <i>i</i> knows $\neg\varphi$ ’ |
| 3. $\mathbf{Agn}_i^k \varphi \wedge \mathbf{B}_i^o \varphi$ | ‘ <i>i</i> saw φ ’ |
| 4. $\mathbf{Agn}_i^k \varphi \wedge \mathbf{B}_i^o \neg\varphi$ | ‘ <i>i</i> saw $\neg\varphi$ ’ |
| 5. $\mathbf{Agn}_i^o \varphi \wedge \mathbf{B}_i^c \varphi$ | ‘ <i>i</i> heard φ ’ |
| 6. $\mathbf{Agn}_i^o \varphi \wedge \mathbf{B}_i^c \neg\varphi$ | ‘ <i>i</i> heard $\neg\varphi$ ’ |
| 7. $\mathbf{Agn}_i^c \varphi \wedge \mathbf{B}_i^d \varphi$ | ‘ <i>i</i> believes φ by default’ |
| 8. $\mathbf{Agn}_i^c \varphi \wedge \mathbf{B}_i^d \neg\varphi$ | ‘ <i>i</i> believes $\neg\varphi$ by default’ |
| 9. $\mathbf{Agn}_i^d \varphi$ | ‘ <i>i</i> is completely agnostic wrt φ ’ |

Given these attitudes and their intuitive interpretation, we introduce the following predicates by definitional abbreviation:

- $\mathbf{Saw}_i \varphi =^{\text{def}} \mathbf{Agn}_i^k \varphi \wedge \mathbf{B}_i^o \varphi$
- $\mathbf{Heard}_i \varphi =^{\text{def}} \mathbf{Agn}_i^o \varphi \wedge \mathbf{B}_i^c \varphi$
- $\mathbf{Jumped}_i \varphi =^{\text{def}} \mathbf{Agn}_i^c \varphi \wedge \mathbf{B}_i^d \varphi$

In fact, the \mathbf{Heard}_i operator does not formalize hearing *per se*, but formalizes *believing on the basis of hearing*. One cannot prevent an agent from being told inconsistencies. However, the agent will deal with these inconsistencies in such a way that at any given time its beliefs grounded in the things that it heard are consistent.

3.10. REMARK. Note that the definition of \mathbf{Agn}_i^x is such that $\mathbf{Agn}_i^x \varphi$ and $\mathbf{Agn}_i^x \neg\varphi$ are equivalent notions. Note furthermore that $\mathbf{Saw}_i \varphi$, $\mathbf{Heard}_i \varphi$ and $\mathbf{Jumped}_i \varphi$ could equivalently — and less complex — be defined as $\mathbf{B}_i^o \varphi \wedge \neg\mathbf{B}_i^k \varphi$, $\mathbf{B}_i^c \varphi \wedge \neg\mathbf{B}_i^o \varphi$ and $\mathbf{B}_i^d \varphi \wedge \neg\mathbf{B}_i^c \varphi$ respectively. However, to emphasize that the agent is genuinely *ignorant* at higher levels, we have chosen to include the \mathbf{Agn}_i^x operator in the definition of the derived belief operators.

The following proposition formalizes the behavior of the derived belief operators with respect to the axioms of modal logic as given by Chellas [6].

3.11. PROPOSITION. *Let φ and ψ be formulae. Let \mathbf{X} be in the set $Bel = \{\mathbf{Saw}_i, \mathbf{Heard}_i, \mathbf{Jumped}_i \mid i \in \mathcal{A}\}$, and let $\mathbf{Y} \in Bel \cup \{\mathbf{K}_i\}$. Define the ordering \succeq' to be the reflexive and transitive closure of $>'$ with $\mathbf{K}_i >' \mathbf{Saw}_i >' \mathbf{Heard}_i >' \mathbf{Jumped}_i$ ². Then we have:*

1. $\models \mathbf{X}\varphi \wedge \mathbf{X}(\varphi \rightarrow \psi) \rightarrow \mathbf{X}\psi$	K
2. $\models \neg(\mathbf{X}\varphi \wedge \mathbf{X}\neg\varphi)$	D
3. $\models \mathbf{Saw}_i\varphi \rightarrow \varphi$	T
4. $\not\models \mathbf{X}\varphi \rightarrow \mathbf{X}\mathbf{X}\varphi$	4
5. $\models \mathbf{X}\varphi \rightarrow \mathbf{K}_i\mathbf{X}\varphi$	4'
6. $\not\models \neg\mathbf{X}\varphi \rightarrow \mathbf{X}\neg\mathbf{X}\varphi$	5
7. $\models \neg\mathbf{X}\varphi \rightarrow \mathbf{K}_i\neg\mathbf{X}\varphi$	5'
8. $\models (\mathbf{X}\varphi \wedge \mathbf{X}\psi) \rightarrow \mathbf{X}(\varphi \wedge \psi)$	C
9. $\not\models \mathbf{X}(\varphi \wedge \psi) \rightarrow (\mathbf{X}\varphi \wedge \mathbf{X}\psi)$	M
10. $\models \mathbf{X}(\varphi \wedge \psi) \rightarrow (\bigvee_{\mathbf{Y} \succeq' \mathbf{X}} \mathbf{Y}\varphi \wedge \bigvee_{\mathbf{Y} \succeq' \mathbf{X}} \mathbf{Y}\psi) \wedge (\mathbf{X}\varphi \vee \mathbf{X}\psi)$	M'
11. $\models \varphi \not\models \mathbf{X}\varphi$ and even $\models \varphi \Rightarrow \models \neg\mathbf{X}\varphi$	N
12. $\models \varphi \rightarrow \psi \not\models \mathbf{X}\varphi \rightarrow \mathbf{X}\psi$	RM
13. $\models \varphi \rightarrow \psi \Rightarrow \models \mathbf{X}\varphi \rightarrow \bigvee_{\mathbf{Y} \succeq' \mathbf{X}} \mathbf{Y}\psi$	RM'
14. $\models \varphi \leftrightarrow \psi \Rightarrow \models \mathbf{X}\varphi \leftrightarrow \mathbf{X}\psi$	RE

3.12. REMARK. That the K-axiom holds for all three derived belief operators might look somewhat surprising at first sight. However, validity of this axiom has everything to do with the *rationality* of agents. Take the example of a rational agent grounding its beliefs of both φ and $\varphi \rightarrow \psi$ in its observations. Being the rational creature that it is, it is obvious that it in any case observationally believes ψ . Again given the rationality of the agent it cannot attach a higher credibility to ψ than to $\varphi \rightarrow \psi$: for the latter is implied by the former, and therefore the credibility attached to it is at least the credibility attached to ψ . Hence both ψ and $\varphi \rightarrow \psi$ are believed with the same strength. Beliefs are consistent (the D-axiom), and beliefs that are grounded in observations are furthermore veridical (the T-axiom). The first property is highly desirable for rational agents, the latter property is the essential characteristic of beliefs acquired through observations (cf. [3], p. 12). None of the derived belief operators validates positive or negative introspection. However, all operators satisfy introspective properties with respect to knowledge: if an agent believes something on the ground of its observations (communicating, default jumps) then it *knows* that it does so. This is as one would intuitively expect: if an agent sees something, then it does not *see* that it sees it, but a rational agent *knows* that it sees it. All derived belief operators distribute over conjunctions (the R-axiom), and none of these satisfy necessitation (the N-rule). The reason for this latter fact is that valid propositions are already known, and are therefore never grounded in observations, communication or default jumps. The derived belief operators are in general not monotonic (the RM-rule), but satisfy some kind of ‘upward

²Note that \succeq' and $>'$ are the obvious modifications of \succeq and $>$ to deal with the derived informational operators.

monotonicity’ (the RM’-rule), corresponding to the idea that whenever some proposition is believed to a certain degree, weaker propositions might already be believed to a higher degree (the degenerate case being one where the weaker propositions are validities that are known to the agent). Finally, all derived belief operators are closed under equivalence of (believed) propositions. Although Barwise argues against this property (cf. [3], p. 18), it seems harmless for artificial agents with perfect reasoning capacities.

4. Combining informative actions

In this section we successively formalize three kinds of informative actions, which correspond to the various ways in which agents can acquire information. The first of these actions is the one that models the observations that agents may make. Through performing an **observe** φ action an agent observes whether φ holds in the state in which it is currently residing. The second action corresponds to information acquisition through communication. By performing an **inform** (φ, i) action, some agent j may inform agent i of the fact φ . The third and last action corresponds to the endogenous source of information acquisition, viz. the reasoning by default. By performing **try_jump** φ actions agents may try and jump to the default conclusion φ . In defining the semantics of these ‘high-level’ informative actions, we use a special ‘low-level’ action that causes the beliefs of an agent to be *revised* in the appropriate manner. Intuitively, a revision is the change of belief through which some formula is — whenever possible — consistently included in the belief set of an agent. If the formula that is to be included is consistent with the already present beliefs of the agent it is simply added, otherwise — in order to maintain consistency of the agent’s belief set — some of the old beliefs of the agent are deleted before including the new formula. The revision implemented by the actions that we present validates the (well known and by now standard) AGM postulates for belief revision [1, 11, 12, 29]. To deal adequately with the classification of the agents information, the low-level revision action is defined to stretch over the various information classes. To emphasize the distinction between the high and the low-level informative actions, we introduce in addition to the set of rational agents a constant e in our language. The intuitive reading of this constant could either be *external environment/observer*, or *supervisor*. The idea is that this external environment performs the low-level actions that are not available to the agents, but that appear in the implementation of actions that are available. As such, the low-level revision actions are not visible for the agents: when they reason about a high-level informative action α in the implementation of which a revision action occurs, then α appears as an indivisible unit to the agent. Furthermore, the ability of agents to perform high-level informative actions is defined without regard to the implementation provided by the low-level revision actions; since these latter actions are not available to the agents it is of no use to discuss the notion of ability for these actions.

4.1. DEFINITION. The class of high-level informative actions Ac_i^h and the class of low-level informative actions Ac_i^l are defined by:

- $Ac_I^h = \{\text{observe } \varphi, \text{inform } (\varphi, i), \text{try_jump } \varphi \mid \varphi \in \mathcal{L}_0, i \in \mathcal{A}\}$
- Ac_I^h is the smallest superset of $\{\text{revise}^x(\varphi, i) \mid x \in \{1, 2, 3, d, c, o\}, i \in \mathcal{A}, \varphi \in \mathcal{L}_0\}$ that is closed under sequential composition.

The class of actions Ac is extended with the class Ac_I^h .

Informally speaking, the $\text{revise}^x, x = d, c, o$ actions that we define below take care of a belief revision in the sense of the AGM postulates but *stretched over different levels*. For example, consider the case where a revision with some formula φ of the observational beliefs of an agent should take place. Given the classification of belief as presented in Sect. 3 this implies that not only the agent's observational beliefs, but also its communicational and default beliefs need to be updated, i.e., the revision should affect other kinds of the agent's beliefs. In the following definition, the $\text{revise}^k, k = 1, 2, 3$ actions take care of a revision per level, whereas the $\text{revise}^x, x = d, c, o$ actions force a belief revision on all appropriate levels.

4.2. DEFINITION. Let $\mathcal{M} = \langle \mathcal{S}, \pi, R, B^o, B^c, B^d, \mathbf{r}, \mathbf{c} \rangle$ be some Kripke model, let $s \in \mathcal{S}$, let $\mathcal{S}' \subseteq \mathcal{S}$, and let $\varphi \in \mathcal{L}_0$.

- $[s]_{B^o(i)}^{\varphi+} =^{\text{def}} [s]_{B^o(i)} \cap \{s' \in \mathcal{S} \mid \mathcal{M}, s \models \varphi \Leftrightarrow \mathcal{M}, s' \models \varphi\}$
- $[s]_{B^o(i)}^{\varphi-} =^{\text{def}} [s]_{B^o(i)} \setminus [s]_{B^o(i)}^{\varphi+}$
- $\text{Cl}_{\text{eq}}(\mathcal{S}') =^{\text{def}} \mathcal{S}' \times \mathcal{S}'$

4.3. DEFINITION. For $\mathcal{M} = \langle \mathcal{S}, \pi, R, B^o, B^c, B^d, \mathbf{r}, \mathbf{c} \rangle$ some Kripke model with $s \in \mathcal{S}$, i some agent, and φ some propositional formula we define:

- The function \mathbf{r} is retyped to $(\mathcal{A} \times Ac) \cup (\{e\} \times Ac_I^h) \rightarrow (\mathcal{M} \times \mathcal{S}) \rightarrow \wp(\mathcal{M} \times \mathcal{S})$.
- $\mathbf{r}(e, \text{revise}^1(i, \varphi))(\mathcal{M}, s) = \mathcal{M}', s$ where
 $\mathcal{M}' = \langle \mathcal{S}, \pi, R, B^o, B^c, B^d, \mathbf{r}, \mathbf{c} \rangle$ with
 $B^d(i', s') = B^d(i', s')$ if $i \neq i'$ or $s' \notin [s]_{R(i)}$
 $B^d(i, s') = \begin{cases} B^d(i, s) \cap \llbracket \varphi \rrbracket & \text{if } B^d(i, s) \cap \llbracket \varphi \rrbracket \neq \emptyset, s' \in [s]_{R(i)} \\ B^c(i, s) \cap \llbracket \varphi \rrbracket & \text{if } B^d(i, s) \cap \llbracket \varphi \rrbracket = \emptyset, s' \in [s]_{R(i)} \end{cases}$
- $\mathbf{r}(e, \text{revise}^2(i, \varphi))(\mathcal{M}, s) = \mathcal{M}', s$ where
 $\mathcal{M}' = \langle \mathcal{S}, \pi, R, B^o, B^c, B^d, \mathbf{r}, \mathbf{c} \rangle$ with
 $B^c(i', s') = B^c(i', s')$ if $i \neq i'$ or $s' \notin [s]_{R(i)}$
 $B^c(i, s') = \begin{cases} B^c(i, s) \cap \llbracket \varphi \rrbracket & \text{if } B^c(i, s) \cap \llbracket \varphi \rrbracket \neq \emptyset, s' \in [s]_{R(i)} \\ [s]_{B^o(i)} \cap \llbracket \varphi \rrbracket & \text{if } B^c(i, s) \cap \llbracket \varphi \rrbracket = \emptyset, s' \in [s]_{R(i)} \end{cases}$
- $\mathbf{r}(e, \text{revise}^3(i, \varphi))(\mathcal{M}, s) = \mathcal{M}', s$ where
 $\mathcal{M}' = \langle \mathcal{S}, \pi, R, B^o, B^c, B^d, \mathbf{r}, \mathbf{c} \rangle$ with
 $B^o(i') = B^o(i')$ if $i' \neq i$
 $B^o(i) = (B^o(i) \setminus \text{Cl}_{\text{eq}}([s]_{B^o(i)})) \cup \text{Cl}_{\text{eq}}([s]_{B^o(i)}^{\varphi+}) \cup \text{Cl}_{\text{eq}}([s]_{B^o(i)}^{\varphi-})$
- $\mathbf{r}(e, \text{revise}^d(i, \varphi))(\mathcal{M}, s) = \mathbf{r}(e, \text{revise}^1(i, \varphi))(\mathcal{M}, s)$
- $\mathbf{r}(e, \text{revise}^c(i, \varphi))(\mathcal{M}, s) = \mathbf{r}(e, \text{revise}^2(i, \varphi); \text{revise}^1(i, \varphi))(\mathcal{M}, s)$
- $\mathbf{r}(e, \text{revise}^o(i, \varphi))(\mathcal{M}, s) = \mathbf{r}(e, \text{revise}^3(i, \varphi); \text{revise}^2(i, \varphi); \text{revise}^1(i, \varphi))(\mathcal{M}, s)$

4.4. REMARKS. Definition 4.3 gives rise to the following remarks.

- The belief revision implemented in Definition 4.3 is the so called *All-is-Good* revision that we previously defined [29]. This kind of revision corresponds to a revision based on *full meet* contraction as it is defined by Gärdenfors [11]. Full meet revision is the most *rigorous* way of revising beliefs under the AGM postulates. Although this rigour of revision is in general undesirable, for our goals it does not constitute much of a problem (see item 10 of Proposition 4.8).
- An important observation concerning Definition 4.3 is that the function \mathbf{r} is in fact not well-defined, neither for the $\mathbf{revise}^m, m = 1, 2, 3$ actions nor for the $\mathbf{revise}^x, x = d, c, o$ actions, i.e., it is possible that the result from $\mathbf{r}(e, \mathbf{revise}^r(i, \varphi))(\mathcal{M}, s), r \in \{1, 2, 3, d, c, o\}$ does not satisfy the demands given in Definition 3.3. For the $\mathbf{revise}^m, m = 1, 2, 3$ actions this is a consequence of the fact that each of these actions works on its own particular level, and does not take ‘lower’ levels into account. Therefore it is for instance possible that performing the \mathbf{revise}^2 action restricts the set of B^c worlds to those satisfying some formula φ whilst the set B^d still contains $\neg\varphi$ worlds, which conflicts the subset ordering imposed in Definition 3.3. In the case of the $\mathbf{revise}^x, x = d, c, o$ actions the non well-definedness follows from the fact that this action could be executed in inappropriate circumstances. For instance, performing a $\mathbf{revise}^c(i, \varphi)$ action in a situation in which $\mathbf{B}_i^c \neg\varphi$ holds results in the set $B^c(i, s)$ being empty, thereby contradicting the demand for non-emptiness given in Definition 3.3. Both problems observed above are however not that serious: we will ensure that the revision actions behave correctly when applied in the implementation of a high-level informative action. Nevertheless, should one consider this deficiency to be insurmountable, one could always expand Definition 4.3 in such a way that the problem disappears (thereby ending up with a definition that is four or five times the size of Definition 4.3).

The following definition considers some properties that are more or less typical for informative actions.

4.5. DEFINITION. Let χ be an arbitrary formula. We distinguish the following properties of actions α :

- $\models \langle \text{do}_i(\alpha) \rangle \mathbf{tt}$ *realizability*
- $\models \langle \text{do}_i(\alpha) \rangle \chi \leftrightarrow [\text{do}_i(\alpha)]\chi$ *determinism*
- $\models \langle \text{do}_i(\alpha; \alpha) \rangle \chi \leftrightarrow \langle \text{do}_i(\alpha) \rangle \chi$ *idempotence*

Realizability of an action implies that agents have the opportunity to perform the action regardless of the situation, determinism of an action means that performing the action results in a unique state of affairs, and idempotence of an action implies that sequentially performing the action an arbitrary number of times has the same effect as performing the action just once.

4.1. Formalizing observations: seeing is believing

Through observations an agent *learns whether* some proposition is true of the state in which it is residing. For artificial agents it seems to be a reasonable assumption to demand that observations are *truthful* [32]. That is, if some observation yields information that φ , then it should indeed be the case that φ ³. Observations form the most trustworthy way of acquiring information: utterances like ‘I’ve seen it with my own eyes’ or ‘Seeing is believing’ support this claim. The formalization that we propose is therefore such that observations overrule any beliefs acquired by other means, i.e., observations will in general lead to belief revisions. The truth-value assigned to φ in the current world determines how the (observational) beliefs of the agent are revised: if φ holds, the agent’s beliefs are revised with φ , otherwise a revision with $\neg\varphi$ takes place. Note that observations will never conflict with an agent’s knowledge: since knowledge is veridical, and observations are truthful, it is not possible that an observation that some formula φ holds contradicts knowledge that $\neg\varphi$ holds.

4.6. DEFINITION. For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ we define:

$$\mathbf{r}(i, \text{observe } \varphi)(\mathcal{M}, s) = \begin{cases} \mathcal{M}, s & \text{if } \mathcal{M}, s \models \neg \mathbf{Agn}_i^k \varphi \\ \mathbf{r}(e, \text{revise}^o(i, \varphi))(\mathcal{M}, s) & \text{if } \mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \varphi \\ \mathbf{r}(e, \text{revise}^o(i, \neg\varphi))(\mathcal{M}, s) & \text{if } \mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \neg\varphi \end{cases}$$

When defining the function \mathbf{r} for the **observe** action as done in Definition 4.6 it is indeed the case that the revision actions are applied correctly.

4.7. PROPOSITION. For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ , $\mathbf{r}(i, \text{observe } \varphi)(\mathcal{M}, s)$ is a well-defined Kripke model.

4.8. PROPOSITION. For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' we have:

1. **observe** φ is deterministic, realizable and idempotent for all $\varphi \in \mathcal{L}_0$
2. $\models \mathbf{K}_{i'} \psi \leftrightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{K}_{i'} \psi$
3. $\models \langle \text{do}_i(\text{observe } \varphi) \rangle \chi \leftrightarrow \langle \text{do}_i(\text{observe } \neg\varphi) \rangle \chi$
4. $\models \langle \text{do}_i(\text{observe } \varphi) \rangle \neg \mathbf{Agn}_i^o \varphi$
5. $\models \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \varphi$
6. $\models \neg\varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \neg\varphi$
7. $\models \varphi \wedge \mathbf{Agn}_i^k \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \varphi$
8. $\models \neg\varphi \wedge \mathbf{Agn}_i^k \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \neg\varphi$
9. $\models \varphi \wedge (\mathbf{Heard}_i \neg\varphi \vee \mathbf{Jumped}_i \neg\varphi) \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \varphi$
10. $\models \varphi \wedge \mathbf{B}_i^c \neg\varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle ((\mathbf{B}_i^c \chi \leftrightarrow \mathbf{B}_i^o \chi) \wedge (\mathbf{B}_i^d \chi \leftrightarrow \mathbf{B}_i^c \chi))$

³Note that this property does not hold for human agents: magicians make a living out of this.

4.9. REMARK. Item 1 of Proposition 4.8 states that observations validate all of the properties introduced in Def. 4.5. Item 2 states that the knowledge fluents — the propositional formulae known to be true — of all agents persist under execution of an **observe** action by one of them. Item 3 states that the **observe** φ action formalizes ‘observing whether φ ’: observing whether φ is equivalent to observing whether $\neg\varphi$. Item 4 states that agents always have the opportunity to make observations, and that after observing whether φ , agents are no longer observationally agnostic with respect to φ . Items 5 and 6 follow by a combination of the property of truthfulness with the idea of agents learning whether some proposition is true: if φ is true, agents learn *that* φ , and analogously for $\neg\varphi$. Item 7 and 8 state that for knowledge-agnostic agents observations actually lead to *learning by seeing*. Item 9 — a special case of item 7 — is a very important one: it states that observations are the most credible source of information. Observations overrule other beliefs acquired through communication or adopted by default, i.e, incorrect communicational or default beliefs are *revised* in favor of observational beliefs. Item 10 sheds some more light on the way in which beliefs are revised: observing something that contradicts communicational beliefs leads to a *reset* of both the latter and the default beliefs of the agent, i.e., after such a revision all beliefs of the agent are at least grounded in its observations. Item 10 is a consequence of the use of All-is-Good revision, and indicates that our agents are rather rigorous in the maintenance of their beliefs.

4.2. Formalizing communication: hearing is believing

Through communication agents can pass on some of the information that they possess to other agents which can use this incoming information to update their beliefs. The formalization of communication that we present here is that of ‘super-cooperative’, didactic agents⁴. These agents always have the opportunity to pass on all of their belief to every other agent. Note that this action is in fact not a genuinely informative one: only the agent that receives the sent formula, and not the agent that performs the action, has the possibility to extend its information (with respect to the sent formula). It is however possible that the sending agent — the one who executes the **inform** action — does acquire information, for instance on the beliefs of the receiving agent.

In our opinion, when modelling communication in multi-agent frameworks, the concept of *trust and dependence relations* between agents deserves considerable attention. These relations are used to model whether some agent is considered to be a credible source on a certain subject by another agent [20, 30]. In communication this notion of credibility is very important: an agent i may or may not accept the information on φ that an agent j sends to it, depending on i ’s trust in j with respect to φ . We formalize the dependence relations between agents using the *dependent* operator $\mathbf{D}_{i,j}$, introduced by Huang [20]. We like to think of $\mathbf{D}_{i,j}\varphi$ as expressing the fact that agent j is a *teacher* of agent i on the subject φ .

⁴This kind of communication is among the ones we previously formalized [30]. Shoham’s agents [41] communicate in this way.

4.10. DEFINITION. We extend Definition 2.1 as follows:

- if $i, j \in \mathcal{A}$ and $\varphi \in \mathcal{L}$ then $\mathbf{D}_{i,j}\varphi \in \mathcal{L}$

The Kripke models from Definition 2.3 and Definition 3.3 are extended with a function $\mathbf{D} : \mathcal{A} \times \mathcal{A} \times \mathcal{S} \rightarrow \wp(\mathcal{L})$. We define:

- $\mathcal{M}, s \models \mathbf{D}_{i,j}\varphi$ iff $\varphi \in \mathbf{D}(i, j, s)$

Besides credibility of the sending agent, formalized through the $\mathbf{D}_{i,j}$ operator, it is furthermore very important *how* the sending agent acquired its information [4]. If the agent itself *observed* the truth of the formula that it is sending, then this formula should be considered more credible than when the agent *heard* the formula, or *adopted it by default*. The formalization that we propose is such that whenever an agent sends its observational beliefs then this will in general lead to an incorporation in the beliefs of the receiving agent. To ensure that the credibility ordering within the agent's beliefs remains meaningful, i.e., to ensure that communicational beliefs are indeed more credible than default beliefs, the source of these communicational beliefs needs to be credible. Now whenever some agent is sending out formulae that it heard, it is in general not possible to determine the credibility of these formulae, since one lost track of their source. Therefore, beliefs that the sending agent itself acquired through communication lead to a revision of the beliefs of the receiving agent only if this agent is completely agnostic with respect to these formulae. Formulae that the sending agent believes by default are considered to be too weak to be incorporated in the receiving agent's beliefs. Agents will therefore not react upon being told a default belief.

4.11. DEFINITION. For all Kripke models \mathcal{M} with state s , and for all agents i, j and propositional formulae φ we define:

$$\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s) = \begin{cases} \emptyset & \text{if } \mathcal{M}, s \models \neg \mathbf{B}_j^d \varphi \\ \mathbf{r}(e, \mathbf{revise}^c(i, \varphi))(\mathcal{M}, s) & \text{if } \mathcal{M}, s \models \mathbf{D}_{i,j}\varphi \wedge ((\mathbf{B}_j^o \varphi \wedge \mathbf{Agn}_i^o \varphi) \vee (\mathbf{Heard}_j \varphi \wedge \mathbf{Agn}_i^d \varphi)) \\ \mathcal{M}, s & \text{otherwise} \end{cases}$$

The definition of \mathbf{r} for the \mathbf{inform} action also applies the revision actions correctly:

4.12. PROPOSITION. *For all Kripke models \mathcal{M} with state s , and for all agents i, j and propositional formulae φ , if $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s) \neq \emptyset$ then it holds that $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s)$ is a well-defined Kripke model.*

The following proposition shows that Definition 4.11 indeed takes care of a formalization of the intuitive ideas that we previously exposed.

4.13. PROPOSITION. *For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' and j we have:*

1. $\mathbf{inform}(\varphi, i)$ is deterministic and idempotent for all $\varphi \in \mathcal{L}_0, i \in \mathcal{A}$
2. $\models \mathbf{B}_i^o \psi \rightarrow [\mathbf{do}_j(\mathbf{inform}(\varphi, i))\mathbf{B}_i^o \psi$

3. $\models \mathbf{B}_j^d \varphi \leftrightarrow \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \text{tt}$
4. $\models \mathbf{B}_j^d \varphi \wedge \neg \mathbf{D}_{i,j} \varphi \rightarrow (\langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$
5. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi \rightarrow \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{B}_i^c \varphi$
6. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi \wedge \mathbf{Agn}_i^o \varphi \rightarrow \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$
7. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \mathbf{Agn}_i^d \varphi \rightarrow \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$
8. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \neg \mathbf{Agn}_i^d \varphi \rightarrow (\langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$
9. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Jumped}_j \varphi \rightarrow (\langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$

4.14. REMARK. Proposition 4.13 formalizes the intuitive ideas concerning communication that we previously exposed. Item 1 deals with the properties introduced at the beginning of Section 4. Item 2 states that the observational belief fluents — and hence also the knowledge fluents — of the agents persist under execution of an **inform** action. Item 3 states that agents have the opportunity to inform other agents of all the formulae that they themselves believe. Attempts to send non-beliefs are doomed to fail. Our agents are therefore not allowed to *gossip*, i.e., it is not allowed for an agent to spread around rumours that it itself does not even believe. Note that item 3 also shows that the **inform** action is not realizable. Item 4 deals with the dependence relation between the agents: if the receiving agent does not trust the sending agent, it lets the information pass without updating its beliefs (and changing anything else for that matter). Item 5 states that if some trustworthy agent j tells another agent i some formula that it either knows or observed, this leads to a state of affairs in which the receiving agent believes φ at least with the credibility attached to communicational beliefs; whenever i is *a priori* agnostic with regard to this formula on the level of communicational beliefs, the receiving agent actually *learns* φ (item 6). Items 7 and 8 deal with the situation of *hearsay*: agent j tells i some formula that j itself has heard. In this case the beliefs of the receiving agent are updated only if the agent is agnostic on the default level with respect to the formula. The ratio behind this validity is that agent i does not necessarily consider the source that agent j used to acquire its belief in φ to be reliable. It might for instance be the case that j heard φ from some agent j' that is trusted by j , but not by i ⁵. Item 9 states that default beliefs of an agent are not transferable: this notion of belief is that weak that no revision takes place upon hearing one of these beliefs.

4.3. Formalizing default jumps: jumping is believing

In a previous paper [31] we proposed a dynamic formalization of default reasoning, in which agents attempt to jump to certain plausible formulae, called *defaults*. A remarkable aspect of this approach is our representation of defaults as *common possibilities*⁶. The idea underlying this representation is that

⁵Huang [20] proposes transitivity of the dependency relation, but in our opinion this is not intuitively acceptable. Consider for example the case of president Reagan, a man to trust. However, some of the things Reagan told were whispered in his ear by his wife Nancy, who on her turn depended strongly on an astrologer for her information. And although we trust president Reagan, it is not the case that we trust Nancy's astrologer.

⁶Note, however, that our formalization of default reasoning does by no means depend on the representation of defaults as common possibilities. If one would wish to do so, one could

defaults are founded in *common sense*, and in our multi-agent architecture, *common sense* is related to the knowledge and lack of knowledge of *all* agents. The modality of common possibility captures this idea of defaults as determined by the (lack of) knowledge of all agents. The intuitive interpretation of some formula φ being a common possibility is that it is considered epistemically possible by all agents, i.e., none of the agents knows φ to be false.

4.15. DEFINITION ([31]). For all formulae φ , the formula $\mathbf{N}\varphi$, for nobody knows not φ , is defined by:

$$\mathbf{N}\varphi =^{\text{def}} \mathbf{M}_1\varphi \wedge \dots \wedge \mathbf{M}_n\varphi$$

The formalization of attempted jumps to defaults in the context of other sources of information is based on the idea that adopting formulae by default is the last resort: only if no other means of acquiring information are available a jump to a default is attempted. As such, default jumps are effective only for agents that are completely agnostic with respect to the default that is jumped to. If the agent is not completely agnostic, the attempted jump reduces to the empty action **skip**.

4.16. DEFINITION. For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ we define:

$$\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s) = \begin{cases} \emptyset & \text{if } \mathcal{M}, s \not\models \mathbf{N}\varphi \\ \mathbf{r}(e, \text{revise}^d(i, \varphi))(\mathcal{M}, s) & \text{if } \mathcal{M}, s \models \mathbf{N}\varphi \wedge \mathbf{Agn}_i^d\varphi \\ \mathcal{M}, s & \text{otherwise} \end{cases}$$

Also for this third kind of informative action, the function \mathbf{r} is defined correctly:

4.17. PROPOSITION. For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ , if $\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s) \neq \emptyset$ then it holds that $\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s)$ is a well-defined Kripke model.

4.18. PROPOSITION. For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' we have:

1. $\text{try_jump } \varphi$ is deterministic and idempotent for all $\varphi \in \mathcal{L}_0$
2. $\models \mathbf{B}_{i'}^c\psi \rightarrow [\text{do}_i(\text{try_jump } \varphi)]\mathbf{B}_{i'}^c\psi$
3. $\models \mathbf{N}\varphi \leftrightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt}$
4. $\models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt} \leftrightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \neg \mathbf{Agn}_i^d\varphi$
5. $\models \mathbf{N}\varphi \wedge \mathbf{Agn}_i^d\varphi \rightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{Jumped}_i\varphi$
6. $\models \mathbf{N}\varphi \wedge \neg \mathbf{Agn}_i^d\varphi \rightarrow (\langle \text{do}_i(\text{try_jump } \varphi) \rangle \chi \leftrightarrow \chi)$

4.19. REMARK. The first item of Proposition 4.18 deals with the properties introduced in Def. 4.5. Again it is obvious that the property of realizability is not validated since an agent may attempt to jump to a non-default, and in item 3

easily predefine the agents' defaults in a syntactical way (see also [31]).

it is formalized that such attempted jumps are doomed to fail. Item 2 states that the knowledge fluents, the observational belief fluents, and the communicational belief fluents of all agents persist under the attempted jump to a formula by one of them. Item 4 states that default jumps that are possible always result in the agent not being default-agnostic with regard to the formula that is jumped to. Item 5 formalizes the idea that agents that are completely agnostic with respect to some formula φ jump to new default beliefs by applying the `try_jump` φ action. The last item states that attempted jumps to default conclusions yield information for completely agnostic agents only.

5. The ability to gather information

The formalization of the abilities of intelligent information agents to execute informative actions is determined by two different notions. On the one hand, the abilities of an agent restrict its practical possibility to acquire information: only the actions that are within the agent’s capacities can be used as means to extend its information. On the other hand, the agent’s abilities steer the way in which it acquires its information: through its abilities an agent can be forced to acquire the most credible information possible. With respect to the latter point, we for instance demand it to be the case that agents resort to the use of default jumps only if they cannot acquire this information through observations.

Since by nature observations provide the most credible source of information, the ability to observe is determined strictly by the fact that the agents’ information gathering is limited. For artificial agents these limits are simply a consequence of the way these agents were built. One can for instance think of a group of robots, each equipped with its personal set of sensors: one robot can observe whether its environment is radioactive, another whether the atmosphere contains oxygen, but no robot can sense both. Being more or less a construction decision, we assume that these observational capacities of robots are determined on beforehand, i.e., with respect to the agents’ abilities observations are treated as atomic actions.

With regard to the capabilities of agents to inform other agents, we basically follow the approach we pursued previously [31]: an agent has the ability to inform all other agents of exactly the things it believes.

The ability to attempt to jump to a default captures two different aspects. Firstly, agents are able to jump to a default only if they *know* it to be a default. Secondly, the capability to jump depends on the observational capacities of the agent in the sense that an agent is (mentally) able to attempt a jump to a formula only if it knows that it is not able to observe whether the formula holds. In this way it is ensured that agents resort to default jumps only if the possibility of acquiring the information through observations is excluded.

5.1. DEFINITION. For Kripke models \mathcal{M} with states s , for agents i, j and formula $\varphi \in \mathcal{L}_0$, the function \mathbf{c} is modified such that

- it is predefined for all `observe` φ actions
- $\mathbf{c}(i, \text{observe } \varphi)(\mathcal{M}, s) = \mathbf{c}(i, \text{observe } \neg\varphi)(\mathcal{M}, s)$

- $c(j, \text{inform}(\varphi, i))(\mathcal{M}, s) = c(j, \text{confirm } \mathbf{B}_j^d \varphi)(\mathcal{M}, s)$
- $c(i, \text{try_jump } \varphi)(\mathcal{M}, s) = c(i, \text{confirm } \mathbf{K}_i(\neg \mathbf{A}_i \text{observe } \varphi \wedge \mathbf{N}\varphi))(\mathcal{M}, s)$

5.2. PROPOSITION. For agents i, j and formulae $\varphi \in \mathcal{L}_0$ we have:

1. $\models \mathbf{A}_j \text{inform}(\varphi, i) \leftrightarrow \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{tt}$
2. $\models \mathbf{A}_i \text{try_jump } \varphi \rightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt}$
3. $\models \mathbf{A}_i \text{observe } \varphi \rightarrow \neg \mathbf{A}_i \text{try_jump } \varphi$

5.3. REMARK. The first two items of Proposition 5.2 state that both the **inform** action and the **try_jump** action are *A-realizable* [18], i.e., being able to perform these action implies having the opportunity to do so. This property is typical for actions without circumstantial prerequisites, of which the **inform** and **try_jump** actions are typical examples. The last item formalizes the idea that agents attempt jumps to a formula only if it is not within their capacities to observe whether the formula holds.

When applying a definition like 5.1 for a given agent in a state of a model, the set of propositional formulae can be divided into four parts. One part consists of the formulae that the agent believes in one way or another, one part contains the formulae that the agent does not believe but of which it is able to observe truth, another part consists of all formulae that are not believed but jumpable for the agent, and yet another part consists of all unbelieved formulae that the agent may somehow be informed of. In this way it can for instance be modelled that an agent is *de facto omniscient* given its context. With this is meant that although the agent does not have complete information on all the things in its universe, given its capacities and opportunities and those of the agents surrounding it, it is not possible to acquire any information that the agent does not already possess.

6. Discussion

In this paper a formal framework is proposed in which various informative actions are combined, corresponding to the different ways in which rational agents can acquire information. In order to solve the various conflicts that could possibly occur when acquiring information from different sources, we propose a classification of the information that an agent possesses according to credibility. Based on this classification, we formalize what it means for agents to have seen or heard something, or to believe something by default. We present a formalization of observations, communication actions, and the attempted jumps to conclusions that constitute default reasoning. To implement these informative actions we use a general belief revision action which satisfies the AGM postulates; dependent on the credibility of the incoming information this revision action acts on one or more parts of the classified belief set of the agents. Through the agents' abilities it is both formalized that the capacities of agents to acquire information are limited, and that agents prefer one kind of (credible) information acquisition to another (less credible) one. The framework presented in this paper might be seen as a first step towards a formalization of intelligent

information agents which will help us overcome the information explosion and will play an important part in life in the information age.

Acknowledgements Thanks are due to Cristiano Castelfranchi for his very valuable suggestions on the psychological and philosophical aspects of the topics discussed in this paper; thanks are also due to Theo Huibers for his critical comments and remarks on a draft version of this paper. This research is partially supported by Esprit III BRWG project No.8319 ‘ModelAge’, Esprit III BRA project No.6156 ‘Drums II’, and the Vrije Universiteit Amsterdam; the third author is furthermore partially supported by the Katholieke Universiteit Nijmegen.

A. Appendix: the proofs

3.5. PROPOSITION. *Define the ordering $>$ on informational operators by $\mathbf{B}_i^k > \mathbf{B}_i^o > \mathbf{B}_i^c > \mathbf{B}_i^d$, and let \succeq be the reflexive, transitive closure of $>$. Then for all $\mathbf{X}, \mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$, and for all formulae φ :*

- $\mathbf{X} \succeq \mathbf{Y} \Rightarrow \models \mathbf{X}\varphi \rightarrow \mathbf{Y}\varphi$

PROOF: Let $i \in \mathcal{A}$ and $\mathbf{X}, \mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$. Since $\mathbf{B}^d(i, s) \subseteq \mathbf{B}^c(i, s) \subseteq [s]_{\mathbf{B}^o(i)} \subseteq [s]_{\mathbf{R}(i)}$ it follows that whenever $\mathbf{X} > \mathbf{Y}$ the set of informational alternatives used to determine truth of \mathbf{Y} is a subset of the set of informational alternatives used to determine truth of \mathbf{X} . Since \subseteq is reflexive and transitive it follows that the same holds whenever $\mathbf{X} \succeq \mathbf{Y}$. Now let φ be an arbitrary formula, and assume that $\mathbf{X} \succeq \mathbf{Y}$. Then it holds for all models \mathcal{M} with state s :

$$\begin{aligned} & \mathcal{M}, s \models \mathbf{X}\varphi \\ \Leftrightarrow & \varphi \text{ holds at all } \mathbf{X}\text{-informational alternatives of } s \\ \Rightarrow & \varphi \text{ holds at all } \mathbf{Y}\text{-informational alternatives of } s \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{Y}\varphi \end{aligned}$$

which suffices to conclude the proof of the proposition.

□

3.6. PROPOSITION. *Let $\mathbf{X} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o\}$ and $\mathbf{Y} \in \{\mathbf{B}_i^k, \mathbf{B}_i^o, \mathbf{B}_i^c, \mathbf{B}_i^d\}$. For all formulae φ we have:*

- | | |
|--|---|
| 1. $\models \mathbf{Y}(\varphi \rightarrow \psi) \rightarrow (\mathbf{Y}\varphi \rightarrow \mathbf{Y}\psi)$ | K |
| 2. $\models \neg(\mathbf{Y}\varphi \wedge \mathbf{Y}\neg\varphi)$ | D |
| 3. $\models \mathbf{X}\varphi \rightarrow \varphi$ | T |
| 4. $\models \mathbf{Y}\varphi \rightarrow \mathbf{Y}\mathbf{Y}\varphi$ | 4 |
| 5. $\models \neg\mathbf{Y}\varphi \rightarrow \mathbf{Y}\neg\mathbf{Y}\varphi$ | 5 |
| 6. $\models \varphi \Rightarrow \models \mathbf{Y}\varphi$ | N |

PROOF: By definition of R and \mathbf{B}^o being equivalence relations, \mathbf{B}_i^k and \mathbf{B}_i^o satisfy an S5 axiomatization, i.e., they validate the K-axiom (item 1), the T-axiom (item 3) — and thereby also the D-axiom (item 2) — the axioms of positive and negative introspection (items 4 and 5), and the rule of necessitation (item 6). A proof of this claim is standard and can be found elsewhere [6, 23, 35].

Defining B^c and B^d using belief clusters straightforwardly yields that B_i^c and B_i^d validate a KD45 axiomatization, corresponding to items 1, 2, 4, 5 and 6 given above (a proof of this claim can also be found elsewhere [34]).

☒

3.8. PROPOSITION. *Let φ be an i -doxastic sequenced formula, and let $\mathbf{Y} \in \{B_i^k, B_i^o, B_i^c, B_i^d\}$. Then:*

- $\models \mathbf{Y}\varphi \leftrightarrow \varphi$

PROOF: Let φ be an i -doxastic sequenced formula $\mathbf{X}\psi$, and let \mathcal{M} be some Kripke model with state s . Slightly abusing notation we define $X(i, s)$ to be $X(i, s)$ for $X = B^c, B^d$ and $[s]_{X(i)}$ for $X = R, B^o$. We distinguish two cases:

- $\mathbf{X} \in \{B_i^k, B_i^o, B_i^c, B_i^d\}$. In this case we have:

$$\begin{aligned} & \mathcal{M}, s \models \varphi \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{X}\psi \\ \Leftrightarrow & \forall s' \in X(i, s)[\mathcal{M}, s' \models \psi] \\ \stackrel{\star}{\Leftrightarrow} & \forall s'' \in Y(i, s)\forall s' \in X(i, s'')[\mathcal{M}, s' \models \psi] \\ \Leftrightarrow & \forall s' \in Y(i, s)[\mathcal{M}, s' \models \mathbf{X}\psi] \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{Y}\mathbf{X}\psi \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{Y}\varphi \end{aligned}$$

- $\mathbf{X} \in \{\neg B_i^k, \neg B_i^o, \neg B_i^c, \neg B_i^d\}$. In this case we have:

$$\begin{aligned} & \mathcal{M}, s \models \varphi \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{X}\psi \\ \Leftrightarrow & \exists s' \in X(i, s)[\mathcal{M}, s' \models \neg\psi] \\ \stackrel{\star}{\Leftrightarrow} & \forall s'' \in Y(i, s)\exists s' \in X(i, s'')[\mathcal{M}, s' \models \neg\psi] \\ \Leftrightarrow & \forall s'' \in Y(i, s)[\mathcal{M}, s'' \models \mathbf{X}\psi] \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{Y}\mathbf{X}\psi \\ \Leftrightarrow & \mathcal{M}, s \models \mathbf{Y}\varphi \end{aligned}$$

The equivalences tagged with \star hold since for $X, Y \in \{R, B^o, B^c, B^d\}$, $Y(i, s) \neq \emptyset$, $Y(i, s) \subseteq [s]_{R(i)}$ and $X(i, s'') = X(i, s)$ for all $s'' \in [s]_{R(i)}$. Since the two items given above capture all possible i -doxastic sequenced formulae φ we conclude that the proposition holds.

☒

3.11. PROPOSITION. *Let φ and ψ be formulae. Let \mathbf{X} be in the set $Bel = \{\mathbf{Saw}_i, \mathbf{Heard}_i, \mathbf{Jumped}_i \mid i \in \mathcal{A}\}$, and let $\mathbf{Y} \in Bel \cup \{\mathbf{K}_i\}$. Define the ordering \succeq' to be the reflexive and transitive closure of $>'$ with $\mathbf{K}_i >' \mathbf{Saw}_i >' \mathbf{Heard}_i >' \mathbf{Jumped}_i$. Then we have:*

- | | |
|---|----|
| 1. $\models \mathbf{X}\varphi \wedge \mathbf{X}(\varphi \rightarrow \psi) \rightarrow \mathbf{X}\psi$ | K |
| 2. $\models \neg(\mathbf{X}\varphi \wedge \mathbf{X}\neg\varphi)$ | D |
| 3. $\models \mathbf{Saw}_i\varphi \rightarrow \varphi$ | T |
| 4. $\not\models \mathbf{X}\varphi \rightarrow \mathbf{X}\mathbf{X}\varphi$ | 4 |
| 5. $\models \mathbf{X}\varphi \rightarrow \mathbf{K}_i\mathbf{X}\varphi$ | 4' |
| 6. $\not\models \neg\mathbf{X}\varphi \rightarrow \mathbf{X}\neg\mathbf{X}\varphi$ | 5 |
| 7. $\models \neg\mathbf{X}\varphi \rightarrow \mathbf{K}_i\neg\mathbf{X}\varphi$ | 5' |
| 8. $\models (\mathbf{X}\varphi \wedge \mathbf{X}\psi) \rightarrow \mathbf{X}(\varphi \wedge \psi)$ | C |

9.	$\not\models \mathbf{X}(\varphi \wedge \psi) \rightarrow (\mathbf{X}\varphi \wedge \mathbf{X}\psi)$	M
10.	$\models \mathbf{X}(\varphi \wedge \psi) \rightarrow (\bigvee_{\mathbf{Y} \succeq \mathbf{X}} \mathbf{Y}\varphi \wedge \bigvee_{\mathbf{Y} \succeq \mathbf{X}} \mathbf{Y}\psi) \wedge (\mathbf{X}\varphi \vee \mathbf{X}\psi)$	M'
11.	$\models \varphi \not\equiv \mathbf{X}\varphi$ and even $\models \varphi \Rightarrow \models \neg \mathbf{X}\varphi$	N
12.	$\models \varphi \rightarrow \psi \not\equiv \mathbf{X}\varphi \rightarrow \mathbf{X}\psi$	RM
13.	$\models \varphi \rightarrow \psi \Rightarrow \models \mathbf{X}\varphi \rightarrow \bigvee_{\mathbf{Y} \succeq \mathbf{X}} \mathbf{Y}\psi$	RM'
14.	$\models \varphi \leftrightarrow \psi \Rightarrow \models \mathbf{X}\varphi \leftrightarrow \mathbf{X}\psi$	RE

PROOF: Since the various case are highly analogous for the different operators, we restrict ourselves to proving the cases for the \mathbf{Heard}_i operator. At various places in the proofs we use the equivalent formulations of the derived belief operators that are proposed in Remark 3.10. Let \mathcal{M} be some Kripke model with state s , and let φ and ψ be formulae.

1. Suppose $\mathcal{M}, s \models \mathbf{Heard}_i\varphi \wedge \mathbf{Heard}_i(\varphi \rightarrow \psi)$, i.e., $\mathcal{M}, s \models \mathbf{B}_i^c\varphi \wedge \neg\mathbf{B}_i^o\varphi \wedge \mathbf{B}_i^c(\varphi \rightarrow \psi) \wedge \neg\mathbf{B}_i^o(\varphi \rightarrow \psi)$. Then since \mathbf{B}_i^c validates the K axiom, it holds that $\mathcal{M}, s \models \mathbf{B}_i^c\psi$, hence to show $\mathcal{M}, s \not\models \mathbf{B}_i^o\psi$. Assume towards a contradiction that $\mathcal{M}, s \models \mathbf{B}_i^o\psi$. Then also $\mathcal{M}, s \models \mathbf{B}_i^o(\varphi \rightarrow \psi)$ since \mathbf{B}_i^o validates the K-axiom. Since this contradicts the fact that $\mathcal{M}, s \models \neg\mathbf{B}_i^o(\varphi \rightarrow \psi)$, we conclude that $\mathcal{M}, s \not\models \mathbf{B}_i^o\psi$. Hence $\mathcal{M}, s \models \mathbf{Heard}_i\psi$.
2. Since \mathbf{B}_i^c satisfies the D-axiom, it holds that $\mathcal{M}, s \models \neg(\mathbf{B}_i^c\varphi \wedge \mathbf{B}_i^c\neg\varphi)$. Hence also $\mathcal{M}, s \models \neg(\mathbf{Heard}_i\varphi \wedge \mathbf{Heard}_i\neg\varphi)$.
3. If $\mathcal{M}, s \models \mathbf{Saw}_i\varphi$ this implies that $\mathcal{M}, s \models \mathbf{B}_i^o\varphi$. Since \mathbf{B}_i^o validates the T-axiom it follows that $\mathcal{M}, s \models \varphi$.
4. This item is most easily proved as a direct consequence of the following one. For whenever $\mathcal{M}, s \models \mathbf{Heard}_i\varphi$ it follows that $\mathcal{M}, s \models \mathbf{K}_i\mathbf{Heard}_i\varphi$. Hence $\mathcal{M}, s \models \mathbf{B}_i^o\mathbf{Heard}_i\varphi$ and thus $\mathcal{M}, s \not\models \mathbf{Heard}_i\mathbf{Heard}_i\varphi$.
5. Assume $\mathcal{M}, s \models \mathbf{Heard}_i\varphi$, i.e., $\mathcal{M}, s \models \mathbf{B}_i^c\varphi \wedge \neg\mathbf{B}_i^o\varphi$. By Proposition 3.8 it follows that $\mathcal{M}, s \models \mathbf{K}_i\mathbf{B}_i^c\varphi \wedge \mathbf{K}_i\neg\mathbf{B}_i^o\varphi$ and since \mathbf{K}_i satisfies the C-axiom we have $\mathcal{M}, s \models \mathbf{K}_i(\mathbf{B}_i^c\varphi \wedge \neg\mathbf{B}_i^o\varphi)$ and thus $\mathcal{M}, s \models \mathbf{K}_i\mathbf{Heard}_i\varphi$.
6. This item is a direct consequence of the following one for the same reason as given in item 4.
7. Assume $\mathcal{M}, s \models \neg\mathbf{Heard}_i\varphi$. Then $\mathcal{M}, s \models \neg\mathbf{B}_i^c\varphi \vee \mathbf{B}_i^o\varphi$. From Proposition 3.8 it follows that $\mathcal{M}, s \models \mathbf{K}_i\neg\mathbf{B}_i^c\varphi \vee \mathbf{K}_i\mathbf{B}_i^o\varphi$. This implies that $\mathcal{M}, s \models \mathbf{K}_i(\neg\mathbf{B}_i^c\varphi \vee \mathbf{B}_i^o\varphi)$ and hence $\mathcal{M}, s \models \mathbf{K}_i\neg\mathbf{Heard}_i\varphi$.
8. Assume $\mathcal{M}, s \models \mathbf{Heard}_i\varphi \wedge \mathbf{Heard}_i\psi$. Then $\mathcal{M}, s \models \mathbf{B}_i^c\varphi \wedge \neg\mathbf{B}_i^o\varphi \wedge \mathbf{B}_i^c\psi \wedge \neg\mathbf{B}_i^o\psi$, and thus $\mathcal{M}, s \models \mathbf{B}_i^c(\varphi \wedge \psi) \wedge \neg\mathbf{B}_i^o(\varphi \wedge \psi)$. Hence $\mathcal{M}, s \models \mathbf{Heard}_i(\varphi \wedge \psi)$.
9. If ψ is a tautology, i.e., $\models \psi$, then also $\models \mathbf{K}_i\psi$. For ψ a tautology, it is easy to see that the formula $\mathbf{Heard}_i(\varphi \wedge \psi) \wedge \mathbf{Heard}_i\varphi \wedge \neg\mathbf{Heard}_i\psi$ is satisfiable for some contingency φ .
10. Assume $\mathcal{M}, s \models \mathbf{Heard}_i(\varphi \wedge \psi)$. Since $\models (\varphi \wedge \psi) \rightarrow \varphi$ and $\models (\varphi \wedge \psi) \rightarrow \psi$ both hold, the first of the conjuncts on the right-hand side follows more or less directly from the RM' rule. For the second conjunct note that $\mathcal{M}, s \models \neg\mathbf{Heard}_i\varphi \wedge \neg\mathbf{Heard}_i\psi$ together with $\mathcal{M}, s \models \mathbf{Heard}_i(\varphi \wedge \psi)$ would imply $\mathcal{M}, s \models \mathbf{B}_i^o(\varphi \wedge \psi)$ which contradicts $\mathcal{M}, s \models \mathbf{Heard}_i(\varphi \wedge \psi)$.
11. If $\models \varphi$ then also $\models \mathbf{K}_i\varphi$ and hence directly $\models \neg\mathbf{Saw}_i\varphi \wedge \neg\mathbf{Heard}_i\varphi \wedge \neg\mathbf{Jumped}_i\varphi$.

12. If ψ is a tautology, i.e., $\models \psi$ then also $\models (\varphi \rightarrow \psi)$ and furthermore by the previous clause $\models \neg \mathbf{Heard}_i \psi$. Then for an appropriate contingency φ the formula $\mathbf{Heard}_i \varphi \wedge \neg \mathbf{Heard}_i \psi$ is easily satisfiable.
13. Assume that $\models (\varphi \rightarrow \psi)$ and $\mathcal{M}, s \models \mathbf{Heard}_i \varphi$, i.e., $\mathcal{M}, s \models \mathbf{B}_i^c \varphi \wedge \neg \mathbf{B}_i^o \varphi$. Then also $\mathcal{M}, s \models \mathbf{B}_i^c \psi$. Now if $\mathcal{M}, s \not\models \mathbf{B}_i^o \psi$, then $\mathcal{M}, s \models \mathbf{Heard}_i \psi$; otherwise if $\mathcal{M}, s \not\models \mathbf{K}_i \psi$ it holds that $\mathcal{M}, s \models \mathbf{Saw}_i \psi$, and else $\mathcal{M}, s \models \mathbf{K}_i \psi$. Hence $\mathcal{M}, s \models \bigvee_{\mathbf{Y} \succ \mathbf{r}} \mathbf{Heard}_i \mathbf{Y} \psi$.
14. Suppose $\models \varphi \leftrightarrow \psi$ and $\mathcal{M}, s \models \mathbf{Heard}_i \varphi$. Then $\mathcal{M}, s \models \mathbf{B}_i^c \varphi \wedge \neg \mathbf{B}_i^o \varphi$ and thus $\mathcal{M}, s \models \mathbf{B}_i^c \psi \wedge \neg \mathbf{B}_i^o \psi$. Hence $\mathcal{M}, s \models \mathbf{Heard}_i \psi$.

⊠

4.7. PROPOSITION. *For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ , $\mathbf{r}(i, \mathbf{observe} \varphi)(\mathcal{M}, s)$ is a well-defined Kripke model.*

PROOF: Assume that \mathcal{M} is a (well-defined) Kripke model with state s , let i be an agent and φ some propositional formula. It is easily seen from Definitions 4.3 and 4.6 that $\mathbf{r}(i, \mathbf{observe} \varphi)(\mathcal{M}, s)$ yields a pair \mathcal{M}', s where \mathcal{M}, s is a tuple in shape similar to a Kripke model. We show that \mathcal{M}' is indeed a well-defined Kripke model. We distinguish three cases, corresponding to the three cases in Definition 4.6.

- $\mathcal{M}, s \models \neg \mathbf{Agn}_i^k \varphi$. In this case $\mathcal{M}' = \mathcal{M}$ and hence \mathcal{M}' is well-defined.
- $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \varphi$. In this case $\mathcal{M}', s = \mathbf{r}(e, \mathbf{revise}^o(i, \varphi))(\mathcal{M}, s)$. From Definition 4.3 it follows that \mathcal{M}' differs from \mathcal{M} only in $\mathbf{B}^{o'}(i)$, $\mathbf{B}^{o'}(i, s')$ and $\mathbf{B}^{d'}(i, s')$, where $s' \in [s]_{\mathbf{R}(i)}$. To check well-definedness of \mathcal{M}' it suffices to consider these three aspects, which we shall do successively.
 - By definition of $\mathbf{r}(e, \mathbf{revise}^3(\varphi, i))$ we have that $\mathbf{B}^{o'}(i) = (\mathbf{B}^o(i) \setminus \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)})) \cup \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi+}) \cup \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi-})$. Since $\mathcal{M}, s \models \varphi$ it follows that $[s]_{\mathbf{B}^o(i)}^{\varphi+} = [s]_{\mathbf{B}^o(i)} \cap \llbracket \varphi \rrbracket$ and $[s]_{\mathbf{B}^o(i)}^{\varphi-} = [s]_{\mathbf{B}^o(i)} \cap \llbracket \neg \varphi \rrbracket$. We show that $\mathbf{B}^{o'}(i)$ is an equivalence relation by successively showing that it is reflexive, symmetrical and transitive.
 - It is obvious that $(s', s') \in \mathbf{B}^{o'}(i)$ for all $s' \in \mathcal{S}$. For if $s' \notin [s]_{\mathbf{B}^o(i)}$, $(s', s') \in \mathbf{B}^o(i) \setminus \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)})$, and otherwise s' is either in $[s]_{\mathbf{B}^o(i)}^{\varphi+}$ or in $[s]_{\mathbf{B}^o(i)}^{\varphi-}$, which implies that either $(s', s') \in \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi+})$ or $(s', s') \in \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi-})$. In either case $(s', s') \in \mathbf{B}^{o'}(i)$ and hence $\mathbf{B}^{o'}(i)$ is reflexive.
 - Let (s_1, s_2) in $\mathbf{B}^{o'}(i)$. If (s_1, s_2) is either in $\text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi+})$ or in $\text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi-})$ it follows by definition of Cl_{eq} that either $(s_2, s_1) \in \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi+})$ or $(s_2, s_1) \in \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}^{\varphi-})$, hence $(s_2, s_1) \in \mathbf{B}^{o'}(i)$. So assume that $(s_1, s_2) \in \mathbf{B}^o(i) \setminus \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)})$. Then $(s_1, s_2) \in \mathbf{B}^o(i)$ and $(s_1, s_2) \notin \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)})$. Since $\mathbf{B}^o(i)$ is symmetrical, it follows that $(s_2, s_1) \in \mathbf{B}^o(i)$ and by definition of Cl_{eq} we have that $(s_2, s_1) \notin \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)})$. Hence $(s_2, s_1) \in \mathbf{B}^o(i) \setminus \text{Cl}_{\text{eq}}([s]_{\mathbf{B}^o(i)}) \subseteq \mathbf{B}^{o'}(i)$ which suffices to collude that $\mathbf{B}^{o'}(i)$ is symmetrical.

- Let $(s_1, s_2) \in B^{o'}(i)$ and $(s_2, s_3) \in B^{o'}(i)$. We distinguish three cases:
 1. $(s_1, s_2) \in \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}^{\varphi+})$. Then $s_2 \in [s]_{B^{o'}(i)}^{\varphi+}$. From $(s_2, s_3) \in B^{o'}(i)$ it follows that also $s_3 \in [s]_{B^{o'}(i)}^{\varphi+}$. Hence $(s_2, s_3) \in \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}^{\varphi+})$. By definition of Cl_{eq} it follows that $(s_1, s_3) \in \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}^{\varphi+}) \subseteq B^{o'}(i)$.
 2. $(s_1, s_2) \in \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}^{\varphi-})$. This case is completely analogous to the case where $(s_1, s_2) \in \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}^{\varphi+})$.
 3. $(s_1, s_2) \in B^{o'}(i) \setminus \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)})$. In this case $(s_1, s_2) \in B^{o'}(i)$ and $(s_1, s_2) \notin \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)})$. Since $B^{o'}(i)$ is an equivalence relation this implies that $s_1 \notin [s]_{B^{o'}(i)}$ and $s_2 \notin [s]_{B^{o'}(i)}$: for if either one of them would be in $[s]_{B^{o'}(i)}$, they would both be due to the transitivity of $B^{o'}(i)$, and this contradicts $(s_1, s_2) \notin \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)})$. From $s_2 \notin [s]_{B^{o'}(i)}$ it follows that $(s_2, s_3) \notin \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)})$, and thus $(s_2, s_3) \in B^{o'}(i) \setminus \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)})$. But then $(s_1, s_3) \in B^{o'}(i)$ and since $s_1 \notin [s]_{B^{o'}(i)}$ it follows that $(s_1, s_3) \in B^{o'}(i) \setminus \text{Cl}_{\text{eq}}([s]_{B^{o'}(i)}) \subseteq B^{o'}(i)$.

Hence $B^{o'}(i)$ is transitive.

Thus $B^{o'}(i)$ is reflexive, symmetrical and transitive, and hence an equivalence relation. Furthermore, $B^{o'}(i) \subseteq B^o(i) \subseteq [s]_{R(i)} = [s]_{R'(i)}$.

- Let $s' \in [s]_{R(i)}$. We distinguish two cases:
 - If $B^{c'}(i, s')$ is equal to $B^c(i, s) \cap \llbracket \varphi \rrbracket$, it follows that $B^{c'}(i, s') \neq \emptyset$. Since $B^c(i, s) \subseteq [s]_{B^{o'}(i)}$, and $[s]_{B^{o'}(i)} = [s]_{B^o(i)} \cap \llbracket \varphi \rrbracket$, it follows that $B^{c'}(i, s') \subseteq [s]_{B^{o'}(i)}$. In this case $B^{d'}(i, s')$ is either equal to $B^d(i, s) \cap \llbracket \varphi \rrbracket$, which would imply by a similar argument as given above that $B^{d'}(i, s') \neq \emptyset$ and $B^{d'}(i, s') \subseteq B^{c'}(i, s')$, or equal to $B^{c'}(i, s')$, which would also imply that $B^{d'}(i, s') \neq \emptyset$ and $B^{d'}(i, s') \subseteq B^{c'}(i, s')$. In both cases the resulting tuple would indeed be a well-defined Kripke model.
 - If $B^{c'}(i, s')$ is equal to $[s]_{B^{o'}(i)} \cap \llbracket \varphi \rrbracket = [s]_{B^{o'}(i)}$, this implies that $B^c(i, s) \cap \llbracket \varphi \rrbracket = \emptyset$, and hence also $B^d(i, s) \cap \llbracket \varphi \rrbracket = \emptyset$. In this case $B^{d'}(i, s') = B^{c'}(i, s') = [s]_{B^{o'}(i)}$, and hence \mathcal{M}' would also be a well-defined Kripke model.

Since it holds in both cases, we conclude that \mathcal{M}' is a well-defined Kripke model.

- $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \neg \varphi$. This case is completely analogous to the case where $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \varphi$.

Since in all three cases \mathcal{M}' is well-defined, we conclude that the proposition holds.

☒

4.8. PROPOSITION. *For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' we have:*

1. *observe φ is deterministic, realizable and idempotent for all $\varphi \in \mathcal{L}_0$*

2. $\models \mathbf{K}_{i'}\psi \leftrightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{K}_{i'}\psi$
3. $\models \langle \text{do}_i(\text{observe } \varphi) \rangle \chi \leftrightarrow \langle \text{do}_i(\text{observe } \neg\varphi) \rangle \chi$
4. $\models \langle \text{do}_i(\text{observe } \varphi) \rangle \neg \mathbf{Agn}_i^o \varphi$
5. $\models \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \varphi$
6. $\models \neg\varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \neg\varphi$
7. $\models \varphi \wedge \mathbf{Agn}_i^k \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \varphi$
8. $\models \neg\varphi \wedge \mathbf{Agn}_i^k \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \neg\varphi$
9. $\models \varphi \wedge (\mathbf{Heard}_i \neg\varphi \vee \mathbf{Jumped}_i \neg\varphi) \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \varphi$
10. $\models \varphi \wedge \mathbf{B}_i^c \neg\varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle ((\mathbf{B}_i^c \chi \leftrightarrow \mathbf{B}_i^o \chi) \wedge (\mathbf{B}_i^d \chi \leftrightarrow \mathbf{B}_i^c \chi))$

PROOF: Let \mathcal{M} be some Kripke model with state s , let φ, ψ be propositional formulae, let χ be an arbitrary formula, and let i, i' be agents.

1. Realizability and determinism follow directly from Proposition 4.7. Idempotence is easily shown by inspection of Definition 4.3.
2. Since observations leave the epistemic accessibility relation in a model intact, it follows that $\models \mathbf{K}_{i'}\psi \rightarrow [\text{do}_i(\text{observe } \varphi)]\mathbf{K}_{i'}\psi$. Due to the realizability and determinism of the **observe** action the box may be replaced by a diamond and the implication by an equivalence.
3. By inspection of Definition 4.6 it is easily seen that $\mathbf{r}(i, \text{observe } \varphi)(\mathcal{M}, s)$ and $\mathbf{r}(i, \text{observe } \neg\varphi)(\mathcal{M}, s)$ are identical.
4. Let $\mathcal{M}', s = \mathbf{r}(i, \text{observe } \varphi)(\mathcal{M}, s)$; \mathcal{M}', s exists due to the realizability of **observe** φ . If $\mathcal{M}, s \models \neg \mathbf{Agn}_i^k \varphi$ then $\mathcal{M}', s = \mathcal{M}, s$ and hence $\mathcal{M}', s \models \neg \mathbf{Agn}_i^k \varphi$ which implies that $\mathcal{M}', s \models \neg \mathbf{Agn}_i^o \varphi$. If $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \varphi$ then $\mathbf{B}^{o'}(i, s) = \mathbf{B}^o(i, s) \cap \llbracket \varphi \rrbracket$ by definition of $\mathbf{r}(e, \text{revise}^o \varphi)$. Hence $\mathcal{M}', s' \models \varphi$ for all $s' \in \mathbf{B}^{o'}(i, s)$ and thus $\mathcal{M}', s \models \mathbf{B}_i^o \varphi$. The case where $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \neg\varphi$ is completely analogous to the case where $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \wedge \varphi$.
5. By item 4 we have that $\mathcal{M}, s \models \langle \text{do}_i(\text{observe } \varphi) \rangle (\mathbf{B}_i^o \varphi \vee \mathbf{B}_i^o \neg\varphi)$. Now using that the truth value of propositional formulae in states of a model is dependent of the valuation π only, we conclude from $\mathcal{M}, s \models \varphi$ that $\mathcal{M}, s \models \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \varphi$.
6. The proof of this item is completely analogous to the previous one.
7. By item 5, we have that if $\mathcal{M}, s \models \varphi \wedge \mathbf{Agn}_i^k \varphi$ then it is also the case that $\mathcal{M}, s \models \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{B}_i^o \varphi$. Since execution of the revision action does not affect the knowledge fluents in the model, it is furthermore the case that $\mathcal{M}, s \models \mathbf{Agn}_i^k \varphi \rightarrow \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Agn}_i^k \varphi$, which suffices to conclude that this item indeed holds.
8. The proof of this item is completely analogous to the previous one.
9. If $\mathcal{M}, s \models \varphi \wedge (\mathbf{Heard}_i \neg\varphi \vee \mathbf{Jumped}_i \neg\varphi)$ then also $\mathcal{M}, s \models \varphi \wedge \mathbf{Agn}_i^k \varphi$. From item 7 it then follows that $\mathcal{M}, s \models \langle \text{do}_i(\text{observe } \varphi) \rangle \mathbf{Saw}_i \varphi$.
10. Suppose $\mathcal{M}, s \models \varphi \wedge \mathbf{B}_i^c \neg\varphi$. Then obviously $\mathcal{M}, s \models \varphi \wedge \mathbf{Agn}_i^k \varphi$. Let $\mathcal{M}', s = \mathbf{r}(i, \text{observe } \varphi)(\mathcal{M}, s)$. By definition of $\mathbf{r}(e, \text{revise}^o(i, \varphi))$ it follows that
 - $\mathbf{B}^{o'}(i, s) = \mathbf{B}^o(i, s) \cap \llbracket \varphi \rrbracket$
 - $\mathbf{B}^{c'}(i, s) = \mathbf{B}^{o'}(i, s)$
 - $\mathbf{B}^{d'}(i, s) = \mathbf{B}^{o'}(i, s)$

Hence $\mathcal{M}', s \models \mathbf{B}_i^o \chi$ iff $\mathcal{M}', s \models \mathbf{B}_i^c \chi$ and $\mathcal{M}', s \models \mathbf{B}_i^c \chi$ iff $\mathcal{M}', s \models \mathbf{B}_i^d \chi$ which suffices to conclude item 10.

☒

4.12. PROPOSITION. *For all Kripke models \mathcal{M} with state s , and for all agents i, j and propositional formulae φ , if $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s) \neq \emptyset$ then it holds that $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s)$ is a well-defined Kripke model.*

PROOF: The proposition is proved by an argument similar to that given in the proof of Proposition 4.7.

☒

4.13. PROPOSITION. *For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' and j we have:*

1. $\mathbf{inform}(\varphi, i)$ is deterministic and idempotent for all $\varphi \in \mathcal{L}_0, i \in \mathcal{A}$
2. $\models \mathbf{B}_i^o \psi \rightarrow [\mathbf{do}_j(\mathbf{inform}(\varphi, i))] \mathbf{B}_i^o \psi$
3. $\models \mathbf{B}_j^d \varphi \leftrightarrow \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \mathbf{tt}$
4. $\models \mathbf{B}_j^d \varphi \wedge \neg \mathbf{D}_{i,j} \varphi \rightarrow (\langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$
5. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi \rightarrow \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \mathbf{B}_i^c \varphi$
6. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi \wedge \mathbf{Agn}_i^o \varphi \rightarrow \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$
7. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \mathbf{Agn}_i^d \varphi \rightarrow \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$
8. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \neg \mathbf{Agn}_i^d \varphi \rightarrow (\langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$
9. $\models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Jumped}_j \varphi \rightarrow (\langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \chi \leftrightarrow \chi)$

PROOF: Let \mathcal{M} be some Kripke model with state s , let i, j be agents, let φ be some propositional and let χ be an arbitrary formula.

1. Determinism and idempotence are easily shown by inspection of Definitions 4.3 and 4.11.
2. This item follows since communication does affect neither the epistemic accessibility relation nor the set of observational belief alternatives.
3. This item is straightforward from Definition 4.11, since $\mathbf{confirm}(\mathbf{B}_j^d \varphi)$ — and thereby $\mathbf{inform}(\varphi, i)$ — fails if and only if $\mathcal{M}, s \not\models \mathbf{B}_j^d \varphi$.
4. If $\mathcal{M}, s \models \mathbf{B}_j^d \varphi \wedge \neg \mathbf{D}_{i,j} \varphi$ then by the third clause of Definition 4.11, $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s) = \mathcal{M}, s$. Hence $\mathcal{M}, s \models \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \chi$ iff $\mathcal{M}, s \models \chi$.
5. Assume that $\mathcal{M}, s \models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi$. From Definitions 4.3 and 4.11 it is obvious that a model \mathcal{M}' exists such that $\mathbf{r}(j, \mathbf{inform}(\varphi, i))(\mathcal{M}, s) = \mathcal{M}', s$. Now let \mathcal{M}' be such a model. We distinguish three cases:
 - $\mathcal{M}, s \models \neg \mathbf{B}_i^c \neg \varphi$. In this case $\mathcal{M}', s = \mathbf{r}(e, \mathbf{revise}^c(i, \varphi))(\mathcal{M}, s)$ is such that $\mathbf{B}^{c'}(i, s) = \mathbf{B}^c(i, s) \cap \llbracket \varphi \rrbracket$. Hence $\mathcal{M}', s' \models \varphi$ for all $s' \in \mathbf{B}^{c'}(i, s)$ and $\mathcal{M}', s \models \mathbf{B}_i^c \varphi$. Thus $\mathcal{M}, s \models \langle \mathbf{do}_j(\mathbf{inform}(i, \varphi)) \rangle \mathbf{B}_i^c \varphi$.
 - $\mathcal{M}, s \models \mathbf{B}_i^c \neg \varphi \wedge \mathbf{Agn}_i^o \varphi$. Then $\mathcal{M}', s = \mathbf{r}(e, \mathbf{revise}^c(i, \varphi))(\mathcal{M}, s)$ is such that $\mathbf{B}^{c'}(i, s) = \mathbf{B}^o(i, s) \cap \llbracket \varphi \rrbracket$. Hence $\mathcal{M}', s' \models \varphi$ for all $s' \in \mathbf{B}^{c'}(i, s)$ and $\mathcal{M}', s \models \mathbf{B}_i^c \varphi$. Thus $\mathcal{M}, s \models \langle \mathbf{do}_j(\mathbf{inform}(i, \varphi)) \rangle \mathbf{B}_i^c \varphi$.
 - $\mathcal{M}, s \models \neg \mathbf{Agn}_i^o \varphi$. In this case $\mathcal{M}', s = \mathcal{M}, s$. Given the fact that $\mathcal{M}, s \models \mathbf{B}_j^o \varphi$ and that observational beliefs are veridical it follows that $\mathcal{M}, s \models \mathbf{B}_i^o \varphi$ and thus $\mathcal{M}, s \models \mathbf{B}_i^c \varphi$. Hence $\mathcal{M}, s \models \langle \mathbf{do}_j(\mathbf{inform}(\varphi, i)) \rangle \mathbf{B}_i^c \varphi$.

Since in all three cases $\mathcal{M}, s \models \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{B}_i^c \varphi$ we conclude that item 5 holds.

6. Assume that $\mathcal{M}, s \models \mathbf{D}_{i,j} \varphi \wedge \mathbf{B}_j^o \varphi \wedge \mathbf{Agn}_i^o \varphi$. By item 5 it follows that $\mathcal{M}, s \models \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{B}_i^c \varphi$. Let $\mathcal{M}', s = \mathbf{r}(j, \text{inform}(\varphi, i))(\mathcal{M}, s)$. By Definition 4.11 it follows that $\mathcal{M}', s = \mathbf{r}(e, \text{revise}^c(i, \varphi))(\mathcal{M}, s)$. Since the revise^c action does not affect observational belief fluents it follows that $\mathcal{M}', s \models \mathbf{Agn}_i^o \varphi$, and hence $\mathcal{M}', s \models \mathbf{Heard}_i \varphi$. Thus $\mathcal{M}, s \models \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$.
7. Assume that $\mathcal{M}, s \models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \mathbf{Agn}_i^d \varphi$. Let \mathcal{M}' be the model such that $\mathcal{M}', s = \mathbf{r}(j, \text{inform}(\varphi, i))(\mathcal{M}, s)$. From clause 2 of Definition 4.11 it follows that $\mathcal{M}', s = \mathbf{r}(e, \text{revise}^c(i, \varphi))(\mathcal{M}, s)$. Since $\mathcal{M}, s \models \mathbf{Agn}_i^d \varphi$ and hence $\mathcal{M}, s \models \neg \mathbf{B}_i^c \neg \varphi$, it follows that $\mathbf{B}^{c'}(i, s) = \mathbf{B}^c(i, s) \cap \llbracket \varphi \rrbracket$. Hence $\mathcal{M}', s \models \mathbf{B}_i^c \varphi$. Since $\mathcal{M}, s \models \mathbf{Agn}_i^d \varphi$ implies $\mathcal{M}, s \models \mathbf{Agn}_i^o \varphi$, it follows that $\mathcal{M}, s \models \mathbf{Agn}_i^o \varphi$, and since the revise^c action does not affect observational belief fluents it follows that $\mathcal{M}', s \models \mathbf{Agn}_i^o \varphi$. Hence $\mathcal{M}', s \models \mathbf{Heard}_i \varphi$ and therefore $\mathcal{M}, s \models \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \mathbf{Heard}_i \varphi$.
8. This item is straightforward from Definition 4.11: if $\mathcal{M}, s \models \mathbf{D}_{i,j} \varphi \wedge \mathbf{Heard}_j \varphi \wedge \neg \mathbf{Agn}_i^d \varphi$ it follows that $\mathbf{r}(j, \text{inform}(\varphi, i))(\mathcal{M}, s) = \mathcal{M}, s$ and hence $\mathcal{M}, s \models \langle \text{do}_j(\text{inform}(\varphi, i)) \rangle \chi$ iff $\mathcal{M}, s \models \chi$.
9. Item 9 is shown by an argument similar to the one proving item 8.

⊠

4.17. PROPOSITION. *For all Kripke models \mathcal{M} with state s , and for all agents i and propositional formulae φ , if $\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s) \neq \emptyset$ then it holds that $\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s)$ is a well-defined Kripke model.*

PROOF: Again the proof of this proposition is analogous to the one given for Proposition 4.7.

⊠

4.18. PROPOSITION. *For all propositional formulae φ, ψ , for arbitrary formulae χ , and for all agents i, i' we have:*

1. $\text{try_jump } \varphi$ is deterministic and idempotent for all $\varphi \in \mathcal{L}_0$
2. $\models \mathbf{B}_{i'}^c \psi \rightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{B}_{i'}^c \psi$
3. $\models \mathbf{N} \varphi \leftrightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt}$
4. $\models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt} \leftrightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \neg \mathbf{Agn}_i^d \varphi$
5. $\models \mathbf{N} \varphi \wedge \mathbf{Agn}_i^d \varphi \rightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{Jumped}_i \varphi$
6. $\models \mathbf{N} \varphi \wedge \neg \mathbf{Agn}_i^d \varphi \rightarrow (\langle \text{do}_i(\text{try_jump } \varphi) \rangle \chi \leftrightarrow \chi)$

PROOF: Let \mathcal{M} be some Kripke model with state s , let i be an agent, let φ be some propositional formula, and let χ be some arbitrary formula.

1. Determinism and idempotence follow by inspection of Definitions 4.3 and 4.16.
2. Since attempted default jumps affect only default belief fluents, the item follows directly.
3. If $\mathcal{M}, s \models \mathbf{N} \varphi$, \mathcal{M}' exists such that $\mathcal{M}', s = \mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s)$: $\mathcal{M}, s \models \neg \mathbf{Agn}_i^d \varphi$, $\mathcal{M}' = \mathcal{M}$ and else \mathcal{M}' exists by definition of \mathbf{r} for the revise^d action. The reverse implication is even so obvious.

4. Suppose $\mathcal{M}, s \models \mathbf{N}\varphi$. Let $\mathcal{M}', s = \mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s)$; \mathcal{M}' exists by item 3. We distinguish two cases:
- If $\mathcal{M}, s \not\models \mathbf{Agn}_i^d \varphi$ then $\mathcal{M}' = \mathcal{M}$. Hence $\mathcal{M}', s \models \neg \mathbf{Agn}_i^d \varphi$ and thus $\mathcal{M}, s \models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \neg \mathbf{Agn}_i^d \varphi$.
 - If $\mathcal{M}, s \models \mathbf{Agn}_i^d \varphi$ then $\mathcal{M}', s = \mathbf{r}(e, \text{revise}^d(i, \varphi))(\mathcal{M}, s)$. By definition of $\mathbf{r}(e, \text{revise}^d(i, \varphi))$ it follows that $\mathbf{B}^{d'}(i, s) = \mathbf{B}^d(i, s) \cap \llbracket \varphi \rrbracket$. Then $\mathcal{M}', s' \models \varphi$ for all $s' \in \mathbf{B}^{d'}(i, s)$ and $\mathcal{M}', s \models \mathbf{B}_i^d \varphi$ which implies that $\mathcal{M}, s \models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \neg \mathbf{Agn}_i^d \varphi$.
- Since in both cases $\mathcal{M}, s \models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \neg \mathbf{Agn}_i^d \varphi$ we conclude that the left-to-right implication of this item holds. The right-to-left implication is trivial.
5. Suppose $\mathcal{M}, s \models \mathbf{N}\varphi \wedge \mathbf{Agn}_i^d \varphi$. Let $\mathcal{M}', s = \mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s)$. From Definition 4.16 it follows that $\mathcal{M}', s = \mathbf{r}(e, \text{revise}^d(i, \varphi))(\mathcal{M}, s)$. Then from the definition of $\mathbf{r}(e, \text{revise}^d(i, \varphi))$ it follows that $\mathcal{M}', s \models \mathbf{B}_i^d \varphi$ by a similar argument as given above. From $\mathcal{M}, s \models \mathbf{Agn}_i^d \varphi$ it follows that $\mathcal{M}, s \models \mathbf{Agn}_i^c \varphi$. Since execution of the revise^d action does not affect the observational belief fluents it follows that $\mathcal{M}', s \models \mathbf{Agn}_i^c \varphi$ and thus $\mathcal{M}', s \models \mathbf{Jumped}_i \varphi$. Then also $\mathcal{M}, s \models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{Jumped}_i \varphi$.
6. Suppose $\mathcal{M}, s \models \mathbf{N}\varphi \wedge \neg \mathbf{Agn}_i^d \varphi$. Then it follows from Definition 4.16 that $\mathbf{r}(i, \text{try_jump } \varphi)(\mathcal{M}, s) = \mathcal{M}, s$. But then $\mathcal{M}, s \models \langle \text{do}_i(\text{try_jump } \varphi) \rangle \chi$ iff $\mathcal{M}, s \models \chi$.

⊠

5.2. PROPOSITION. *For agents i, j and formulae $\varphi \in \mathcal{L}_0$ we have:*

1. $\models \mathbf{A}_j \text{inform } (\varphi, i) \leftrightarrow \langle \text{do}_j(\text{inform } (\varphi, i)) \rangle \mathbf{tt}$
2. $\models \mathbf{A}_i \text{try_jump } \varphi \rightarrow \langle \text{do}_i(\text{try_jump } \varphi) \rangle \mathbf{tt}$
3. $\models \mathbf{A}_i \text{observe } \varphi \rightarrow \neg \mathbf{A}_i \text{try_jump } \varphi$

PROOF: All cases are rather straightforward by combining Definitions 4.11, 4.16 and 5.1.

⊠

References

- [1] C.E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.
- [2] Y. Arens, C.Y. Chee, C-N. Hsu, and C.A. Knoblock. Retrieving and integrating data from multiple information sources. *International Journal on Intelligent and Cooperative Information Systems*, 2(2):127–158, 1993.
- [3] J. Barwise. *The Situation in Logic*, volume 17 of *CSLI Lecture Notes*. CSLI, Stanford, 1989.
- [4] C. Castelfranchi. Private communication. 1994.

- [5] C. Castelfranchi. Guaranties for autonomy in cognitive agent architecture. In M. Wooldridge and N.R. Jennings, editors, *Intelligent Agents. Proceedings of the 1994 ECAI Workshop on Agent Theories, Architectures, and Languages*, volume 890 of *Lecture Notes in Artificial Intelligence*, pages 56–70. Springer-Verlag, 1994.
- [6] B.F. Chellas. *Modal Logic. An Introduction*. Cambridge University Press, Cambridge, 1980.
- [7] P.R. Cohen and H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.
- [8] B. Dunin-Keplicz and A. Radzikowska. Epistemic approach to actions with typical effects. Submitted, 1995.
- [9] B. Dunin-Keplicz and J. Treur. Compositional formal specification of multi-agent systems. In M. Wooldridge and N.R. Jennings, editors, *Intelligent Agents. Proceedings of the 1994 ECAI Workshop on Agent Theories, Architectures, and Languages*, volume 890 of *Lecture Notes in Artificial Intelligence*, pages 102–117. Springer-Verlag, 1994.
- [10] D. Elgesem. *Action Theory and Modal Logic*. PhD thesis, Institute for Philosophy, University of Oslo, Oslo, Norway, 1993.
- [11] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. The MIT Press, Cambridge, Massachusetts and London, England, 1988.
- [12] P. Gärdenfors, editor. *Belief Revision*. Cambridge University Press, 1992.
- [13] R. Goldblatt. *Logics of Time and Computation*, volume 7 of *CSLI Lecture Notes*. CSLI, Stanford, 1992. Second edition.
- [14] J.Y. Halpern and Y. Moses. A guide to completeness and complexity for modal logics of knowledge and belief. *Artificial Intelligence*, 54:319–379, 1992.
- [15] D. Harel. Dynamic logic. In D.M. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic*, volume 2, chapter 10, pages 497–604. D. Reidel, Dordrecht, 1984.
- [16] J. Hintikka. *Knowledge and Belief*. Cornell University Press, Ithaca, NY, 1962.
- [17] W. van der Hoek. Systems for knowledge and beliefs. *Journal of Logic and Computation*, 3(2):173–195, 1993.
- [18] W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer. A logic of capabilities. In A. Nerode and Yu. V. Matiyasevich, editors, *Proceedings of the Third International Symposium on the Logical Foundations of Computer Science (LFCS'94)*, volume 813 of *Lecture Notes in Computer Science*, pages 366–378. Springer-Verlag, 1994.

- [19] W. van der Hoek, B. van Linder, and J.-J. Ch. Meyer. Unravelling non-determinism: On having the ability to choose (extended abstract). In P. Jorrand and V. Sgurev, editors, *Proceedings of the Sixth International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA '94)*, pages 163–172. World Scientific, 1994.
- [20] Z. Huang. Logics for belief dependence. In E. Börger, H. Kleine Büning, M.M. Richter, and W. Schönfeld, editors, *Computer Science Logic, 4th Workshop CSL'90*, volume 533 of *Lecture Notes in Computer Science*, pages 274–288. Springer-Verlag, 1991.
- [21] Z. Huang. *Logics for Agents with Bounded Rationality*. PhD thesis, Universiteit van Amsterdam, 1994.
- [22] Z. Huang, M. Masuch, and L. Pólos. ALX, an action logic for agents with bounded rationality. Technical Report 92-70, Center for Computer Science in Organization and Management, University of Amsterdam, October 1992. To appear in *Artificial Intelligence*.
- [23] G.E. Hughes and M.J. Cresswell. *An Introduction to Modal Logic*. Routledge, London, 1968.
- [24] A. Kenny. *Will, Freedom and Power*. Basil Blackwell, Oxford, 1975.
- [25] S. Kraus and D. Lehmann. Knowledge, belief and time. *Theoretical Computer Science*, 58:155–174, 1988.
- [26] Y. Lesperance, H. Levesque, F. Lin, D. Marcu, R. Reiter, and R. Scherl. A logical approach to high-level robot programming – a progress report. To appear in *Control of the Physical World by Intelligent Systems*, Working Notes of the 1994 AAAI Fall Symposium, New Orleans, LA., 1994.
- [27] H. Levesque. Knowledge, action and ability in the situation calculus. Overheads from invited talk at TARK 1994.
- [28] A.Y. Levy, Y. Sagiv, and D. Srivastava. Towards efficient information gathering agents. In *Working Notes of the AAAI Spring Symposium on Software Agents*, pages 64–70, Stanford, California, 1994.
- [29] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Actions that make you change your mind: Belief revision in an agent-oriented setting. Technical Report UU-CS-1994-53, Utrecht University, December 1994.
- [30] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Communicating rational agents. In B. Nebel and L. Dreschler-Fischer, editors, *KI-94: Advances in Artificial Intelligence*, volume 861 of *Lecture Notes in Computer Science (subseries LNAI)*, pages 202–213. Springer-Verlag, 1994.
- [31] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. The dynamics of default reasoning. Technical Report UU-CS-1994-48, Utrecht University, October 1994.

- [32] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. Tests as epistemic updates. In A.G. Cohn, editor, *Proceedings of the 11th European Conference on Artificial Intelligence (ECAI'94)*, pages 331–335. John Wiley & Sons, 1994.
- [33] B. van Linder, W. van der Hoek, and J.-J. Ch. Meyer. How to motivate your agents - on preferences, goals and commitments. Technical report, Utrecht University, 1995.
- [34] J.-J. Ch. Meyer and W. van der Hoek. A modal logic for nonmonotonic reasoning. In W. van der Hoek, J.-J. Ch. Meyer, Y.H. Tan, and C. Witteveen, editors, *Non-Monotonic Reasoning and Partial Semantics*, pages 37–77. Ellis Horwood, Chichester, 1992.
- [35] J.-J. Ch. Meyer and W. van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995. To appear.
- [36] R.C. Moore. Reasoning about knowledge and action. Technical Report 191, SRI International, 1980.
- [37] R.C. Moore. A formal theory of knowledge and action. Technical Report 320, SRI International, 1984.
- [38] A.S. Rao and M.P. Georgeff. Asymmetry thesis and side-effect problems in linear time and branching time intention logics. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI'91)*, pages 498–504, 1991.
- [39] A.S. Rao and M.P. Georgeff. Modeling rational agents within a BDI-architecture. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484, San Mateo CA, 1991. Morgan Kaufmann.
- [40] A.S. Rao and M.P. Georgeff. A model-theoretic approach to the verification of situated reasoning systems. In *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence (IJCAI'93)*, pages 318–324, 1993.
- [41] Y. Shoham. Agent-oriented programming. *Artificial Intelligence*, 60:51–92, 1993.
- [42] Y. Shoham and S.B. Cousins. Logics of mental attitudes in AI. In G. Lakemeyer and B. Nebel, editors, *Foundations of Knowledge Representation and Reasoning*, volume 810 of *Lecture Notes in Computer Science (sub-series LNAI)*, pages 296–309. Springer-Verlag, 1994.
- [43] S.R. Thomas. *PLACA, An Agent Oriented Programming Language*. PhD thesis, Department of Computer Science, Stanford University, Stanford CA, September 1993. Appeared as technical report STAN-CS-93-1487.

- [44] M. Wooldridge. *The Logical Modelling of Computational Multi-Agent Systems*. PhD thesis, Department of Computation, UMIST, Manchester, October 1994. Appeared as technical report MMU-DOC-94-01.
- [45] M. Wooldridge and M. Fisher. A first-order branching time logic of multi-agent systems. In B. Neumann, editor, *Proceedings of the 10th European Conference on Artificial Intelligence (ECAI'92)*, pages 234–238. John Wiley & Sons, 1992.
- [46] G.H. von Wright. *Norm and Action*. Routledge & Kegan Paul, London, 1963.