TU Clausthal
Clausthal University of Technology

Agents, Beliefs, and Plausible Behavior in a Temporal Setting

Nils Bulling and Wojciech Jamroga

IfI Technical Report Series **IfI-06-05**

IfI

Department of Informatics
Clausthal University of Technology

Impressum

Publisher: Institut für Informatik, Technische Universität Clausthal
Julius-Albert Str. 4, 38678 Clausthal-Zellerfeld, Germany
Editor of the series: Jürgen Dix
Technical editor: Wojciech Jamroga
Contact: wjamroga@in.tu-clausthal.de
URL: http://www.in.tu-clausthal.de/forschung/technical-reports/
ISSN: 1860-8477

The IfI Review Board

Agents, Beliefs, and Plausible Behavior in a Temporal Setting

Nils Bulling and Wojciech Jamroga

Department of Informatics,
Clausthal University of Technology, Germany
{bulling,wjamroga}@in.tu-clausthal.de

Abstract

In this paper, we add a notion of *plausible behavior* to the branching-time logic CTL so that we obtain a language to reason about what can (or must) plausibly happen. Moreover, we propose a non-standard notion of beliefs, which is defined in terms of epistemic relations and plausibility – and we investigate properties of plausibility, knowledge and beliefs in this new framework. In particular, we show that knowledge is still an $S5$ modality, and that beliefs satisfy axioms $K45$ in general, and $KD45$ for the class of so called plausibly serial models. Finally, we show that the relationship between knowledge and beliefs for plausibly serial models is very natural and reflects the initial intuition well.

Keywords: multi-agent systems, temporal logic, plausibility, beliefs.

1 Introduction

Notions like time, knowledge, and beliefs seem to be very important for analyzing the behavior of agents and multi-agent systems. Modal logics have proved successful in providing a natural and intuitive theoretical framework, in which these (and other) notions can be modeled and investigated. In this paper, we extend modal logics of time and knowledge to consider a concept of *plausible behavior*. To this end, we add the concept to the models and language of CTLK [11], which is a straightforward combination of the branching-time temporal logic CTL [3, 2] and standard epistemic logic [6]. In our approach, plausibility is seen as a temporal property of behaviors. That is, some behaviors of the system can be assumed plausible and others implausible, with the underlying idea that the latter should be perhaps ignored in practical reasoning about possible future courses of action. Moreover, behaviors can be formally understood as temporal paths in the Kripke structure modeling a multi-agent system. In consequence, we obtain a language to reason about what can (or must) plausibly happen. We propose a non-standard notion of beliefs (inspired by [12]), defined in terms of epistemic relations and plausibility. The main intuition is that beliefs are facts *that an agent would know if he assumed that only plausible things could happen*.

We imply that humans use such a concept of plausibility and "practical beliefs" quite often in their everyday reasoning. Restricting one's reasoning to plausible possibilities is essential to make the reasoning feasible, as the space of *all* possibilities is exceedingly large in real life. We investigate some important properties of plausibility, knowledge, and beliefs in this new framework. In particular, we show that knowledge is an $S5$ modality, and that beliefs satisfy axioms $K45$ in general, and $KD45$ for the class of *plausibly serial models*. Finally, we show that the relationship between knowledge and beliefs for plausibly serial models is natural and reflects the initial intuition well. We also propose how plausibility assumptions can be specified in the object language via a *plausibility update operator*, and we study properties of such updates.

2 Branching Time and Agents' Knowledge

In this paper we build a framework for agents' beliefs about how the world can (or must) evolve. Thus, we need a notion of time and change, plus a notion of what the agents are supposed to know in particular situations. The logic of CTLK [11] seems to capture both dimensions in a natural way, and we will use it as the basis. CTLK is a straightforward combination of the computation tree logic CTL [3, 2] and standard epistemic logic [6]. CTL, on one hand, includes operators for temporal properties of systems: i.e., path quantifiers E ("there is a path") and A ("for every path"), together with temporal operators: $\bigcirc$ ("in the next state"), $\Box$ ("always from now on") and $\mathcal{U}$ ("until").[1] Every occurrence of a temporal operator is preceded by exactly one path quantifier in CTL (this variant of the language is sometimes called "vanilla" CTL). The broader language of CTL*, in which no such restriction is imposed, is not discussed here in order to keep things simpler. Epistemic logic, on the other hand, uses operators for representing agents' knowledge: $K_a\varphi$ is read as "agent a knows that φ".[2]

Let Π be a set of atomic propositions with a typical element p, and $\mathbb{A}\mathrm{gt} = \{1, ..., k\}$ be a set of agents with a typical element a. The language of CTLK consists of formulae φ, given as follows:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathsf{E}\gamma \mid \mathsf{A}\gamma \mid K_a\varphi$$
$$\gamma ::= \bigcirc\varphi \mid \Box\,\varphi \mid \varphi\,\mathcal{U}\,\varphi.$$

We will sometimes refer to formulae φ as ("vanilla") *state* formulae and to formulae γ as ("vanilla") *path* formulae.

The semantics of CTLK is based on Kripke models $\mathcal{M} = \langle Q, R, \sim_1, ..., \sim_k, \pi\rangle$, which include a nonempty set of states Q, a state transition relation $R \subseteq Q \times Q$, epistemic indistinguishability relations $\sim_a \subseteq Q \times Q$ (one per agent), and a valuation of

[1] An additional operator $\Diamond$ ("sometime") can be defined as $\Diamond\,\varphi \equiv \top\mathcal{U}\,\varphi$.
[2] We do not consider collective knowledge operators for the sake of simplicity.

propositions $\pi : \Pi \to \mathcal{P}(Q)$. We assume that relation R is serial and that all $\sim_a$ are equivalences. A *path* λ in $\mathcal{M}$ refers to a possible behavior (or computation) of system $\mathcal{M}$, and can be represented as an infinite sequence of states that follow relation R, that is, a sequence $q_0 q_1 q_2 \ldots$ such that $q_i R q_{i+1}$ for every $i = 0, 1, 2, \ldots$. We denote the ith state in path λ by $\lambda[i]$. A q*-path* is a path that starts from q, i.e., $\lambda[0] = q$. The set of all paths in $\mathcal{M}$ is denoted by $\Lambda_{\mathcal{M}}$ and the set of all q-paths by $\Lambda_{\mathcal{M}}(q)$ (if the model is clear from the context, $\mathcal{M}$ will be omitted). Now, the semantics of CTLK can be defined as below:

$\mathcal{M}, q \models p$	iff $q \in \pi(p)$;
$\mathcal{M}, q \models \neg\varphi$	iff $\mathcal{M}, q \not\models \varphi$;
$\mathcal{M}, q \models \varphi \wedge \psi$	iff $\mathcal{M}, q \models \varphi$ and $\mathcal{M}, q \models \psi$;
$\mathcal{M}, q \models \mathsf{E}\bigcirc\varphi$	iff there is a q-path λ such that $\mathcal{M}, \lambda[1] \models \varphi$;
$\mathcal{M}, q \models \mathsf{E}\Box\varphi$	iff there is a q-path λ such that $\mathcal{M}, \lambda[i] \models \varphi$ for every $i \geq 0$;
$\mathcal{M}, q \models \mathsf{E}\varphi\,\mathcal{U}\,\psi$	iff there is a q-path λ and $i \geq 0$ such that $\mathcal{M}, \lambda[i] \models \psi$ and $\mathcal{M}, \lambda[j] \models \varphi$ for every $0 \leq j < i$;
$\mathcal{M}, q \models K_a\varphi$	iff $\mathcal{M}, q \models \varphi$ for every q' such that $q \sim_a q'$.

The semantics of the universal path quantifier A is defined analogously.

3 The Concept of Plausibility

In this section we discuss the central concept of this paper, i.e. the concept of plausibility. First, we present related work [5, 12, 10]. Next, we introduce our own approach.

3.1 Friedman and Halpern: Plausibility Spaces

The work of Friedman and Halpern [5] extends the concepts of knowledge and belief with the notion of *plausibility*; i.e., some worlds can be more plausible for an agent than others. To implement this idea, Kripke models are extended with function P which assigns a *plausibility space* $P(q,a) = (\Omega_{(q,a)}, \preceq_{(q,a)})$ to every state $q \in Q$ and agent $a \in \mathbb{A}\mathrm{gt}$. The plausibility space is just a partially ordered subset of states; that is, $\Omega(q,a) \subseteq Q$, and $\preceq_{(q,a)} \subseteq Q \times Q$ is a reflexive and transitive relation. Let $S, T \subseteq \Omega_{(q,a)}$ be finite subsets of states; now, T is defined to be *plausible given S with respect to $P(q,a)$*, denoted by $S \to_{P(q,a)} T$, iff all minimal points/states in S (with respect to $\preceq_{(q,a)}$) are also in T.

The *language of knowledge and plausibility* is defined by the following grammar:

$$\varphi ::= p \mid \varphi \wedge \varphi \mid \neg\varphi \mid K_a\varphi \mid \varphi \to_a \varphi,$$

where the semantics of all operators except $\to_a$ is given as usual, and formulae $\varphi \to_a \psi$ have the meaning that ψ is true in the most plausible worlds in which φ holds. Formally, the semantics for $\to_a$ is given as:

$$\mathcal{M}, q \models \varphi \to_a \psi \text{ iff } S^{\varphi}_{P(q,a)} \to_{P(q,a)} S^{\psi}_{P(q,a)},$$

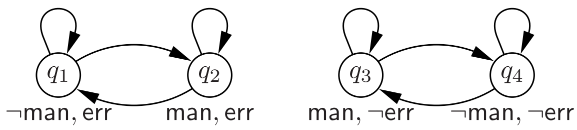

Figure 1: Kripke model for a communication domain.

where $S^{\varphi}_{(q,a)} = \{q' \in \Omega_{(q,a)} \mid \mathcal{M}, q' \models \varphi\}$ are the states in $\Omega_{(q,a)}$ that satisfy φ.

The idea of defining beliefs is given by the assumption that an agent *believes* in something if he *knows that it is true in the most plausible worlds of* $\Omega_{(q,a)}$; formally, this can be stated as $B_a\varphi \equiv K_a(\top \rightarrow_a \varphi)$.

Remark 1 *Note that:*

$$\begin{aligned}\mathcal{M}, q \models B_a\varphi \textit{ iff } & \mathcal{M}, q \models K_a(\top \rightarrow_a \varphi)\\ \textit{iff } & \forall q' \in Q(q \sim_a q' \Rightarrow \mathcal{M}, q' \models \top \rightarrow_a \varphi)\\ \textit{iff } & \forall q' \in Q(q \sim_a q' \Rightarrow S^{\top}_{P(q',a)} \rightarrow_{P(q',a)} S^{\varphi}_{P(q',a)})\\ \textit{iff } & \forall q' \in Q(q \sim_a q' \Rightarrow \Omega_{(q',a)} \rightarrow_{P(q',a)} S^{\varphi}_{P(q',a)}).\end{aligned}$$

The last line has the interpretation that all minimal points (with respect to $\preceq_{(q',a)}$) in $\Omega_{(q',a)}$ must be in $S^{\varphi}_{P(q',a)}$ for all states q' with $q \sim_a q'$.

We will provide an example to clarify the idea behind this concept.

Example 1 Consider an agent a who can receive messages. Using a check digit, he is able to recognize whether a received message contains an error (proposition error) or it is error-free (¬error). An error can have two sources. First, the message could have been manipulated (man) by someone, or second, the error might have occurred because of an inaccurate transmission (¬man). Agent a cannot distinguish between these two possible causes. Note that, even if the check digit is OK, a smart intruder could have manipulated the message (whereas a faulty transmission is impossible in the case of an error-free message).

Let $\mathcal{M}_1$ be the model shown in Figure 1. The epistemic relation is given by $\sim_a = \{(q_1, q_1), (q_2, q_2), (q_3, q_3),$
$(q_4, q_4), (q_1, q_2), (q_2, q_1), (q_3, q_4)(q_4, q_3)\}$. Suppose now that agent a receives a faulty message. In this case the current state is q_1 or q_2 (as the agent cannot distinguish between these states). Obviously, the agent does not know if the message was manipulated or not; i.e., $q_1 \not\models K_a\mathsf{man}$ and $q_2 \not\models K_a\mathsf{man}$.

We now define the plausibility space $P(q_1, a) = (\Omega_{(q_1,a)}, \preceq_{(q_1,a)})$ for state q_1. Suppose that messages are transmitted through a network with its own error correction mechanism. Then, bad transmission can be considered less plausible than manipulation by an intruder. That is, we have $q_2 \prec_{(q_1,a)} q_1$ for $\Omega_{(q_1,a)} = \{q_1, q_2\}$ (state q_2 is more plausible than state q_1). On the other hand, if the message is all right, it is not plausible that someone did manipulate it because up-to-date encryption software is used. In this case we define the following plausibility space for state q_3: $P(q_3, a) = (\Omega_{(q_3,a)}, \preceq_{(q_3,a)})$ with $\Omega_{(q_3,a)} = \{q_3, q_4\}$ and $q_4 \prec_{(q_3,a)} q_3$. Finally, we define the plausibility space for q_2 by $P(q_2, a) = P(q_1, a)$ and for state q_4 by $P(q_4, a) = P(q_3, a)$.

With these plausibility orderings the agent believes that someone manipulated a faulty message, but he but does not know it; i.e.,

$$\models (\mathsf{err} \rightarrow (B_a \mathsf{man} \wedge \neg K_a \mathsf{man})).$$

In the same way, he believes that an error-free message is not manipulated (but he but does not know it):

$$\models (\neg\mathsf{err} \rightarrow (B_a \neg\mathsf{man} \wedge \neg K_a \neg\mathsf{man})).$$

This is because only the most plausible worlds are considered for beliefs.

Friedman and Halpern have shown that the $KD45$ axioms are valid for operator B_a if plausibility spaces satisfy *consistency* (for all states $q \in Q$ it holds that $\Omega_{(q,a)} \subseteq \{ q' \in Q \mid q \sim_a q' \}$) and *normality* (for all states $q \in Q$ it holds that $\Omega_{(q,a)} \neq \emptyset$).[3] They also extended the language with time, using the interpreted systems approach developed in [7, 4].

3.2 Su et al.: KBC Logic

Su et al. [12] have developed a multi-modal, computationally grounded logic with modalities K, B, and C (knowledge, belief, and certainty). The semantics is given by an extension of interpreted systems. The computational model consists of (global) states $q = (q^{vis}, q^{inv}, q^{per}, Q^{pls})$ where the environment is divided into a visible (q^{vis}) and an invisible part (q^{inv}), and q^{per} captures the agent's perception of the visible part of the environment. External sources may provide the agent with information about the invisible part of a state, which result in a set of states Q^{pls} that are plausible for the agent. Given a global state q, we additionally define $Vis(q) = q^{vis}$, $Inv(q) = q^{inv}$, $Per(q) = q^{per}$, and $Pls(q) = Q^{pls}$.

A KBC *system* $\mathcal{R}$ is given by *runs*, where a run $r : \mathbb{N} \rightarrow Q$ is a function from time moments (modeled by $\mathbb{N}$) to global states, and a *point* (r, i) is given by a time point $i \in \mathbb{N}$ and a run r. An *interpreted* KBC *system* $\mathcal{M} = (\mathcal{R}, \pi)$ is given by a system $\mathcal{R}$ and a valuation of propositions π. KBC formulae are defined as $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi \mid B\varphi \mid C\varphi$. The epistemic relation $\sim_{vis}$ is captured in the following way: $(r, i) \sim_{vis} (r', i')$ iff $Vis(r(i)) = Vis(r'(i'))$. The semantic clauses are given below:

[3] Note that this "normality" is essentially seriality of states wrt plausibility spaces.

$\mathcal{M}, r, i \models p$	iff $p \in \Pi$ and $p \in \pi(Vis(r(i)), Inv(r(i)))$
$\mathcal{M}, r, i \models \neg\varphi$	iff $\mathcal{M}, r, i \not\models \varphi$
$\mathcal{M}, r, i \models \varphi \wedge \psi$	iff $\mathcal{M}, r, i \models \varphi$ and $\mathcal{M}, r, i \models \psi$
$\mathcal{M}, r, i \models K\varphi$	iff $\mathcal{M}, r', i' \models \varphi$ for all (r', i') with $(r, i) \sim_{vis} (r', i')$
$\mathcal{M}, r, i \models B\varphi$	iff $\mathcal{M}, r', i' \models \varphi$ for all (r', i') with $Vis(r'(i')) = Per(r(i))$ and $Inv(r'(i')) \in Pls(r(i))$
$\mathcal{M}, r, i \models C\varphi$	iff $\mathcal{M}, r', i' \models \varphi$ for all $(r'(i'))$ with $Vis(r'(i')) = Per(r(i))$

Thus, an agent believes that φ if, and only if, φ is true in all states *which look like what he sees now and seem plausible in the current state*. Certainty is stronger: if an agent is certain about φ, the formula must hold in all states with a visible part equal to the current perception, regardless of whether the invisible part is plausible or not.

The logic does not include temporal formulae, although it can be easily extended with temporal operators, as as time is already present in KBC models.

3.3 Moses and Shoham: Beliefs as Conditional Knowledge

In [5, 12], as well as in our approach (which will be introduced in Section 3.4), plausibility is used as a primitive semantic concept that helps to define beliefs on top of agents' knowledge. A similar idea was introduced by Moses and Shoham in [10]. In fact, their work preceded both [5] and [12] – and although Moses and Shoham do not mention the term "plausibility" in their paper, it seems appropriate to summarize the idea here.

In [10], beliefs are relativized with respect to a formula α (which can be seen as a plausibility assumption expressed in the object language). This concept is expressed via symbols $B_i^\alpha\varphi$; the index $i \in \{1, 2, 3\}$ is used to distinguish between three different implementations of beliefs. The first version is given by $B_1^\alpha\varphi \equiv K(\alpha \rightarrow \varphi)$.[4] A drawback of this version is that if α is false, then everything will be believed with respect to α. The second version overcomes this problem: $B_2^\alpha\varphi \equiv K(\alpha \rightarrow \varphi) \wedge (K\neg\alpha \rightarrow K\varphi)$; now φ is only believed if it is known that φ follows from assumption α, and φ must be known if assumption α is known to be false. Finally, $B_3^\alpha\varphi \equiv K(\alpha \rightarrow \varphi) \wedge \neg K\neg\alpha$: if the assumption α is known to be false, nothing should be believed with respect to α. The strength of these different notions is given as follows: $B_3^\alpha\varphi$ implies $B_2^\alpha\varphi$, and $B_2^\alpha\varphi$ implies $B_1^\alpha\varphi$. In this approach belief is strongly connected to knowledge in the sense that belief is knowledge with respect to a given assumption.

3.4 Our Approach: Plausibility as a Temporal Property

Plausibility can serve as a primitive concept that helps to define the semantics of beliefs, in a similar way as indistinguishability of states (represented by relation $\sim_a$) is the

[4] Unlike in most approaches, K is interpreted over *all* worlds and not only over the indistinguishable worlds.

semantic concept that underlies knowledge. In this sense, our work follows [5, 12]: essentially, beliefs are what an agent would know if he took only plausible options into account.

The work in [5, 12], however, attributes plausibility to states (possible worlds). [5] assumes orderings on worlds, and [12] provides agents with additional assumptions about the "invisible part" of each state. Thus, plausibility in [5, 12] is a *static property of states*. In our approach, plausibility is seen as a *temporal property*. That is, we do not consider states to be more plausible than others but rather define some behaviors to be plausible (and others implausible). Moreover, as we propose in Section 4, behaviors can be formally understood as temporal paths in the Kripke structure modeling a multi-agent system.

An actual notion of plausibility (that is, a particular set of plausible paths) can emerge in many different ways. It may result from observations and learning from the environment; an agent can learn from his observations and see specific patterns of events as plausible ("a lot of people wear black shoes if they wear a suit"). Knowledge exchange is another possibility (e.g., an agent a can tell agent b that "player c always bluffs when he is smiling"). Last but not least, folk knowledge is an important source of plausibility-related classifications of behavior ("players normally want to win a game", "people want to live").

We (i.e., the authors) point out that we (i.e., humans) use this (or a similar) concept of plausibility quite often in our everyday reasoning. Of course, we know that people do commit suicides, that players may sometimes be indifferent or even want to lose, and that there are some guys who really wear white sport shoes to a suit – but we usually disregard these possibilities when analyzing potential outcomes of our actions. Restricting the reasoning to plausible possibilities is essential to make the reasoning feasible, as the space of *all* possibilities (we call them "physical" possibilities in the rest of the paper) is exceedingly large in real life. A more extensive analysis must be conducted only in emergency, e.g. when our plausibility assumptions do not seem accurate any more ("my girlfriend looks depressed, I'd better take more care of her or she might do something bad to her"), or when the cost of inaccurate assumptions can be too high (like in the case of high-budget business decisions). And even in these cases, we do not get rid of plausibility assumptions completely – we only revise them to make them more cautious.[5]

To formalize this idea, we extend models of CTLK with *sets of plausible paths* and add plausibility operators $\mathbf{Pl}_a$, physical paths operator $\mathbf{Ph}$, and belief operators B_a to the language of CTLK. Now, it is possible to make statements that refer to plausible paths only, as well as statements that regard all paths that may occur in the system.

[5] That is, when planning to open an industrial plant in the UK, we will probably consider the possibility of our main contractor taking his life, but we will still *not* take into account the possibilities of: an invasion of UFO, England being destroyed by a meteorite, Fidel Castro becoming the British Prime Minister etc. Note that this is fundamentally different from using a probabilistic model in which all these unlikely scenarios are assigned very low probabilities: in that case, they also have a very small influence on our final decision, but we must process the *whole* space of physical possibilities to evaluate the options.

For instance, we may claim it is plausible to assume that a shop is closed after the opening hours, though the manager may be physically able to open it at any time: $\mathbf{Pl}_a \mathsf{A}\Box\,(\mathsf{late} \rightarrow \neg\mathsf{open}) \wedge \mathbf{Ph}\,\mathsf{E}\Diamond\,(\mathsf{late} \wedge \mathsf{open})$.

Finally, we want to point out that we see plausibility as a *subjective* property; i.e. every agent has his own notion of plausibility encoded in a model.

4 Extending Time and Knowledge with Plausibility and Beliefs

In this section, we extend the logic of CTLK with the notion of plausibility. We will call the resulting logic CTLKP. To implement our concept, we add *plausible path operators* $\mathbf{Pl}_a$ and *physical path operator* $\mathbf{Ph}$ to CTLK. Formula $\mathbf{Pl}_a\varphi$ has the intended meaning: *according to agent a, it is plausible that φ holds*; formula $\mathbf{Ph}\,\varphi$ reads as: *φ holds in all "physically" possible scenarios* (i.e., even in implausible ones). The plausible path operator restricts statements only to these paths which are defined to be "sensible", whereas the physical path operator generates statements about all paths that may theoretically occur. Furthermore, we define beliefs on top of plausibility and knowledge, as the facts *that an agent would know if he assumed that only plausible things could happen*.

4.1 CTLK with Plausibility

Formally, the language of CTLKP is defined as:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathsf{E}\gamma \mid \mathsf{A}\gamma \mid \mathbf{Pl}_a\varphi \mid \mathbf{Ph}\,\varphi \mid K_a\varphi \mid B_a\varphi$$
$$\gamma ::= \bigcirc\varphi \mid \Box\,\varphi \mid \varphi\,\mathcal{U}\,\varphi.$$

For example, we can now express the property that it is plausible to expect that an agent will not commit suicide; on the other hand, an agent is (always) physically able to commit that, and it is also plausible to expect that he has this physical ability:

$$\mathbf{Pl}_a\mathsf{A}\Box\,\neg\mathsf{suicide} \wedge \mathsf{A}\Box\,\mathbf{Ph}\,\mathsf{E}\Diamond\,\mathsf{suicide} \wedge \mathbf{Pl}_a\mathsf{A}\Box\,\mathbf{Ph}\,\mathsf{E}\Diamond\,\mathsf{suicide}.$$

The semantics of CTLKP extends that of CTLK as follows. First, we augment the models with *sets of plausible paths*. A *model with plausibility* is given as $\mathcal{M} = \langle Q, R, \sim_1, ..., \sim_k, \Upsilon_1, ..., \Upsilon_k, \pi\rangle$, where $\langle Q, R, \sim_1, ..., \sim_k, \pi\rangle$ is a CTLK model, and $\Upsilon_a \subseteq \Lambda_{\mathcal{M}}$ is the set of paths in $\mathcal{M}$ that are plausible according to agent a. If we want to make it clear that Υ_a is taken from model $\mathcal{M}$, we will write $\Upsilon_a^{\mathcal{M}}$. It seems worth emphasizing that this notion of plausibility is *subjective* and *global*. It is subjective because Υ_a represents *agent a's subjective view on what is plausible* – and indeed, different agents may have different ideas on plausibility (i.e., Υ_a may differ from Υ_b). It is global because Υ_a represents agent a's idea of the plausible behavior *of the whole system* (including the behavior of other agents).

Second, we use a non-standard satisfaction relation $\models_P$, which we call *plausible satisfaction*. Let $\mathcal{M}$ be a CTLKP model and $P \subseteq \Lambda_{\mathcal{M}}$ be an arbitrary subset of paths in $\mathcal{M}$ (not necessarily $\Upsilon^{\mathcal{M}}$). $\models_P$ restricts the evaluation of temporal formulae to the paths given in P only. The "absolute" satisfaction relation $\models$ is defined as $\models_{\Lambda_{\mathcal{M}}}$.

Let $\partial(P)$ be the set of all states from which at least one path in P starts, i.e. $\partial(P) = \{q \in Q \mid \exists\lambda \in P\ \lambda[0] = q)\}$. Now, the semantics of CTLKP can be given via the following clauses:

$\mathcal{M}, q \models_P p$	iff $q \in \pi(p)$;
$\mathcal{M}, q \models_P \neg\varphi$	iff $\mathcal{M}, q \not\models_P \varphi$;
$\mathcal{M}, q \models_P \varphi \wedge \psi$	iff $\mathcal{M}, q \models_P \varphi$ and $\mathcal{M}, q \models_P \psi$;
$\mathcal{M}, q \models_P \mathsf{E}\bigcirc\varphi$	iff there is a q-path $\lambda \in P$ such that $\mathcal{M}, \lambda[1] \models_P \varphi$;
$\mathcal{M}, q \models_P \mathsf{E}\Box\,\varphi$	iff there is a q-path $\lambda \in P$ such that $\mathcal{M}, \lambda[i] \models_P \varphi$ for every $i \geq 0$;
$\mathcal{M}, q \models_P \mathsf{E}\varphi\mathcal{U}\,\psi$	iff there is a q-path $\lambda \in P$ and $i \geq 0$ such that $\mathcal{M}, \lambda[i] \models_P \psi$, and $\mathcal{M}, \lambda[j] \models_P \varphi$ for every $0 \leq j < i$;
$\mathcal{M}, q \models_P \mathbf{Pl}_a\varphi$	iff $\mathcal{M}, q \models_{\Upsilon_a} \varphi$;
$\mathcal{M}, q \models_P \mathbf{Ph}\,\varphi$	iff $\mathcal{M}, q \models \varphi$;
$\mathcal{M}, q \models_P K_a\varphi$	iff $\mathcal{M}, q \models \varphi$ for every q' such that $q \sim_a q'$;
$\mathcal{M}, q \models_P B_a\varphi$	iff for all $q' \in \partial(\Upsilon_a)$ with $q \sim q'$, we have that $\mathcal{M}, q' \models_{\Upsilon_a} \varphi$.

Again, the semantics of the universal path quantifier A is defined analogously. One of the main reasons for using the concept of plausibility is that we want to define agents' *beliefs* out of more primitive concepts – in our case, these are plausibility and indistinguishability – in a way analogous to [12]. If an agent *knows* that φ, he must be "sure" about it. However, *beliefs* of an agent are not necessarily about reliable facts, and they can obviously be wrong. In spite of that, they should make sense to the agent; if he believes that φ, then the formula should at least hold in all futures that he envisages as plausible. Thus, beliefs of an agent may be seen as *things known to him if he disregards all non-plausible possibilities*.

We say that φ is $\mathcal{M}$*-true* ($\mathcal{M} \models \varphi$) if $\mathcal{M}, q \models \varphi$ for all $q \in Q_{\mathcal{M}}$. φ is *valid* ($\models \varphi$) if $\mathcal{M} \models \varphi$ for all models $\mathcal{M}$. φ is $\mathcal{M}$*-strongly true* ($\mathcal{M} \models\!\!\equiv \varphi$) if $\mathcal{M}, q \models_P \varphi$ for all $q \in Q_{\mathcal{M}}$ and all $P \subseteq \Lambda_{\mathcal{M}}$. φ is *strongly valid* ($\models\!\!\equiv \varphi$) if $\mathcal{M} \models\!\!\equiv \varphi$ for all models $\mathcal{M}$. Ultimately, we are going to be interested in normal (not strong) validity, as parameterizing the satisfaction relation with a set P is just a technical device for propagating sets of plausible paths Υ_a into the semantics of nested formulae.

Proposition 2 *Strong truth and strong validity imply truth and validity, respectively. The reverse does not hold in general.*

Proof. Strong truth and validity holds especially for $P = \Lambda$. For the reverse implication, see e.g. the proof of Axiom T in Theorem 6. ■

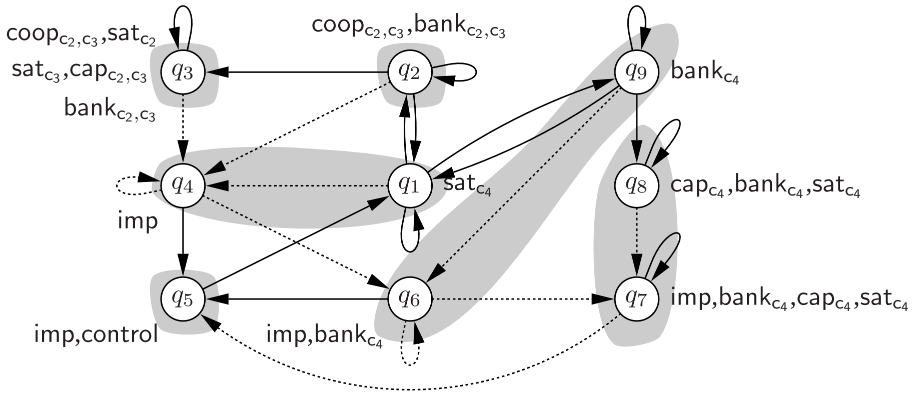

Figure 2: $CTLKP$ model, where all paths which do not contain a dotted line represent plausible paths of c_1 and the grey areas model incomplete information of c_1.

Corollary 3 *If φ is not valid, then φ is not strongly valid, and if φ is not $\mathcal{M}$-true, then φ is not $\mathcal{M}$-strongly true.*

Example 2 Figure 2 shows CTLKP model $\mathcal{M}_2$ which represents the following scenario. Company c_1 is insolvent and firms $c_2,\ c_3$, and c_4 are interested in taking over c_1. To this end, c_2 and c_3 may cooperate; on the other hand, c_4 can impend the other companies to prevent their cooperation. All firms need additional money for the takeover. Company c_1 has incomplete information about the world, modeled by relation $\sim_{c_1}$. The *set of plausible paths* Υ_{c_1}, according to company c_1, is given by all (infinite) paths that do *not* contain a dotted edge, e.g. $q_1 q_2 q_2 \ldots$[6] The following propositions are used:

sat$_\mathsf{i}$: company i is satisfied

bank$_\mathsf{i,j}$: companies i, j get money from the bank (j is optional)

coop$_\mathsf{i,j}$: cooperation of i and j.

cap$_\mathsf{i,j}$: companies i, j capture (take over) company c_1 (j is optional)

imp: company c_4 impends the cooperation of c_2 and c_3 (e.g. by dumping prices against c_2 and c_3)

control: companies c_2 and c_3 consult a control instance to check the conduct of c_4 (e.g. check on violations of market rules)

In q_1, company c_4 is satisfied because it remains the biggest company if nothing happens. In q_2, c_2 and c_3 have cooperated and borrowed money from the bank. Because

[6] Note that in general it is not possible to use this (finite) representation to capture an (infinite) set of plausible paths because plausibility of a transition often depends on previous transitions.

the manager of company c_4 does not like the newly created cooperation, he may decide to impend c_2 or c_3 so that they will break up their joining (this leads to state q_4; note that this course of events is implausible according to c_1). The threatened company can decide to consult a control instance to check for violation of market rules, which is modeled by state q_5. Then, the control instance has to deliberate about the case because it is not so obvious that a violation of the law occurred, but even c_4 becomes unsure of its possibly dubious acting, so all parties decide to "forget about everything" which results in a transition to q_1 (and, again, only company c_4 is satisfied).

Note that, for example, formula

$$\mathbf{Pl}_{c_1} \mathsf{E}\Diamond\, \mathsf{A}\Box\, (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4})$$

is true in $\mathcal{M}_2$ but

$$\mathbf{Ph}\, \mathsf{E}\Diamond\, \mathsf{A}\Box\, (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4}).$$

is not. We will carefully show these properties for any $q \in Q$:

$$\begin{aligned}\mathcal{M}, q \models & \mathbf{Pl}_{c_1} \mathsf{E}\Diamond\, \mathsf{A}\Box\, (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4}) \\ & \text{iff there is } \lambda \in \Upsilon_{c_1}(q) \text{ and } i \in \mathbb{N}_0 \\ & \quad \text{such that } \mathcal{M}, \lambda[i] \models_{\Upsilon_{c_1}} \mathsf{A}\Box\, (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4}) \\ & \text{iff there is } \lambda \in \Upsilon_{c_1}(q) \text{ such that for all } \lambda' \in \Upsilon_{c_1}(\lambda[i]),\ i \in \mathbb{N}_0, \\ & \quad \text{and for all } j \in \mathbb{N}_0 \text{ we have } \mathsf{sat}_{\mathsf{c}_4} \in \pi(\lambda'(j)) \text{ or } \mathsf{sat}_{\mathsf{c}_1} \in \pi(\lambda'(j)).\end{aligned}$$

On the other hand, $\mathsf{A}\Box\ (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4})$ is plausibly satisfied in states q_3, q_7, and q_8. Furthermore, it is easy to see that from all states $q \in Q\backslash\{q_7\}$ there is a plausible path from q to q_3 or q_8, and in q_7 the only plausible q_7-path is $q_7q_7q_7\ldots$. In the case of all possible scenarios, states in which c_2 or c_4 are satisfied can always be left, and therefore, $\mathbf{Ph}\, \mathsf{E}\Diamond\, \mathsf{A}\Box\, (\mathsf{sat}_{\mathsf{c}_1} \vee \mathsf{sat}_{\mathsf{c}_4})$ is not valid.

Another interesting property is that company c_4 is always physically able to impend the cooperation of c_2 and c_3, but it is not plausible that c_4 would ever impend the cooperation according to c_1's view of plausibility:

$$\models \mathsf{A}\Box\, \mathbf{Ph}\, (\mathsf{coop}_{\mathsf{c}_2,\mathsf{c}_3} \rightarrow \mathsf{E}\Diamond\, \mathsf{imp}) \wedge \mathbf{Pl}_{c_1}(\neg\mathsf{imp} \rightarrow \mathsf{A}\Box\, \neg\mathsf{impend}).$$

Furthermore, c_1 – having been captured by c_4 – believes that c_4 will always be satisfied, but does not know it for sure: $\models (\mathsf{cap}_{\mathsf{c}_4} \rightarrow B_{c_1}\mathsf{A}\Box\, \mathsf{sat}_{\mathsf{c}_4} \wedge \neg K_{c_1}\mathsf{A}\Box\, \mathsf{sat}_{\mathsf{c}_4})$.

4.2 Defining Plausible Paths with Formulae

So far, we have assumed that sets of plausible paths are somehow given in a model. In this section we present a dynamic approach where an actual notion of plausibility can be specified in the object language. Note that we want to specify (usually infinite) sets of infinite paths, and we need a finite representation of these structures. One logical solution is given by using path formulae γ. These formulae describe properties of paths;

therefore, a specific formula can be used to characterize a set of paths. For instance, think about a country in Africa where it should never snow; then plausible paths might be defined as ones in which it never snows, i.e., all paths that satisfy $\Box\,\neg\mathsf{snows}$. Formally, we define $|\gamma|_{\mathcal{M}}$ to be the set of paths that satisfy path formula γ in model $\mathcal{M}$ (when the model is clear from the context, the subscript will be omitted):

$$\begin{aligned}
|\bigcirc\varphi|_{\mathcal{M}} &= \{\lambda \mid \mathcal{M},\lambda[1] \models \varphi\}\\
|\Box\,\varphi|_{\mathcal{M}} &= \{\lambda \mid \forall i\,(\mathcal{M},\lambda[i] \models \varphi)\}\\
|\varphi_1\mathcal{U}\,\varphi_2|_{\mathcal{M}} &= \{\lambda \mid \exists i\big(\mathcal{M},\lambda[i] \models \varphi_2 \wedge \forall j(0 \le j < i \Rightarrow \mathcal{M},\lambda[j] \models \varphi_1)\big)\}.
\end{aligned}$$

Moreover, we define the *plausible paths model update* as follows. Let $\mathcal{M} = \langle Q, R, \sim_1, ..., \sim_k, \Upsilon_1, ..., \Upsilon_k, \pi\rangle$ be a CTLKP model, and let $P \subseteq \Lambda_{\mathcal{M}}$ be a set of paths. Then $\mathcal{M}^{a,P} = \langle Q, R, \sim_1, ..., \sim_k, \Upsilon_1, ..., \Upsilon_{a-1}, P, \Upsilon_{a+1}, ..., \Upsilon_k, \pi\rangle$ denotes model $\mathcal{M}$ with a's set of plausible paths reset to P. Note that the set of all paths remains the same in both models because the transition relation does not change, i.e., $\Lambda_{\mathcal{M}} = \Lambda_{\mathcal{M}^{a,P}}$.

Now we can extend the language of CTLKP with formulae $(\textbf{set-pl}_a\ \gamma)\varphi$ with the intuitive reading: "suppose that γ exactly characterizes the set of plausible paths, then φ holds", and formal semantics given below:

$$\mathcal{M}, q \models_P (\textbf{set-pl}_a\ \gamma)\varphi \text{ iff } \mathcal{M}^{a,|\gamma|}, q \models_P \varphi.$$

We observe that this update scheme is similar to the one proposed in [8].

Remark 4 *Note that the set of paths with which the satisfaction relation is annotated does not change after a plausible path update. Consider a CTLKP model* $\mathcal{M} = \langle Q, R, \sim_1, \ldots, \sim_k, \Upsilon_1, \ldots, \Upsilon_k, \pi\rangle$ *and statement*

$$\mathcal{M}, q \models_P (\textbf{\textit{set-pl}}_a\ \gamma)\varphi.$$

The semantic rules transform the formula into the equivalent notation

$$\mathcal{M}^{a,|\gamma|}, q \models_P \varphi.$$

But the set of paths P*, with which the satisfaction relation is indexed, is still the same as before. If we want set* $\Upsilon_a^{\mathcal{M}^{a,|\gamma|}}$ *to be referred to, plausible operator* $\mathbf{Pl}_a$ *must occur within formula* φ*.*

Example 3 Consider the scenario from Example 2. Suppose that it becomes implausible (according to c_1) that companies c_2 and c_3 will ever cooperate. Moreover, it is not likely that c_4 may impend another company (there is no reason for c_4 for such an action any more). Thus, the plausible paths (from c_1's perspective) can be now described by the path formula $\gamma_1 \equiv \Box\,(\neg\mathsf{coop}_{\mathsf{c_2,c_3}} \wedge \neg\mathsf{imp})$. Under this assumption: if c_4 is satisfied now, then it will be always either satisfied or have a way of becoming satisfied in the next moment. That is, formula $(\textbf{set-pl}_a\ \gamma_1)\mathbf{Pl}_{c_1}(\mathsf{sat_{c_4}} \rightarrow \mathsf{A}\Box\ (\mathsf{sat_{c_4}} \vee \mathsf{E}\bigcirc\mathsf{sat_{c_4}}))$ is true in the model from Fig. 2.

4.3 Plausible Paths: Discussion

So far, we did not assume anything about plausibility sets. Does every set of plausible paths make equal sense? Probably not. Here, we are going to suggest that there should be at least one plausible path starting in each state of the system. In fact, it is hard to imagine a situation with no outgoing plausible paths because it would mean that, if such a situation occurs, the agent will see *no plausible future at all*. Even when we consider a state which does not seem to plausibly happen from the perspective of our current state (that is, a state q' which is not reachable via a plausible path from the current state q): still, there should be a plausible path going out of q'. Though it seems now incredible that q' ever occurs, *if this does happen*, it should be accepted as a fact, and some outgoing paths should be seen as more plausible than the others. We formalize this restriction through the notion of *plausible seriality* of models.

A CTLKP model is *plausibly serial* (or *p-serial*) if every state of the system has an outgoing plausible path, i.e. $\partial(\Upsilon_a) = Q$. As we will see in Section 5, a weaker requirement is sometimes sufficient. We call a model *weakly p-serial* if every state has at least one indistinguishable counterpart from which a plausible path starts, i.e. for each $q \in Q$ there is a $q' \in Q$ such that $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$. Obviously, p-seriality implies weak p-seriality.

5 Investigating Plausibility, Knowledge, and Beliefs in CTLKP

In this section we study some relevant properties of plausibility, knowledge, and beliefs; in particular, axioms $K, D, T, 4, 5$ are examined.

5.1 Axiomatic Properties of Knowledge and Beliefs

We start with a slightly non-standard characterization of equivalence relations.

Lemma 5 *Relation $\sim$ is an equivalence relation (i.e., $\sim$ is transitive, reflexive, and symmetric) if and only if it is reflexive, symmetric, and euclidean. Moreover, an equivalence relation is serial.*

Proof. We show that equivalence relation $\sim$ is also euclidean. Let $x \sim y$ and $x \sim z$. Symmetry ($x \sim y \Rightarrow y \sim x$) and transitivity ($y \sim x \wedge x \sim z \Rightarrow y \sim z$) implies that $y \sim z$.

Now, we assume that $\sim$ is reflexive, symmetric, and euclidean. Let $x \sim y$ and $y \sim z$. This implies transitivity (i.e., $x \sim z$), because we have $y \sim x$ (symmetry) and $y \sim x \wedge y \wedge z \Rightarrow x \sim z$ (euclidity).

Seriality follows from reflexivity. ■

Now, the following result can be proved.

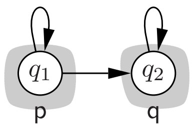

Figure 3: Model in which axiom T is not strongly valid.

Theorem 6 *Axioms $K, D, 4,$ and 5 for knowledge are strongly valid, and axiom T is valid. That is, modalities K_a form system $S5$ (in the sense of normal validity; and $KD45$ in the sense of strong validity).*

Proof. Let $\mathcal{M}$ be a CTLKP model, $P \subseteq \Lambda_{\mathcal{M}}$, and $q \in Q_{\mathcal{M}}$.

Axiom K: $\models K_a\varphi \wedge K_a(\varphi \to \psi) \to K_a\psi$**.** We have to show that $\mathcal{M} \models_P K_a\varphi \wedge K_a(\varphi \to \psi)$ implies $\mathcal{M} \models_P K_a\psi$. Assume $\mathcal{M}, q \models_P K_a\varphi \wedge K_a(\varphi \to \psi)$; it holds if and only if $\forall q' \in Q\big(q \sim_a q' \Rightarrow (q' \models \varphi \text{ and } q' \models \varphi \to \psi)\big)$ if $\forall q' \in Q\big(q \sim_a q' \Rightarrow q' \models \psi\big)$ iff $q \models_P K_a\psi$.

Axiom D: $\models K_a\varphi \to \neg K_a\neg\varphi$**.** Suppose $\mathcal{M}, q \models_P K_a\varphi$. We have to show that $\exists q' \in Q_{\mathcal{M}}(q \sim_a q' \wedge q' \models \varphi)$; this is true for $q' = q$, due to the assumption and reflexivity of $\sim_a$.

Axiom 4: $\models K_a\varphi \to K_aK_a\varphi$**.** Suppose $\mathcal{M}, q \models_P K_a\varphi$. We show that $\forall q' \in Q_{\mathcal{M}}\big(q \sim_a q' \Rightarrow \forall q'' \in Q_{\mathcal{M}}(q' \sim_a q'' \Rightarrow q'' \models \varphi)\big)$. Because of transitivity, we have $q \sim_a q''$, and because of the assumption, we obtain $q'' \models \varphi$.

Axiom 5: $\models \neg K_a\varphi \to K_a\neg K_a\varphi$**.** Suppose $\mathcal{M}, q \models \neg K_a\varphi$; that is, $\exists q' \in Q_{\mathcal{M}}(q \sim_a q' \wedge q' \not\models \varphi)$; let $q' = q^*$ be such a state. Then, we have $\forall q' \in Q_{\mathcal{M}}\big(q \sim_a q' \Rightarrow \exists q'' \in Q_{\mathcal{M}}(q' \sim_a q'' \wedge q'' \not\models \varphi)\big)$ because of euclidity ($q \sim_a q^*$ and $q \sim q'$ implies $q' \sim q^*$).

Axiom T: $\models K_a\varphi \to \varphi$**.** Suppose $\mathcal{M}, q \models K_a\varphi$; i.e., $\forall q' \in Q_{\mathcal{M}}(q \sim_a q' \Rightarrow q' \models \varphi)$. Because of reflexivity, we have $q \sim_a q$, and thus, $\mathcal{M}, q \models \varphi$.

A counterexample against *strong* validity of T is given in Example 4.

■

Example 4 Consider CTLKP model $\mathcal{M}_3$ shown in Figure 3, with the epistemic relation $\sim_a = \{(q_1, q_1), (q_2, q_2)\}$, and any set of plausible paths. Axiom T is not strongly valid if there is a $q \in Q_{\mathcal{M}}$ and a set of paths $P \subseteq \Lambda$ so that $q \not\models_P K_a\varphi \to \varphi$. That is, if $q \models_P K_a\varphi$ and $q \models_P \neg\varphi$ which is equivalent to $\forall q' \in Q_{\mathcal{M}}(q \sim_a q' \Rightarrow q' \models \varphi)$ and $q \models_P \neg\varphi$. From the reflexivity of $\sim_a$ it follows that $q \models \varphi$ and $q \models_P \neg\varphi$ must be satisfiable. Let $q = q_1$, $P = \{(q_1q_2q_2\ldots), (q_2q_2\ldots)\}$, and $\varphi \equiv \mathsf{E}\Box\, \mathsf{p}$. Then we have

$$q_1 \models \varphi \quad \text{and} \quad q_1 \models_P \neg\varphi,$$

so T is not strongly valid in $\mathcal{M}_3$.

A similar study for beliefs brings the following results.

Proposition 7 *Axioms K, 4, and 5 for beliefs are strongly valid. That is, we have:* $\models (B_a\varphi \wedge B_a(\varphi \to \psi)) \to B_a\psi$, $\models (B_a\varphi \to B_aB_a\varphi)$, *and* $\models (\neg B_a\varphi \to B_a\neg B_a\varphi)$.

Proof. Let $\mathcal{M}$ be a CTLKP model, $P \subseteq \Lambda$, and $q, q', q'' \in Q$.

Axiom K: $q \models_P B_a\varphi \wedge B_a(\varphi \to \psi)$ iff $\forall q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow q' \models_{\Upsilon_a} \varphi \wedge q' \models_{\Upsilon_a} (\varphi \to \psi)\big)$ iff $\forall q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow q' \models_{\Upsilon_a} \varphi \wedge q' \models_{\Upsilon_a} \psi\big)$, so $q \models_P B_a\psi$.

Axiom 4: Assume that $q \models_P B_a\varphi$; i.e., $\forall q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow q' \models_{\Upsilon_a} \varphi\big)$. We have to show that also $q \models_P B_aB_a\varphi$ which is the case if and only if $\forall q \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow \forall q'' \in Q\big(q' \sim_a q'' \wedge q'' \in \partial(\Upsilon_a) \Rightarrow q'' \models_{\Upsilon_a} \varphi\big)\big)$. This condition holds because if $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ and $q' \sim_a q''$ and $q'' \in \partial(\Upsilon_a)$ then also $q \sim_a q''$ (transitivity of $\sim_a$) and certainly still $q'' \in \partial(\Upsilon_a)$; therefore, $q'' \models_{\Upsilon_a} \varphi$ holds by the assumption.

Axiom 5: The proof is similar to the previous one. Assume that $q \models_P \neg B_a\varphi$. This is equivalent to $\exists q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \wedge q' \not\models_{\Upsilon_a} \varphi\big)$; let $q' = q^*$ be such a state (there is such a state because we assumed that $\neg B_a\varphi$ holds). We have to show that also $q \models_P B_a\neg B_a\varphi$; i.e., the condition $\forall q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow \exists q'' \in Q\big(q' \sim_a q'' \wedge q'' \in \partial(\Upsilon_a) \wedge q'' \not\models_{\Upsilon_a} \varphi\big)\big)$. We show that q^* is also such a required state for all q'. If $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ then also $q' \sim_a q^*$ (because $\sim_a$ is euclidean and we have $q \sim_a q^*$ and $q \sim_a q'$), and moreover, $q^* \in \partial(\Upsilon_a)$. By the assumption, it follows that $q^* \not\models_{\Upsilon_a} \varphi$ for $q'' = q^*$.

■

The next proposition concerns the "consistency" axiom D: $B_a\varphi \to \neg B_a\neg\varphi$. It is easy to see that the axiom is not valid in general: as we have no restrictions on plausibility sets Υ_a, it may be as well that $\Upsilon_a = \emptyset$. In that case we have $B_a\varphi \wedge B_a\neg\varphi$ for all formulae φ, because the set of states to be considered becomes empty. However, it turns out that D *is* valid for a very natural class of models.

Proposition 8 *Axiom D for beliefs is not valid in the class of all CTLKP models. However, it is strongly valid in the class of weak p-serial models (and therefore also in the class of p-serial models).*

Proof. Let $\mathcal{M}$ be a CTLKP model, $P \subseteq \Lambda$, and $q, q' \in Q$. First, let $\mathcal{M}$ be weakly p-serial. Axiom D is strongly valid if, for all states in which $B_a\varphi$ is true, $\neg B_a\neg\varphi$ is also true; hence, we assume that

$$q \models_P B_a\varphi; \qquad (*)$$

i.e., for all $q' \in Q$ with $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ it holds that $q' \models_{\Upsilon_a} \varphi$. Now, the conclusion must be shown:

$$\begin{aligned} q \models_P \neg B_a \neg\varphi \text{ iff } & \text{not for all } q' \in Q \text{ with } q \sim_a q' \\ & \text{and } q' \in \partial(\Upsilon_a) \text{ we have } q' \models_{\Upsilon_a} \neg\varphi \\ \text{iff } & \text{there is a } q' \in Q \text{ with } q \sim_a q' \\ & \text{and } q' \in \partial(\Upsilon_a) \text{ and } q' \models_{\Upsilon_a} \varphi. \end{aligned} \qquad (**)$$

Because the model is weakly p-serial, there is a state q' with $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$; hence, because of $(*)$, it holds that $q' \models_{\Upsilon_a} \varphi$ and therefore $(**)$.

Next, we show that the axiom is not valid in the class of all CTLKP models. Let Υ_a be empty. With this definition $(*)$ is true but $(**)$ is false. ■

Moreover, as one may expect, beliefs do not have to be always true.

Proposition 9 *Axiom T for beliefs is not valid; i.e., $\not\models (B_a\varphi \rightarrow \varphi)$. The axiom is not even valid in the class of p-serial models.*

Proof. For a counterexample consider Figure 3 once more. Let Υ_a be given as $\{(q_1 q_1 \dots)\}$, and a's epistemic relation as $\sim_a = \{(q_1, q_1), (q_2, q_2)\}$. Then we have $q_1 \models B_a \mathsf{A}\Box\, \mathsf{p}$ but not $q_1 \models \mathsf{A}\Box\, \mathsf{p}$ which contradicts the validity of axiom T. ■

All the above results are summarized in the next theorem.

Theorem 10 *Belief modalities B_a form system $K45$ in the class of all models, and $KD45$ in the class of weakly plausibly serial models (in the sense of both normal and strong validity). Axiom T is not even valid for p-serial models.*

An additional (but nevertheless interesting) property of B_a is that an agent believes that his beliefs are true:

Proposition 11 *Formula $B_a(B_a\varphi \rightarrow \varphi)$ is strongly valid.*

Proof. The formula holds iff $\forall q' \in Q\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow \big(\exists q'' \in Q(q' \sim_a q'' \wedge q'' \in \partial(\Upsilon_a) \wedge q'' \not\models_{\Upsilon_a} \varphi) \vee q' \models_{\Upsilon_a} \varphi\big)\big)$ holds. This is certainly the case because with $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ we also have $q' \sim_a q'$, and therefore, $q' \not\models_{\Upsilon_a} \varphi$ or $q' \models_{\Upsilon_a} \varphi$. ■

5.2 Interaction between Plausibility, Knowledge, and Beliefs

First, we investigate the relationship between knowledge and plausibility/physicality operators. Then, we look at the interaction between knowledge and beliefs, examining some axioms presented in [9].

Proposition 12 *Let φ be a CTLKP formula, and $\mathcal{M}$ be a CTLKP model. We have the following strong validities:*

(i) $\models \mathbf{Pl}_a K_a\varphi \leftrightarrow K_a\varphi$

(ii) $\models \mathbf{Ph}\, K_a\varphi \leftrightarrow K_a\mathbf{Ph}\,\varphi \leftrightarrow K_a\varphi$

Proof. Let $P \subseteq \Lambda_{\mathcal{M}}$.

(i) It follows from the fact that the definition of K_a "overwrites" the set of paths P with which the satisfaction relation is annotated.

(ii) We have that:

$$\begin{aligned}\mathcal{M}, q \models_P \mathbf{Ph}\, K_a\varphi \text{ iff } & \mathcal{M}, q \models K_a\varphi \\ \text{iff } & \text{for all } \mathcal{M}, q' \in Q \text{ with } q \sim_a q' \text{ it holds that } \mathcal{M}, q' \models \varphi \\ \text{iff } & \mathcal{M}, q \models_P K_a\varphi \\ \text{iff } & \text{for all } \mathcal{M}, q' \in Q \text{ with } q \sim_a q' \text{ it holds that } \mathcal{M}, q' \models_P \mathbf{Ph}\,\varphi \\ \text{iff } & \models_P K_a\mathbf{Ph}\,\varphi.\end{aligned}$$

■

We now want to examine the relationship between knowledge and beliefs. For instance, if agent a believes in something, he knows that he believes it. Or, if he knows a fact, he also believes that he knows it. On the other hand, for instance, an agent does not necessarily believe in all the things he knows.

Proposition 13 *The following formulae are strongly valid:*
(i) $B_a\varphi \to K_aB_a\varphi$ *(ii)* $K_aB_a\varphi \to B_a\varphi$ *(iii)* $K_a\varphi \to B_aK_a\varphi$
The following formulae are not *valid:*
(iv) $B_a\varphi \to B_aK_a\varphi$ *(v)* $K_a\varphi \to B_a\varphi$

Proof. Let $\mathcal{M} = (Q, R, \sim_1, \dots, \sim_k, \Upsilon_1, \dots \Upsilon_k, \pi)$, and $q, q', q'' \in Q$.

(a) (i) Assume that $B_a\varphi$ holds; we show that $K_aB_a\varphi$. The latter formula does not hold iff $\exists q'\big(q \sim_a q' \wedge \exists q''(q' \sim_a q'' \wedge q'' \in \partial(\Upsilon_a) \wedge q'' \not\models_{\Upsilon_a} \varphi)\big)$. But this condition is never fulfilled because if $q \sim_a q'$ and $q' \sim_a q''$ and $q'' \in \partial(\Upsilon)$ then also $q \sim_a q''$ and by assumption it follows $q'' \models_{\Upsilon_a} \varphi$.

(ii) Assume that $K_aB_a\varphi$ holds; i.e., formula $\forall q'\big(q \sim_a q' \Rightarrow \forall q''(q' \sim_a q'' \wedge q'' \in \partial(\Upsilon_a) \Rightarrow q'' \models_{\Upsilon_a} \varphi)\big)$. Hence, $B_a\varphi$ is true because otherwise there must exist a q^* with $(q \sim_a q^* \wedge q^* \in \partial(\Upsilon_a) \wedge q^* \not\models_{\Upsilon_a} \varphi)$. But this would yield a contradiction, since $\sim_a$ is reflexive (so the assumption would apply to $q'' = q^*$ and we would obtain $q' \models_{\Upsilon_a} \varphi$).

(iii) The same reason as for (i); if $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ and $q' \sim_a q''$ then also $q \sim_a q''$.

(b) (i) Assume that $B_a\varphi$ holds. Then, $\forall q'\big(q \sim_a q' \wedge q' \in \partial(\Upsilon_a) \Rightarrow \forall q''(q' \sim_a q'' \Rightarrow q'' \models \varphi)\big)$ does not hold for all states because q'' may not be a beginning of a plausible path so that the assumption does not secure $q'' \models_{\Upsilon_a} \varphi$ and especially not $q'' \models \varphi$.

(ii) See the counterexample in Example 5.

■

Example 5 Formula $K_a\varphi \rightarrow B_a\varphi$ has the meaning that everything that is known should also be believed. In our approach this is not the case. The axiom would hold if and only if the following statement would hold: $\forall q' \in Q$ with $q \sim_a q'$ it holds that $q' \models \varphi$ implies that $\forall q' \in Q$ with $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ we have $q' \models_{\Upsilon_a} \varphi$. Because generally $\Upsilon_a \subset \Lambda$, there could be a path in $\Lambda\backslash\Upsilon_a$ so that φ is fulfilled on that path. We will now specify φ and provide an example for our assumption.

Consider model $\mathcal{M}_3$ from Figure 3, and formula $\varphi \equiv \mathsf{E}\Box\ \mathsf{p}$ again. The agent knows that φ is true in q_1:

$$\mathcal{M}_3, q_1 \models K_a\varphi$$

because only $q_1 \sim_a q_1$, and $\mathcal{M}_3, q_1 \models_\Lambda \mathsf{E}\Box\ \mathsf{p}$ (note that p is true along path $q_1q_1q_1q_1\ldots$). Furthermore, q_1 is in $\partial(\Upsilon_a) = \{q_1, q_2\}$ but $\mathcal{M}_3, q_1 \not\models_{\Upsilon_a} \mathsf{E}\Box\ \mathsf{p}$ since p does not hold in q_2, and the only plausible q_1-path is $q_1q_2q_2\ldots$. Thus,

$$\mathcal{M}_3, q_1 \not\models_P B_a\varphi$$

which shows that $K_a\varphi \not\rightarrow B_a\varphi$.

The last invalidity is especially important: it is *not* the case that knowing something implies believing in it. For example, we may know that an invasion from another galaxy is in principle possible ($K_a\mathsf{E}\Diamond$ invasion), but if we do not take this possibility as plausible ($\neg\mathbf{Pl}_a\mathsf{E}\Diamond$ invasion), then we reject the corresponding belief in consequence ($\neg B_a\mathsf{E}\Diamond$ invasion). This emphasizes that we study a specific concept of beliefs here. Note that this specific is not due to the notion of plausibility itself; the reason lies rather in the fact that we investigate knowledge and beliefs *in a temporal framework*. This observation is formalized in the next proposition. After that, we show how the relationship between knowledge and beliefs can be characterized for the class of p-serial models.

Proposition 14 *Let φ be a CTLKP formula that does not include any temporal operators. Then $K_a\varphi \rightarrow B_a\varphi$ is strongly valid, and in the class of p-serial models we have even that $\models K_a\varphi \leftrightarrow B_a\varphi$.*

Proof. Assume that $\mathcal{M}, q \models K_a\varphi$ holds; i.e., for all q' with $q \sim_a q'$ we have that $\mathcal{M}, q' \models \varphi$. We show that $B_a\varphi$ also holds. First, let $q' \in \partial(\Upsilon_a)$; then, $q' \models_{\Upsilon_a} \varphi$ holds because no temporal operator occurs in φ (which makes the set of plausible paths irrelevant). On the other hand, if there is no state q' with $q \sim_a q'$ and $q' \in \partial(\Upsilon_a)$ then $B_a\varphi$ is trivially true.

In the class of p-serial models, we have $\partial(\Upsilon_a) = Q$, and therefore, the condition $q \in \partial(\Upsilon_a)$ is always true for all $q \in Q$. Furthermore, we have $q \models_{\Upsilon_a} \varphi$ if and only if $q \models \varphi$ because φ does not contain any temporal operator (and therefore no path quantifier). ■

Theorem 15 *The following formulae are strongly valid in the class of plausibly serial CTLKP models:*

(i) $K_a\mathbf{Pl}_a\varphi \leftrightarrow B_a\varphi$ *(ii)* $K_a\varphi \leftrightarrow B_a\mathbf{Ph}\,\varphi$

Proof.

(i)

$$\begin{aligned}\mathcal{M}, q \models_P B_a\varphi \text{ iff } & \forall q' \in Q(q \sim_a q' \wedge \underbrace{q' \in \partial(\Upsilon_a)}_{\top} \Rightarrow \mathcal{M}, q' \models_{\Upsilon_a} \varphi)\\ \text{iff } & \forall q' \in Q\big(q \sim_a q' \Rightarrow \mathcal{M}, q' \models_\Lambda \mathbf{Pl}_a\varphi\big)\\ \text{iff } & \mathcal{M}, q \models_P K\mathbf{Pl}_a\varphi.\end{aligned}$$

(ii)

$$\begin{aligned}\mathcal{M}, q \models_P K_a\varphi \text{ iff } & \forall q' \in Q(q \sim_a q' \Rightarrow \mathcal{M}, q' \models \varphi)\\ \text{iff } & \forall q' \in Q(q \sim_a q' \Rightarrow \mathcal{M}, q' \models_{\Upsilon_a} \mathbf{Ph}\,\varphi)\\ \text{iff } & \forall q' \in Q(q \sim_a q' \wedge \underbrace{q' \in \partial(\Upsilon_a)}_{\top} \Rightarrow \mathcal{M}, q' \models_{\Upsilon_a} \mathbf{Ph}\,\varphi)\\ \text{iff } & \mathcal{M}, q \models_P B_a\mathbf{Ph}\,\varphi\end{aligned}$$

■

Note that this characterization has a strong commonsense reading: *believing is knowing that φ plausibly holds*, and *knowing is believing that it holds physically*.

Finally, we observe an important feature of our plausibility operators. If a sequence of plausibility operators occurs in a formula, then only then only the last of them matters.

Proposition 16 $\models \mathbf{Pl}_i\mathbf{Pl}_j\varphi \leftrightarrow \mathbf{Pl}_j\varphi$ *for any agents i, j and formula φ.*

Proof. Let $\mathcal{M}$ be a CTLKP model, $P \subseteq \Lambda_\mathcal{M}$ and $\mathcal{M}, q \in Q_\mathcal{M}$. Then, we have: $\mathcal{M}, q \models_P \mathbf{Pl}_i\mathbf{Pl}_j\varphi$ iff $\mathcal{M}, q \models_{\Upsilon_i} \mathbf{Pl}_j\varphi$ iff $\mathcal{M}, q \models_{\Upsilon_j} \varphi$ iff $\mathcal{M}, q \models_P \mathbf{Pl}_j\varphi$. ■

Note also that the above feature does not extend to beliefs, i.e., $\not\models B_iB_j\varphi \leftrightarrow B_j\varphi$.

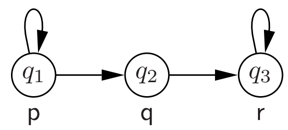

Figure 4: Model $\mathcal{M}_4$ with 3 states, and propositions p, q, and r

5.3 Properties of the Update

The plausibility update influences only formulae in which plausibility plays a role, i.e. ones in which belief or plausibility modalities occur.

Proposition 17 *Let φ be a CTLKP formula that does not include operators* $\mathbf{Pl}_a$ *and* B_a*, and γ be a CTLKP path formula. Then, we have* $\models \varphi \leftrightarrow (\textbf{set-pl}_a\ \gamma)\varphi$.

Proof. Let $\mathcal{M}$ be a CTLKP model, $q \in Q_{\mathcal{M}}$ and $P \subseteq \Lambda_{\mathcal{M}}$. Then, we have: $\mathcal{M}, q \models_P (\textbf{set-pl}_a\ \gamma)\varphi$ iff $\mathcal{M}^{a,|\gamma|}, q \models_P \varphi$. Because φ does not contain the B_a and $\mathbf{Pl}_a$ operator, the sets of plausible paths in the models are irrelevant; thus, we have $\mathcal{M}^{a,|\gamma|}, q \models_P \varphi$ iff $\mathcal{M}, q \models_P \varphi$. ■

What can be said about the result of an update? At first sight, formula $(\textbf{set-pl}_a\ \gamma)\mathbf{Pl}_a\mathsf{A}\gamma$ seems a natural characterization; unfortunately, it is not valid. In short, this is because, by leaving the other paths out of the scope, we may change properties of the paths that used to satisfy γ – in particular, they may cease to satisfy γ after that. The next example provides a more concrete argument.

Example 6 Consider model $\mathcal{M}_4$ from Figure 4. Let $\gamma \equiv \Box\, \mathsf{E}\bigcirc\mathsf{q}$. The set of paths described by γ is $\{\lambda \in \Lambda \mid \forall i \in \mathbb{N}_0 \big(\mathcal{M}, \lambda[i] \models \mathsf{E}\bigcirc\mathsf{q}\big)\} = \{(q_1 q_1 q_1 \dots)\}$. This set will become the set of plausible paths $\Upsilon_a^{\mathcal{M}'}$ in model $\mathcal{M}' = \mathcal{M}_4^{a,|\gamma|}$. Now, we can show that $\mathcal{M}_4, q_1 \not\models (\textbf{set-pl}_a\ \gamma)\mathbf{Pl}_a\mathsf{A}\gamma$:

$\mathcal{M}_4, q_1 \models (\textbf{set-pl}_a\ \gamma)\mathbf{Pl}_a\mathsf{A}\gamma$

iff $\mathcal{M}', q_1 \models_{\Upsilon_a^{\mathcal{M}'}} \mathsf{A}\gamma$ (where $\mathcal{M}' = \mathcal{M}_4^{a,\{(q_1 q_1 q_1 \dots)\}}$)

iff $\forall \lambda \in \Upsilon_a^{\mathcal{M}'}(q_1)$ it holds that $\mathcal{M}', \lambda[i] \models_{\Upsilon_a^{\mathcal{M}'}} \mathsf{E}\bigcirc\mathsf{q}$ for all $i \in \mathbb{N}_0$

iff $\mathcal{M}', q_1 \models_{\Upsilon_a^{\mathcal{M}'}} \mathsf{E}\bigcirc\mathsf{q}$

iff $\exists \lambda \in \Upsilon_a^{\mathcal{M}'}(q_1)$ with $\mathcal{M}', \lambda[1] \models_{\Upsilon_a^{\mathcal{M}'}} \mathsf{q}$

iff $\mathcal{M}', q_1 \models_{\Upsilon_a^{\mathcal{M}'}} \mathsf{q}$

Clearly, q does not hold in q_1 which proves that the formula is not valid.

We propose two alternative ways out: the first one restricts the language of the update similarly to [13]; the other refers to physical possibilities, in a way analogous to [8].

Proposition 18 *The CTLKP formula* $(\textbf{\textit{set-pl}}_a\ \gamma)\mathbf{Pl}_a\mathsf{A}\gamma$ *is not valid. However, we have the following strong validities:*

(i) $\models\!\!\equiv (\textbf{\textit{set-pl}}_a\ \gamma_u)\mathbf{Pl}_a\mathsf{A}\gamma_u$, *where* γ_u *is a* universal *path formula, defined as:*

$$\gamma_u ::= \bigcirc\varphi_u \mid \Box\,\varphi_u \mid \varphi_u\mathcal{U}\,\varphi_u,$$
$$\varphi_u ::= p \mid \neg p \mid \varphi_u \wedge \varphi_u \mid \varphi_u \vee \varphi_u \mid \mathsf{A}\gamma_u \mid K_a\varphi_u.$$

(ii) If φ *is an arbitrary CTL formula, then:*

$$\models\!\!\equiv (\textbf{\textit{set-pl}}_a\ \bigcirc\varphi)\mathbf{Pl}_a\mathsf{A}\bigcirc(\mathbf{Ph}\,\varphi),$$
$$\models\!\!\equiv (\textbf{\textit{set-pl}}_a\ \Box\,\varphi)\mathbf{Pl}_a\mathsf{A}\Box\,(\mathbf{Ph}\,\varphi), \text{ and}$$
$$\models\!\!\equiv (\textbf{\textit{set-pl}}_a\ \varphi_1\mathcal{U}\,\varphi_2)\mathbf{Pl}_a\mathsf{A}(\mathbf{Ph}\,\varphi_1)\mathcal{U}\,(\mathbf{Ph}\,\varphi_2).$$

Proof. Let $\mathcal{M}$ be a CTLKP model, $q \in Q_{\mathcal{M}}$ and $P \subseteq \Lambda$.

(i) We will provide a proof for $\gamma_u = \Box\,\varphi_u$; proofs for the other temporal operators are analogous. Let $\mathcal{M}' = \mathcal{M}^{a,|\gamma|}$. $\mathcal{M}, q \models_P (\textbf{set-pl}_a\ \gamma_u)\mathbf{Pl}_a\,\mathsf{A}\gamma_u$ holds if and only if we have

$$\forall\lambda \in \Lambda(q)\forall i \in \mathbb{N}_0\big(\mathcal{M}, \lambda[i] \models \varphi_u \Rightarrow \mathcal{M}', \lambda[i] \models \varphi\big).$$

The set of paths, with which the satisfaction relation is indexed, is only relevant if φ contains the universal quantifier A. Note also that $\Upsilon_a^{\mathcal{M}'} \subseteq \Lambda$. In consequence, if $\mathcal{M}, q \models_{\Upsilon^{\mathcal{M}'}} \varphi$, then also $\mathcal{M}, q \models_\Lambda \varphi$. Furthermore, the sets of plausible paths $\Upsilon_a^{\mathcal{M}}, \Upsilon_a^{\mathcal{M}'}$ inside models $\mathcal{M}, \mathcal{M}'$ are irrelevant because φ contains neither B_a nor $\mathbf{Pl}_a$.

(ii) We prove that $\mathcal{M}, q \models_P (\textbf{set-pl}_a\ \Box\,\varphi)\mathbf{Pl}_a\mathsf{A}\Box\ (\mathbf{Ph}\,\varphi)$; proofs for the other temporal operators are analogous. We have to show the following:

$$\forall\lambda \in \Lambda(q)\forall i \in \mathbb{N}_0\big(\mathcal{M}, \lambda[i] \models \varphi \Rightarrow \mathcal{M}^{|\gamma|}, \lambda[i] \models \varphi\big).$$

This statement is true because φ is just a CTL path formula; i.e., the set of plausible path in the model is irrelevant.

■

6 Conclusions and Further Work

In this paper a notion of *plausible behavior* is considered, with the underlying idea that *im*plausible options should be usually ignored in practical reasoning about possible future courses of action. In contrast to previous approaches [5, 12], we see plausibility as a temporal property. We add the new notion to the logic of CTLK [11], and obtain a language which enables reasoning about what can (or must) plausibly happen. As a technical device to define the semantics of the resulting logic, we use a non-standard satisfaction relation $\models_P$ that allows to propagate the "current" set of plausible paths into subformulae. Furthermore, we propose a non-standard notion of beliefs, defined in terms of indistinguishability and plausibility. We also propose how plausibility assumptions can be specified in the object language via a *plausibility update operator* (in a way similar to [8]).

Next, we use this new framework to investigate some important properties of plausibility, knowledge, beliefs, and updates. In particular, we show that knowledge is an $S5$ modality, and that beliefs satisfy axioms $K45$ in general, and $KD45$ for the class of *plausibly serial models*. Moreover, we prove that, for plausibly serial models, *believing that* φ is *knowing that* φ *plausibly holds*, and *knowing* φ is *believing that it holds physically*. That is, for these models, the relationship between knowledge and beliefs is very natural and reflects the initial intuition precisely.

In our opinion, this paper opens up several interesting directions for further work:

1. In our discourse on knowledge and plausibility, we only considered individual knowledge of agents. It can be interesting to consider collective knowledge as well (e.g., mutual, common and distributed knowledge). Plausibility can be treated in a similar way; i.e., we can think of "mutual", "common", and "distributed plausibility" too. Consequently, these concepts may be used to define collective beliefs in terms of collective knowledge and collective plausibility.

2. Instead of specifying sets of plausible paths by "vanilla" path formulae, one may think of a more general (yet still finite) representation. Note that there is no general solution to this problem, as CTL models usually include *uncountably* many paths.

3. Until now, we considered neither satisfiability checking nor model checking for our logic. This is another interesting topic for further research.

4. Alternating-time Temporal Logic ATL can be used (instead of CTL) as the basis for further studies on plausibility and beliefs. Some preliminary work on this topic has been already reported in [1]. In particular, we would like to describe and investigate various notions of *rationality* using this new framework.

5. Axiomatization of plausibility might also be studied in the future.

Finally, we would like to stress that we do not see this contribution as a mere technical exercise in formal logic. In our opinion, human agents use a similar concept of plausibility and "practical" beliefs in their everyday reasoning in order to reduce the search space and make the reasoning feasible. As a consequence, we suggest that the framework we propose may prove suitable for modeling, design, and analysis of resource-bounded agents in general.

References

[1] N. Bulling. Modal logics for games, time and beliefs. Master thesis, Clausthal University of Technology, 2006.

[2] E. A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 995–1072. Elsevier Science Publishers, 1990.

[3] E.A. Emerson and J.Y. Halpern. "sometimes" and "not never" revisited: On branching versus linear time temporal logic. In *Proceedings of the Annual ACM Symposium on Principles of Programming Languages*, pages 151–178, 1982.

[4] R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. MIT Press: Cambridge, MA, 1995.

[5] N. Friedman and J.Y. Halpern. A knowledge-based framework for belief change, Part I: Foundations. In *Proceedings of TARK*, pages 44–64, 1994.

[6] J.Y. Halpern. Reasoning about knowledge: a survey. In *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 4: Epistemic and Temporal Reasoning*, pages 1–34. Oxford University Press, Oxford, UK, 1995.

[7] J.Y. Halpern and R. Fagin. Modelling knowledge and action in distributed systems. *Distributed Computing*, 3(4):159–177, 1989.

[8] W. Jamroga, W. van der Hoek, and M. Wooldridge. Intentions and strategies in game-like scenarios. In Carlos Bento, Amílcar Cardoso, and Gaël Dias, editors, *Progress in Artificial Intelligence: Proceedings of EPIA 2005*, volume 3808 of *Lecture Notes in Artificial Intelligence*, pages 512–523. Springer Verlag, 2005.

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

[10] Y. Moses and Y. Shoham. Belief as defeasible knowledge. *Artificial Intelligence*, 64(2):299–321, 1993.

[11] W. Penczek and A. Lomuscio. Verifying epistemic properties of multi-agent systems via bounded model checking. In *Proceedings of AAMAS'03*, pages 209–216, New York, NY, USA, 2003. ACM Press.

[12] K. Su, A. Sattar, G. Governatori, and Q. Chen. A computationally grounded logic of knowledge, belief and certainty. In *Proceedings of AAMAS'05*, pages 149–156, New York, NY, USA, 2005. ACM Press.

[13] W. van der Hoek, M. Roberts, and M. Wooldridge. Social laws in alternating time: Effectiveness, feasibility and synthesis. *Synthese*, 2005.