Logic and Logical Philosophy (2015)
DOI: 10.12775/LLP.2015.014
Published online: August 18, 2015

Mattias Skipper Rasmussen

DYNAMIC EPISTEMIC LOGIC AND LOGICAL OMNISCIENCE

Abstract. Epistemic logics based on the possible worlds semantics suffer from the problem of logical omniscience, whereby agents are described as knowing all logical consequences of what they know, including all tautologies. This problem is doubly challenging: on the one hand, agents should be treated as logically non-omniscient, and on the other hand, as moderately logically competent. Many responses to logical omniscience fail to meet this double challenge because the concepts of knowledge and reasoning are not properly separated. In this paper, I present a dynamic logic of knowledge that models an agent's epistemic state as it evolves over the course of reasoning. I show that the logic does not sacrifice logical competence on the altar of logical non-omniscience.

Keywords: epistemic logic; dynamic epistemic logic; logical omniscience; resource-bounded reasoning

1. Introduction

The standard modal approach to epistemic logic, which dates back to Hintikka's *Knowledge and Belief* [24], models knowledge in terms of a possible worlds semantics. On this approach, an agent is said to know a proposition φ just in case φ is true in all possible worlds that are epistemically possible for the agent. Given this semantics, it follows that agents are characterized as *logically omniscient*: they know all logical consequences of what they know, including all classical tautologies.

Informally, the proof goes as follows: Suppose an agent knows a proposition φ and consider any proposition ψ that is logically entailed by φ. Since the agent knows φ, φ is true in all epistemically possible worlds

Received March 18, 2015. *Revised* August 9, 2015

for the agent. And since φ logically entails ψ, ψ is true in all possible worlds in which φ is true. Given that all epistemically possible worlds are *logically* possible, it follows by transitivity of set inclusion that ψ is true in all epistemically possible worlds for the agent. In turn, the agent knows ψ. So if an agent knows φ, she knows any logical consequence of φ. In other words, knowledge is described as closed under logical consequence.

If we aim to model *real-world* agents, as opposed to *ideal* agents, describing knowledge as closed under logical consequence is clearly inadmissible. For whereas the knowledge of unrealistically intelligent and powerful agents may well be closed under logical consequence, ordinary agents like human beings, computers, and robots generally fall short of logical omniscience. Standard modal epistemic logics thus suffer from what was dubbed *the problem of logical omniscience* by Hintikka [25].

Logical omniscience is a problem insofar as we aim to model the *explicit* knowledge of real-world agents — that is, the kind of knowledge that agents can act upon and answer questions about. Some theorists have adopted a distinction due to Levesque [32] between *explicit* and *implicit* knowledge, whereby an agent implicitly knows everything that follows logically from what she explicitly knows. By definition, then, logical omniscience is unproblematic in the case of implicit knowledge. And there may often be good reasons for examining the consequences of what an agent knows, even if the agent cannot make those consequences count in her practical deliberation. For instance, if we aim to model the information that is stored in an agent's epistemic state, we are looking for a theory of implicit knowledge. Or, as pointed out by Levesque [32], if we are interested in, not what an agent knows directly, but what the *world* would be like given the agent's knowledge, we are looking for a theory of implicit knowledge. However, my concern in this paper is with the explicit knowledge of real-world agents. As such, logical omniscience is a genuine problem in need of a solution.

Lack of logical omniscience may stem from various sources, the majority of which can be seen as special cases of resource-boundedness. Bounded agents like you and me simply do not have the time, memory, and computational power to infer all the — generally infinitely many — logical consequences of what they know. But even given unlimited resources, logical non-omniscience may arise from an incomplete reasoning mechanism on part of the agent. For instance, even an infinitely powerful chess computer may fail to decide whether white has a winning strategy, if it is deprived of one or more of the rules of chess. This is not to say that all real-world agents are rule-based. In particular, I leave it an open question to what extent human beings can be said to reason in accordance with inference rules. To circumvent this tricky empirical question, I will focus on resource-bounded agents who, by assumption, reason by applying inference rules.

To model such agents accurately, careful attention must be paid to the kind of logical non-omniscience that bounded resources give rise to. A claim that has received substantial support in the literature is that the knowledge of resource-bounded agents is not generally closed under any logical law, where a set Γ of propositions is closed under a logical law λ iff Γ contains every proposition that can be derived from Γ by any number of applications of λ.[1] The crucial point in favour of this claim is that for any logical law, it is not hard to imagine a resource-bounded agent whose knowledge is not closed under that law. For instance, it is not hard to imagine an agent who knows φ and ψ, but fails to know some complicated proposition $((\varphi \wedge \psi) \wedge \psi) \wedge (\psi \wedge ((\varphi \wedge \psi) \wedge (\psi \wedge \varphi)))$ that can be derived from φ and ψ by a large number of applications of conjunction introduction. So we cannot assume that the knowledge of resource-bounded agents is closed under conjunction. The same goes for other inference rules. Given this, I will take the following to be a desideratum for solving the problem of logical omniscience:

> NON-CLOSURE (NC): The knowledge of resource-bounded agents is not closed under any non-trivial logical law.

Notice that (NC) does not say that agents are *incapable* of deriving logical consequences of their knowledge. In fact, for all (NC) says, some resource-bounded agents may be highly logically competent. The modest claim is that we cannot expect a resource-bounded agent's knowledge to obey any closure principle (with the exception of *identity*, $\varphi \vdash \varphi$).

A solution to the problem of logical omniscience should satisfy (NC) in order to treat agents as logically non-omniscient in the right way. However, as has been pointed out by Chalmers [9], Bjerring [5, 6, 7], Duc [18], among others, we must not sacrifice logical competence on the altar of logical non-omniscience. For although we cannot expect real-world agents to *close* their knowledge under logical laws, we can still expect them to engage in bounded, but non-trivial, inferential reasoning. For

[1] Refer, amongst others, to Wansing [38], Duc [16, 18], Ågotnes [2], and Jago [26].

instance, I currently know that it is snowing and that it is cold outside whenever it snows. Do I know that it is cold outside? It seems so! Of course, there may be circumstantial reasons why I may fail to actually form the belief that it is cold outside. But at least I have the capacity to form the belief given my ability to reliably perform a single application of modus ponens. Similarly, artificially intelligent agents such as computers and robots obviously *do* perform non-trivial deductions, despite limited resources. For instance, a proof generator for classical propositional logic can easily verify that ψ follows from $\{\varphi, \varphi \to \psi\}$. So an epistemic logic for agents who are entirely logically inept is hardly of any interest for computer scientists or artificial intelligence researchers. Given this, a solution to the problem of logical omniscience should treat agents as logically competent or non-ignorant in the following minimal sense:

> Non-ignorance (NI): If a resource-bounded agent knows the premises of a valid inference and knows the relevant inference rule, then, given sufficient resources, the agent can infer the conclusion.

Over the past five decades, many attempts have been made to solve the problem of logical omniscience. However, most such attempts fail to jointly satisfy (NC) and (NI). The reason for this, as I will argue, is that many proposed solutions to the problem of logical omniscience remain within a static framework: they do not describe an agent's reasoning process, but only what the agent knows at the end point of a (more or less idealized) reasoning process. Static approaches to epistemic logic cause problems in the context of logical omniscience, because it is an agent's limited — but not absent — ability to reason logically that makes the agent logically non-omniscient and logically non-ignorant at the same time. So to properly solve the problem of logical omniscience, we cannot abstract away from the reasoning processes that underlie much belief (and knowledge) formation.

In this paper, I develop a dynamic logic of knowledge that models an agent's epistemic state as it evolves over the course of reasoning. I show that the logic jointly satisfies (NC) and (NI) — and so it properly solves the problem of logical omniscience.

I proceed as follows. In Section 2, I examine three prominent static responses to logical omniscience and show that each of them fails to jointly satisfy (NC) and (NI). This serves to motivate Section 3 in which I present a novel dynamic framework to deal with logical omniscience. I present a number of desirable results of the framework and argue that it jointly satisfies (NC) and (NI). I also compare the framework to other dynamic approaches to epistemic logic in the literature. In Section 4, I conclude and discuss future work.

2. Responses to Logical Omniscience

We have seen that logical omniscience arises in Hintikka-style logics from defining knowledge as "truth in all epistemically possible worlds". Given this, it would seem natural to try to avoid logical omniscience by modifying either the notion of *knowledge*, the notion of *truth*, or the notion of a *world*. These three general strategies underlie many proposed solutions to the problem of logical omniscience. In this section, I examine the three strategies (in reverse order) and argue that each of them fails to jointly satisfy (NC) and (NI). I focus on responses that retain the core model of knowledge as "truth in all epistemically possible worlds". As such, I will not discuss syntactic approaches to epistemic logic that take knowledge as primary, rather than analyse knowledge in terms of worlds. In such logics, agents represent their knowledge symbolically as sets of sentences. Later, in Section 3, I briefly discuss a number of dynamic logics of syntactic knowledge that have been proposed to deal with logical omniscience.

2.1. Impossible Worlds

A popular strategy to avoid logical omniscience, proposed in various forms by Cresswell [10, 11, 12], Hintikka [25], Rantala [34], Wansing [38], and Fagin *et al.* [22], is to augment the set of worlds that an agent can consider epistemically possible by a set of *impossible worlds*, where the rules of classical logic fail to hold. Formally, this strategy is implemented by replacing the classical, recursively defined valuation function with a syntactic valuation σ that assigns arbitrary truth-values to all sentences in all worlds. σ is required to behave classically in possible worlds, but may behave arbitrarily in impossible worlds. For example, $\varphi\wedge\psi$ need not be true in an impossible world that verifies φ and ψ, unless we demand that σ closes impossible worlds under conjunction. In turn, an agent can know φ and ψ, but fail to know $\varphi \wedge \psi$. So knowledge is no longer described as closed under classical entailment.

Although the inclusion of impossible worlds can help us avoid logical omniscience, impossible worlds approaches generally fail to jointly satisfy

(NC) and (NI). To see why this is so, proceed by cases. Assume, first, that σ is allowed to behave completely arbitrarily in impossible worlds. That is, impossible worlds may be arbitrarily logically ill-behaved. This means that agents will be modelled as altogether logically ignorant: they need never know *any* logical consequences of what they know. To see this, suppose that an agent knows φ and consider any logical consequence ψ of φ. Since the agent knows φ, φ is true in all epistemically possible worlds for the agent. And given that impossible worlds can be arbitrarily logically ill-behaved, there will be impossible worlds that verify φ but not ψ. In turn, there is nothing to ensure that ψ is true in all epistemically possible worlds just because φ is. So the agent need not know ψ. In turn, (NI) is violated. Assume, next, that σ is required to close impossible worlds under one or more logical laws. This means that an agent's knowledge will be described as closed under those laws. For example, if σ is required to close impossible worlds under conjunction, knowledge will be described as closed under conjunction. To see this, suppose that an agent knows φ and ψ. Given this, φ and ψ are true in all epistemically possible worlds for the agent. And since all worlds that verify φ and ψ also verify $\varphi \wedge \psi$, all epistemically possible worlds for the agent verify $\varphi \wedge \psi$. So the agent knows $\varphi \wedge \psi$. In turn, (NC) is violated. On the assumption that σ either behaves arbitrarily in impossible worlds or closes them under at least one inference rule, we conclude that the impossible worlds approach fails to jointly satisfy (NC) and (NI).

To avoid this dilemma, one may attempt to "partly" close impossible worlds under logical consequence. More specifically, we may aim to close impossible worlds under *easy* but not *full* logical consequence. For if every world that verifies a proposition φ also verifies at least the easy logical consequences of φ, we will have ensured that agents know at least the easy logical consequences of what they know. In this way, some measure of logical competence is retained. The trouble, as has been pointed out by Bjerring [5, 6, 7] and Jago [27, 28], is that no world can be closed under easy logical consequence without being closed under full logical consequence. To see this, consider a world w that is not closed under full logical consequence. This means that, for some proposition ψ that follows from premises $\varphi_1, \ldots, \varphi_n$, w verifies all of the premises $\varphi_1, \ldots, \varphi_n$, but not the conclusion ψ. Consider now any proof P from $\varphi_1, \ldots, \varphi_n$ to ψ. Since w verifies $\varphi_1, \ldots, \varphi_n$, but not ψ, w must violate at least *one* of the inferences in P. And on any relevant interpretation of what counts as an "easy inference", a single application of a simple inference rule will count as an easy inference. In turn, w is not closed under easy logical consequence. So if a world is not closed under full logical consequence, it is not closed under easy logical consequence. Equivalently, if a world is closed under easy logical consequence, it is closed under full logical consequence. So the strategy of partly closing impossible worlds under logical consequence does not help us resolve the dilemma between logical omniscience and logical ignorance.

2.2. Non-Classical Worlds

A different strategy to avoid logical omniscience is to change the notion of *truth* by replacing the semantics of classical logic with the semantics of a weaker non-classical logic. The hope is that the problem of logical omniscience can be alleviated somewhat by describing agents as omniscient with respect to a sufficiently weak non-classical logic instead of classical logic. Numerous non-classical logics have been investigated in the literature, including 4-valued logic, intuitionistic logic, and relevance logic.[2] I shall here consider an approach proposed by Levesque [32], Lakemeyer [30], and Fagin *et al.* [21, 22] — with only minor differences among them — that adopts a non-classical treatment of negation. The idea is to replace the classical truth-functional semantics of negation, by which $\neg\varphi$ is true iff φ is false, with a non-truth-functional semantics that takes $\neg\varphi$ and φ to have independent truth-values. One way of capturing this idea formally is to replace the classical truth assignment which assigns truth-values to all *atomic* sentences in all worlds with a non-classical truth assignment that assigns truth-values to all *literals* in all worlds (where a literal is an atomic sentence or its negation). The familiar satisfaction clauses of modal epistemic logic are modified accordingly to define separately what it means for a sentence and its negation to be true.[3] Fagin *et al.* [22, pp. 321-25] show that this non-classical treatment of negation yields a logic in which $\{\varphi \to \psi, \varphi\}$ no longer implies ψ, but in which, for example, $\{\varphi, \psi\}$ still implies $\varphi \wedge \psi$ and $\neg\neg\varphi$ still implies φ. In turn, knowledge is no longer described as closed under material implication, but is still described as closed under conjunction and double negation.

[2] See Priest [33] for an overview of non-classical logic. 4-valued logic was studied by Belnap [4], intuitionistic logic by Heyting [23], and relevance logic by Anderson and Belnap [1], and Routley and Meyer [35, 36, 37].

[3] I omit some of the formal details here. Refer to Fagin *et al.* [22, pp. 321–32] for a comprehensive exposition.

Although non-classical epistemic logics do not describe knowledge as closed under classical entailment, they nevertheless fail to jointly satisfy (NC) and (NI). To see why this is so, proceed again by cases. Assume, first, that the chosen non-classical logic is non-trivial (i.e. has at least *one* logical law). This means that an agent's knowledge will be described as closed under the non-trivial law(s) of that logic. In the example above, an agent's knowledge would be described as closed under e.g. conjunction and double negation. In turn, (NC) is violated. The relevant informal proof is structurally similar to the ones presented in the previous section. Assume, next, that the chosen non-classical logic is trivial (i.e. has *no* logical laws). In effect, worlds may then be arbitrarily logically ill-behaved. In turn, as I argued in the previous section, agents will be characterized as being incapable of performing even the most elementary logical derivations which means that (NI) is violated. So any non-classical epistemic logic fails to jointly satisfy (NC) and (NI).

2.3. Awareness

A third strategy to avoid logical omniscience, first proposed by Fagin and Halpern [20], is to redefine the concept of *knowledge* in the following way: An agent knows a proposition φ just in case φ is true in all possible worlds that are epistemically possible for the agent and the agent is *aware* of φ. The underlying intuition is that "it is necessary to be *aware* of a concept before one can have beliefs about it. One cannot know something of which one is unaware" [22, p. 337]. Formally, this idea is captured by introducing, for each agent i, a syntactic awareness function $\mathcal{A}_i$ that in each world w yields a set of sentences that agent i is aware of in w. Agent i's awareness set may be completely arbitrary, unless closure conditions are placed on $\mathcal{A}_i$. For example, agent i may be aware of $\varphi \wedge \psi$ in w without being aware of φ or ψ in w, unless a condition is placed on $\mathcal{A}_i$ to the effect that $\varphi, \psi \in \mathcal{A}_i(w)$ if $\varphi \wedge \psi \in \mathcal{A}_i(w)$. In turn, an agent may know $\varphi \wedge \psi$ without knowing φ or ψ which means that knowledge is no longer described as closed under classical entailment.

The awareness approach fails to jointly satisfy (NC) and (NI) for much the same reason that the impossible worlds approach falters in this respect. If awareness sets are allowed to behave completely arbitrarily, agents need never be capable of deriving *any* logical consequences of what they know. To see this, suppose that an agent knows φ, and consider any logical consequence ψ of φ. Since the agent knows φ, φ is true in all epistemically possible worlds for the agent and the agent is aware of φ. And since all epistemically possible worlds are *logically* possible, ψ is true in all epistemically possible worlds for the agent. However, given that awareness sets may behave completely arbitrarily, there is nothing to ensure that the agent is aware of ψ just because she is aware of φ. In turn, the agent need not know ψ just because she knows φ. So agents need never know any logical consequences of what they know which means that (NI) is violated. To avoid this result, we may specify certain closure properties on awareness sets. For instance, we may require that awareness sets are closed under subformulas: whenever an agent is aware of φ, she is aware of all subformulas of φ. As Fagin *et al.* [22, p. 341] show, this ensures that an agent's knowledge is closed under material implication: If an agent knows φ and $\varphi \to \psi$, the agent also knows ψ. In turn, some measure of logical competence is retained. However, this means that (NC) is violated instead. So the awareness approach violates either (NC) or (NI).

To avoid this dilemma, we may aim to close awareness sets under easy but not full logical consequence. For if an agent is aware of at least all the easy logical consequences of what she is aware of, we will have ensured that the agent's epistemic state is closed under easy logical consequence. However, just as we cannot close impossible worlds under easy logical consequence without closing them under full logical consequence, so we cannot close awareness sets under easy logical consequence without closing them under full logical consequence. To see this, we simply go through the same line of reasoning that we went through when discussing the impossible worlds response. So the strategy of partly closing awareness sets under logical consequence does not help us resolve the dilemma between logical omniscience and logical ignorance.

3. Dynamic Epistemic Logic

The underlying reason why the three examined responses to logical omniscience fail to jointly satisfy (NC) and (NI) is that they all remain within a static framework: they do not model an agent's reasoning process, but only what the agent knows at the end point of a (more or less idealized) reasoning process. As mentioned in the introduction, static approaches to epistemic logic cause problems in the context of logical omniscience, because it is an agent's ability to perform bounded, but non-trivial, logical reasoning that makes the agent logically non-omniscient and logically non-ignorant at the same time. It therefore seems natural to try to resolve the dilemma between logical omniscience and logical ignorance by modelling the relationship between knowledge and reasoning. This is what I aim to do in this section. I present a dynamic epistemic logic that models an agent's epistemic state as it evolves over the course of reasoning.

A number of dynamic theories of knowledge have already been proposed in the literature to deal with logical omniscience. However, these theories typically model *syntactic* knowledge: instead of modeling knowledge in a world-involving framework they take knowledge as *primary*. An early dynamic model of syntactic knowledge is Konolige's *Deduction Model of Belief* [29] which models an agent's epistemic state as a set of sentences that is closed under a set of deduction rules. Given that the set of deduction rules may be incomplete, the problem of logical omniscience can be alleviated to some extent. More recently, *active logics* (formerly known as *step logics*) have been developed to model the evolution of syntactic knowledge over the course of reasoning (see [19, 14, 15]). Here one step of reasoning is taken to correspond to a single application of an inference rule. Ågotnes have proposed a *Logic of Finite Syntactic Epistemic States* [2] that centers around two syntactic operators Δ_i and ∇_i that take sets of sentences as arguments: $\Delta_i\{\varphi_1, \ldots, \varphi_n\}$ says that agent i knows *at least* the formulae $\varphi_1, \ldots, \varphi_n$, whereas $\nabla_i\{\varphi_1, \ldots, \varphi_n\}$ says that agent i knows *at most* the formulae $\varphi_1, \ldots, \varphi_n$. The language also introduces expressions for knowing an inference rule — or, more generally, a *mechanism* — analogue to knowing a formula. An agent can then apply a known mechanism to obtain a new epistemic state if she chooses to. In particular, if an agent knows an inference rule and the relevant premises, then the agent may derive the conclusion if she chooses to.

Duc's *dynamic epistemic logic* [16, 17, 18] adds a dynamic operator $\langle \mathsf{F}_i \rangle$ to the standard epistemic language such that $\langle \mathsf{F}_i \rangle \varphi$ intuitively means that φ is true after some reasoning process performed by agent i. This allows him to model agents who, despite being logically non-omniscient, are capable of eventually teasing out any logical consequence of what they know, if only they think hard enough. In particular, such agents can derive any tautology of classical logic, since the rule "from φ infer $\langle \mathsf{F}_i \rangle\, \mathsf{K}_i\, \varphi$" is derivable in Duc's logic (see [17, p. 643]). Since the standard necessitation rule "from φ infer $\mathsf{K}_i\, \varphi$" is not derivable in Duc's logic, logical omniscience is avoided. Duc provides no semantics for his logic, but Alechina and Ågotnes [3] have developed a framework for modelling the dynamics of syntactic knowledge for which Duc's logic is sound and complete. Later, in this section, I will discuss the more technical details of Duc's theory, since his framework bears similarities to the one I present below.

To my knowledge, no one has yet presented a dynamic logic of knowledge that properly solves the problem of logical omniscience while retaining the core model of knowledge as "truth in all epistemically possible worlds". This is a problem, since the worlds-based account has proven to be a highly successful framework for modelling epistemic notions such as belief, knowledge, credence, and information. Indeed, this framework has been widely adopted not only by philosophers, but also by linguists, computer scientists, and artificial intelligence researchers. In light of this popularity, it would be desirable if we could solve the problem of logical omniscience within a world-involving framework. As such, my eventual goal is to develop a dynamic logic of knowledge defined both axiomatically and model theoretically, using a world-involving semantics. In this paper, I develop this logic from an axiomatic point of view. The corresponding model theory will be left for future work. However, in Section 4, I provide an outline of how I intend to proceed to develop the model theory.

To get started, consider the following mundane case of a reasoning process.

> The Mary Case: Mary is about to leave the house and ponders whether to put on a rain jacket. She can see that it is raining outside and knows that if it is raining then she should wear a rain jacket. Since Mary is an ordinary person with ordinary cognitive resources, she manages to apply modus ponens and conclude that she should put on a rain jacket before leaving the house.

Mary's reasoning process has at least two central features. The first feature is the *inference rule*, modus ponens, which Mary has to apply to derive the conclusion. Of course, more complex reasoning processes may involve several inference rules. In such cases, we need to pay attention not only to the involved inference rules, but also to the *chronology* of these inference rules. For instance, if an agent knows $\neg\neg\varphi$ and $\varphi \to \psi$, the agent will have to apply first the rule of double negation elimination and then modus ponens to derive ψ. The second feature of Mary's reasoning process is the *cognitive cost* of applying modus ponens. We can think of the cognitive cost of an inference rule as the time it takes for an agent to apply the inference rule. As is common, I shall take time to be discrete, which means that the cognitive cost of, say, modus ponens simply is a natural number $\mu \in \mathbb{N}$.

A complete description of Mary's reasoning process should pay equal attention to the two features above. Of course, we may not always be interested in a *complete* description of an agent's reasoning process. For instance, if we aim to model whether an agent can derive a conclusion within a given time span, we need only pay attention to the total cognitive cost of the agent's reasoning process and can abstract away from the specific inference rules involved in the reasoning process. Sometimes we may even simply be interested in whether an agent *can* or *cannot* eventually derive a given proposition, regardless of the cognitive cost of doing so. For the sake of generality, I will however develop a fully detailed logic that takes into account (i) the specific applications of inference rules involved in a reasoning process, (ii) the chronology of these applications of inference rules, and (iii) the cognitive cost of each application of an inference rule.

The remainder of this section is organized as follows. In Section 3.1, I define a dynamic epistemic language that augments the standard epistemic language over which Hintikka-style logics are defined. In Section 3.2, I provide an axiomatization of the dynamic epistemic logic that I have set out to develop and present a number of desirable results of this logic. I also compare the logic to Duc's dynamic epistemic logic. Finally, in Section 3.3, I argue that the presented logic jointly satisfies (NC) and (NI) — and so it properly solves the problem of logical omniscience.

3.1. The Dynamic Epistemic Language

The logic presented in what follows is defined over a propositional modal language $\mathcal{L}_{\mathrm{D}}(\Phi)$ ('D' for *dynamic*) defined as follows.

DEFINITION 1 (Language). The dynamic epistemic language $\mathcal{L}_{\mathrm{D}}(\Phi)$ is defined inductively from a set Φ of atomic sentences, an adequate set of truth-functional connectives $\Pi = \{\neg, \rightarrow\}$ (from which the connectives $\wedge$, $\vee$, and $\leftrightarrow$ are defined as usual), a knowledge operator K, and a set of dynamic operators $\langle \mathsf{R}_i \rangle^{\lambda_i}$ for $1 \leqslant i \leqslant n$:

- $\Phi \subseteq \mathcal{L}_{\mathrm{D}}(\Phi)$.
- If $\varphi \in \mathcal{L}_{\mathrm{D}}(\Phi)$ then $\neg\phi \in \mathcal{L}_{\mathrm{D}}(\Phi)$.
- If $\varphi, \psi \in \mathcal{L}_{\mathrm{D}}(\Phi)$ then $\varphi \rightarrow \psi \in \mathcal{L}_{\mathrm{D}}(\Phi)$.

- If $\varphi \in \mathcal{L}_{\mathrm{D}}(\Phi)$ then $\mathsf{K}\,\varphi \in \mathcal{L}_{\mathrm{D}}(\Phi)$.
- If $\varphi \in \mathcal{L}_{\mathrm{D}}(\Phi)$ then $\langle \mathsf{R}_i \rangle^{\lambda_i} \varphi \in \mathcal{L}_{\mathrm{D}}(\Phi)$, for $1 \leqslant i \leqslant n$.

The dual modality is defined in the usual way:

$$[\mathsf{R}_i]^{\lambda_i} \varphi := \neg \langle \mathsf{R}_i \rangle^{\lambda_i} \neg \varphi\,.$$

Intuitive readings of the various sentence types in $\mathcal{L}_{\mathrm{D}}(\Phi)$ follow here:

- $\mathsf{K}\,\varphi$: The agent knows φ.
- $\langle \mathsf{R}_i \rangle^{\lambda_i} \varphi$: After some application of R_i at cognitive cost λ_i, φ is the case.
- $[\mathsf{R}_i]^{\lambda_i} \varphi$: After any application of R_i at cognitive cost λ_i, φ is the case.

The following remarks will help us gain an intuitive understanding of the language $\mathcal{L}_{\mathrm{D}}(\Phi)$. First, we can think of $\{\mathsf{R}_1, \ldots, \mathsf{R}_n\}$ as a set of inference rules, where λ_i is the cognitive cost of R_i. For reasons of generality, I will leave the set of inference rules unspecified, but as a heuristic exercise we may keep in mind the inference rules of a natural deduction system such as Lemmon's system L [31]. In an epistemic context, the inference rules should be sound in order to secure veridicality of knowledge. In a doxastic context, this requirement may be dropped. Second, notice the difference between $\langle \mathsf{R}_i \rangle^{\lambda_i}$ and $[\mathsf{R}_i]^{\lambda_i}$: $\langle \mathsf{R}_i \rangle^{\lambda_i} \varphi$ says that φ is the case after *some* application of R_i at cognitive cost λ_i, whereas $[\mathsf{R}_i]^{\lambda_i} \varphi$ says that φ is the case after *any* application of R_i at cognitive cost λ_i. To illustrate this difference, $\langle \mathrm{MP} \rangle^{\mu}\, \mathsf{K}\,\psi$ is true of Mary, since there is some application of modus ponens after which Mary knows ψ—namely the instance that involves the premises 'it rains' and 'if it rains then Mary should put on a rain jacket'. However, given that Mary knows other sentences on which modus ponens can be applied, Mary does not know ψ after *any* application of modus ponens. So $[\mathrm{MP}]^{\mu}\, \mathsf{K}\,\psi$ is not true of Mary. Having defined $\mathcal{L}_{\mathrm{D}}(\Phi)$, we now turn to axiomatizing our dynamic epistemic logic.

3.2. Axiomatization

In axiomatizing our dynamic epistemic logic $\mathbf{L}_{\mathrm{D}}$, it will be helpful to have a notationally convenient way of representing arbitrary sequences of dynamic operators:

$$\langle \ddagger \rangle^i := \langle \mathsf{R}_i \rangle^{\lambda_i} \ldots \langle \mathsf{R}_j \rangle^{\lambda_j},$$

$$[\ddagger]^i := [\mathsf{R}_i]^{\lambda_i} \ldots [\mathsf{R}_j]^{\lambda_j},$$

where $\mathsf{R}_i, \ldots, \mathsf{R}_j$ are arbitrary inference rules, and $i = \lambda_i + \cdots + \lambda_j$. $\langle \ddagger \rangle^i \varphi$ says that "after some application of R_i at cognitive cost λ_i followed by ... followed by some application of R_j at cognitive cost λ_j, φ is the case". The intuitive reading of $[\ddagger]^i$ is obtained by replacing 'some' by 'any'. Intuitively, a sequence of dynamic operators represents a reasoning process consisting of a sequence of applications of inference rules. Such a sequence may be arbitrarily long or it may be empty. Also, it may contain the same inference rule multiple times.

Given these notational preliminaries, we are ready to axiomatize $\mathbf{L}_{\mathrm{D}}$.

Definition 2 (Axiomatization of $\mathbf{L}_{\mathrm{D}}$). Let $\varphi, \psi \in \mathcal{L}_{\mathrm{D}}(\Phi)$, let $\Gamma \subseteq \mathcal{L}_{\mathrm{D}}(\Phi)$, and let $\langle \ddagger \rangle^i, \langle \dagger \rangle^j$ (and $[\ddagger]^i, [\dagger]^j$) denote arbitrary sequences of dynamic operators. The logic $\mathbf{L}_{\mathrm{D}}$ has the following axiom schemata:

PC	All substitution instances of propositional tautologies,	
(A1)	$\langle \ddagger \rangle^i \, \mathsf{K}\, \varphi \rightarrow \varphi$	(Veridicality)
(A2)	$\langle \ddagger \rangle^i \, \mathsf{K}\, \varphi \rightarrow \langle \ddagger \rangle^i [\dagger]^j \, \mathsf{K}\, \varphi$	(Persistence)
(A3)	$\langle \ddagger \rangle^i \varphi \wedge \langle \dagger \rangle^j \psi \rightarrow \langle \ddagger \rangle^i \langle \dagger \rangle^j (\varphi \wedge \psi)$	(Succession)
(A4)	$\langle \ddagger \rangle^i (\varphi \wedge \psi) \rightarrow \langle \ddagger \rangle^i \varphi$	(Elimination)

$\mathbf{L}_{\mathrm{D}}$ has the following inference rule:

MP If $\Gamma \vdash \varphi$ and $\Gamma \vdash \varphi \rightarrow \psi$ then $\Gamma \vdash \psi$.

The axiom **PC** together with **MP** axiomatize classical propositional logic. The axiom (A1) is a dynamic version of the well-known veridicality axiom **T** of standard epistemic logic, whereby knowledge entails truth. (A1) ensures that only true sentences can be derived. Notice that since $\mathsf{K}\, \varphi \rightarrow \varphi$ is a substitution instance of (A1), knowledge entails truth in the logic $\mathbf{L}_{\mathrm{D}}$. In a doxastic context, (A1) may be dropped, since belief is not veridical. (A2) says that known sentences remain known over the course of reasoning. Such a persistence axiom is needed to ensure that known premises remain available to the agent during an inference. Without some sort of persistence axiom, an agent need never successfully perform an inference. However, (A2) holds only under two assumptions. First, the agent's memory has to be infallible: the agent cannot forget what she knows, since this would mean that knowledge does not persist. Second, the agent's environment has to be static: the truth-values of objective sentences cannot change over the course of reasoning, since knowledge cannot be assumed to persist in a changing environment given that knowledge is veridical. Although both assumptions are implausible in many situations, I shall assume them for current purposes. (A3) says that reasoning processes can succeed each other in the expected way. For instance, if $\mathsf{K}\,\varphi$ is the case after some application of modus ponens and $\mathsf{K}\,\psi$ is the case after some application of modus tollens, then $\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi$ is the case after some application of modus ponens followed by some application of modus tollens. Notice that (A3) does not imply that knowledge is closed under conjunction, since it does not allow us to infer $\mathsf{K}(\varphi \wedge \psi)$ from $\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi$. That is, $\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi \rightarrow \mathsf{K}(\varphi \wedge \psi)$ is not a substitution instance of (A3). This is important, since closure under conjunction would violate (NC). (A4) says that if both φ and ψ are the case after a reasoning process then, in particular, φ is the case. Notice, again, that (A4) does not imply that knowledge is closed under conjunction elimination, since it does not allow us to infer $\mathsf{K}\,\varphi$ from $\mathsf{K}(\varphi \wedge \psi)$. That is, $\mathsf{K}(\varphi \wedge \psi) \rightarrow \mathsf{K}\,\varphi$ is not a substitution instance of (A4). Again, this is important, since closure under conjunction elimination would violate (NC).

There are axioms aside from (A1)–(A4) that one could take to characterize the concepts of knowledge and reasoning. For example, one could introduce dynamic versions of the introspection axioms **4** and **5** from standard epistemic logic. However, for the narrow purposes of defining a logic that jointly satisfies (NC) and (NI), the included axioms will suffice.

The logic $\mathbf{L}_{\mathrm{D}}$ describes agents with no inference rules available to them. In order to equip agents with inference rules, $\mathbf{L}_{\mathrm{D}}$ is extended with appropriate axioms — one axiom per inference rule. If Λ is a set of inference rules, let $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ denote the logic that extends $\mathbf{L}_{\mathrm{D}}$ with axioms that equip agents with the rules in Λ. To illustrate, I will here provide axioms that equip agents with modus ponens (**MP**), conjunction introduction (**CI**), and double negation elimination (**DN**). In stating these axioms, it will be helpful to have a notationally convenient way of representing arbitrary conjunctions of sentences in $\mathcal{L}_{\mathrm{D}}(\Phi)$: $\Delta := \varphi \wedge \cdots \wedge \psi$. The same sentence is allowed to appear multiple times in Δ. Given this notation, we can axiomatize the logic $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ as follows.

Definition 3 (Axiomatization of $\mathbf{L}_{\mathrm{D}}^{\Lambda}$). Let $\Lambda = \{\mathbf{MP}, \mathbf{CI}, \mathbf{DN}\}$, and let Δ be an arbitrary conjunction of sentences in $\mathcal{L}_{\mathrm{D}}(\Phi)$. Furthermore, let μ, κ, and ν denote the cognitive costs of **MP**, **CI**, and **DN**, respectively. $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ extends $\mathbf{L}_{\mathrm{D}}$ with the following axiom schemata:

$$(\mathrm{MP}_{\mathrm{D}}) \quad \langle\ddagger\rangle^{i}(\Delta \wedge \mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \rightarrow \psi)) \rightarrow \langle\ddagger\rangle^{i}\langle \mathrm{MP}\rangle^{\mu}(\Delta \wedge \mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \rightarrow \psi) \wedge \mathsf{K}\,\psi) \qquad (\text{MP-success})$$

$$(\mathrm{CI_D}) \qquad \langle\ddagger\rangle^i(\Delta \wedge \mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi) \rightarrow \langle\ddagger\rangle^i\langle\mathrm{CI}\rangle^\kappa(\Delta \wedge \mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi \wedge \mathsf{K}(\varphi \wedge \psi)) \quad \text{(CI-success)}$$

$$(\mathrm{DN_D}) \qquad \langle\ddagger\rangle^i(\Delta \wedge \mathsf{K}\,\neg\neg\varphi) \rightarrow \langle\ddagger\rangle^i\langle\mathrm{DN}\rangle^\nu(\Delta \wedge \mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}\,\varphi) \quad \text{(DN-success)}$$

The axiom ($\mathrm{MP_D}$) says, roughly, that if an agent knows φ and $\varphi \rightarrow \psi$, then the agent can derive ψ by applying modus ponens at cognitive cost μ. Less roughly, ($\mathrm{MP_D}$) says that if an agent knows φ and $\varphi \rightarrow \psi$ after a reasoning process $\langle\ddagger\rangle^i$ then, regardless of what else is the case after that reasoning process (denoted by Δ), the agent knows ψ, in addition to φ and $\varphi \rightarrow \psi$, after having extended $\langle\ddagger\rangle^i$ with some application of modus ponens at cognitive cost μ. Similar readings apply to ($\mathrm{CI_D}$) and ($\mathrm{DN_D}$). I would like to stress that the inference rules **MP**, **CI**, and **DN** are chosen merely for illustrative purposes. Axioms corresponding to other inference rules can be formulated in accordance with the following general scheme, where R is an inference rule with premises $\varphi_1, \ldots, \varphi_n$, a conclusion ψ, and a cognitive cost λ:

$$(\mathrm{R_D}) \qquad \langle\ddagger\rangle^i(\Delta \wedge \mathsf{K}\,\varphi_1 \wedge \cdots \wedge \mathsf{K}\,\varphi_n) \rightarrow \langle\ddagger\rangle^i\langle R\rangle^\lambda(\Delta \wedge \mathsf{K}\,\varphi_1 \wedge \cdots \wedge \mathsf{K}\,\varphi_n \wedge \mathsf{K}\,\psi)$$

Notice that the deduction theorem holds for $\mathbf{L}_\mathrm{D}^\Lambda$ (the proof is standard and can be found in [8, p. 203]).

Theorem 1 (Deduction theorem). *If $\Gamma \cup \{\varphi\} \vdash \psi$ then $\Gamma \vdash \varphi \rightarrow \psi$.*

The deduction theorem allows us to prove from assumptions, which significantly simplifies the construction of proofs. To show the logic $\mathbf{L}_\mathrm{D}^\Lambda$ at work, consider the following theorem:

Theorem 2. *The following formula is a theorem of $\mathbf{L}_\mathrm{D}^\Lambda$:*

$$\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}(\varphi \rightarrow \psi) \rightarrow \langle\mathrm{DN}\rangle^\nu\langle\mathrm{MP}\rangle^\mu\langle\mathrm{CI}\rangle^\kappa\,\mathsf{K}(\varphi \wedge \psi)$$

Proof. I shall refer to the deduction theorem as (DT). Aside from (DT), the proof involves conjunction introduction (**CI**) and conjunction elimination (**CE**).

1.	$\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}(\varphi \rightarrow \psi)$	~~Ass.~~[19]
2.	$\mathsf{K}\,\neg\neg\varphi$	1, **CE**
3.	$\mathsf{K}(\varphi \rightarrow \psi)$	1, **CE**
4.	$\mathsf{K}\,\neg\neg\varphi \rightarrow \langle\mathrm{DN}\rangle^\nu(\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}\,\varphi)$	($\mathrm{DN_D}$)
5.	$\langle\mathrm{DN}\rangle^\nu(\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}\,\varphi)$	2, 4, **MP**
6.	$\langle\mathrm{DN}\rangle^\nu(\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}\,\varphi) \rightarrow \langle\mathrm{DN}\rangle^\nu\,\mathsf{K}\,\varphi$	(A4)
7.	$\langle\mathrm{DN}\rangle^\nu\,\mathsf{K}\,\varphi$	5, 6, **MP**

8. $\langle\mathrm{DN}\rangle^{\nu}\,\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi)$ 3, 7, **CI**
9. $\langle\mathrm{DN}\rangle^{\nu}\,\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi) \to \langle\mathrm{DN}\rangle^{\nu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi))$ (A3)
10. $\langle\mathrm{DN}\rangle^{\nu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi))$ 8, 9, **MP**
11. $\langle\mathrm{DN}\rangle^{\nu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi)) \to \langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi) \wedge \mathsf{K}\,\psi)$ (MP_{D})
12. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi) \wedge \mathsf{K}\,\psi)$ 10,11,**MP**
13. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \to \psi) \wedge \mathsf{K}\,\psi) \to \langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi)$ (A4)
14. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi)$ 12, 13, **MP**
15. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi) \to \langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi \wedge \mathsf{K}(\varphi \wedge \psi))$ (CI_{D})
16. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi \wedge \mathsf{K}(\varphi \wedge \psi))$ 14, 15, **MP**
17. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}(\mathsf{K}\,\varphi \wedge \mathsf{K}\,\psi \wedge \mathsf{K}(\varphi \wedge \psi))$ $\to \langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}\,\mathsf{K}(\varphi \wedge \psi)$ (A4)
18. $\langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}\,\mathsf{K}(\varphi \wedge \psi)$ 16, 17, **MP**
19. $\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}(\varphi \to \psi) \to \langle\mathrm{DN}\rangle^{\nu}\langle\mathrm{MP}\rangle^{\mu}\langle\mathrm{CI}\rangle^{\kappa}\,\mathsf{K}(\varphi \wedge \psi)$ 1, 18, **DT** □

The formula from Theorem 2 says that if an agent knows $\neg\neg\varphi$ and $\varphi \to \psi$, then after some application of double negation elimination at cognitive cost ν followed by some application of modus ponens at cognitive cost μ followed by some application of conjunction introduction at cognitive cost κ, the agent knows $\varphi \wedge \psi$. Theorem 2 is a desirable result which serves to illustrate how $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ models the dynamics of epistemic states on a high level of detail: attention is paid both to the involved applications of inference rules, to the chronology of these applications of inference rules, and to the cognitive cost of each application of an inference rule. This high level of detail distinguishes the logic $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ from Duc's previously mentioned dynamic epistemic logic which is otherwise similar to $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ in several respects. Recall that Duc's logic centers around the dynamic operator $\langle\mathsf{F}_i\rangle$ which he gives the following reading: $\langle\mathsf{F}_i\rangle\varphi$ means that φ is the case after some reasoning process performed by agent i. This allows him to model whether an agent *can* or *cannot* eventually derive a given proposition. For instance, since $\langle\mathsf{F}_i\rangle\,\mathsf{K}_i\,\varphi$ is a theorem of Duc's logic for any classical tautology φ, agents are described as being capable of deriving any tautology, regardless of the complexity of the deduction. [17, p. 643] However, Duc's logic abstracts away from the particular inference rules involved in a reasoning process as well as the cognitive cost of applying these inference rules. This abstraction has the convenient consequence that the operator $\langle\mathsf{F}_i\rangle$ behaves much like a future operator of standard tense logic of transitive time. In fact, $\langle\mathsf{F}_i\rangle\varphi$

can be read as "φ is true at some future time", whereas $[\mathsf{F}_i]\varphi$ can be read as "φ is true at any future time". This allows Duc to utilize well-known axioms from tense logic in the context of epistemic logic. However, given that we are interested in a detailed model of the evolution of epistemic states over the course of reasoning, the abstractions made in Duc's logic seem inadmissible. In particular, if we want to capture the fact that not all deductions are equally hard or easy for bounded agents to perform, Duc's logic is too abstract. By contrast, the logic $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ captures this fact quite naturally by keeping track of the applied inference rules as well as their corresponding cognitive costs.

Furthermore, if we are interested in reasoning about the epistemic lives of agents with an incomplete reasoning mechanism — such as the one's described by Konolige's previously mentioned *deduction model of belief* [29] — Duc's logic seems inadequate. For, as far as I can tell, there is nothing in Duc's formalism that allows us to model agents who can apply some inference rules, but not others. The reason for this is that the way in which Duc ensures that agents are characterized as logically competent is by including the following axioms:

- $\mathsf{K}_i\,\varphi \wedge \mathsf{K}_i(\varphi \rightarrow \psi) \rightarrow \langle \mathsf{F}_i \rangle\, \mathsf{K}_i\,\psi$
- $\langle \mathsf{F}_i \rangle\, \mathsf{K}_i(\varphi \rightarrow (\psi \rightarrow \varphi))$
- $\langle \mathsf{F}_i \rangle\, \mathsf{K}_i((\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \chi)))$
- $\langle \mathsf{F}_i \rangle\, \mathsf{K}_i((\neg\psi \rightarrow \neg\varphi) \rightarrow (\varphi \rightarrow \psi))$

Together, these four axioms state that agents are capable of applying modus ponens as well as the three standard Hilbert axioms of classical propositional logic. So, on Duc's model, agents do not reason by applying inference rules such as the ones known from natural deduction calculi. Instead, they reason using a Hilbert-style axiomatic proof system. As such, Duc's logic does not allow us to equip agents with a (possibly incomplete) set of inference rules. By contrast, the logic $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ provides us with exactly this kind of flexibility.

To further illustrate the kinds of formulae that can be proven in $\mathbf{L}_{\mathrm{D}}^{\Lambda}$, here follows a sample of theorems of $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ stated without proof.

Theorem 3. *The following formulae are theorems of* $\mathbf{L}_{\mathrm{D}}^{\Lambda}$:

- $\langle \mathrm{DN} \rangle^{\nu}\, \mathsf{K}\,\varphi \wedge \mathsf{K}(\varphi \rightarrow \psi) \rightarrow \langle \mathrm{DN} \rangle^{\nu} \langle \mathrm{MP} \rangle^{\mu}\, \mathsf{K}\,\psi$
- $\mathsf{K}\,\neg\neg\varphi \wedge \mathsf{K}\,\psi \rightarrow \langle \mathrm{DN} \rangle^{\nu} \langle \mathrm{CI} \rangle^{\kappa}\, \mathsf{K}(\varphi \wedge \psi)$
- $\mathsf{K}(\neg\neg(\varphi \wedge \psi)) \wedge \mathsf{K}((\varphi \wedge \psi) \rightarrow \chi) \rightarrow \langle \mathrm{DN} \rangle^{\nu} \langle \mathrm{MP} \rangle^{\mu}\, \mathsf{K}\,\chi$
- $\mathsf{K}\,\neg\varphi \wedge \mathsf{K}(\neg\varphi \rightarrow \psi) \wedge \mathsf{K}(\psi \rightarrow \neg\neg\chi) \rightarrow \langle \mathrm{MP} \rangle^{\mu} \langle \mathrm{MP} \rangle^{\mu} \langle \mathrm{DN} \rangle^{\nu}\, \mathsf{K}\,\chi$

Which formulae can be proven in $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ obviously depends on the specific content of Λ. The more inference rules agents are equipped with, the stronger the resulting logic. In particular, if Λ is a complete deduction system, agents are characterized as fully capable of teasing out logical consequences of what they know.

3.3. Non-closure and Non-ignorance Revisited

In the previous section, I presented a dynamic epistemic logic $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ that allows us to model an agent's knowledge as it evolves over the course of reasoning. It remains to show that this logic properly solves the problem of logical omniscience. To show this, it suffices to show that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ jointly satisfies (NC) and (NI), regardless of the content of Λ. Proceed by cases on Λ (empty or not).

Case 1: $\Lambda = \emptyset$

$\mathbf{L}_{\mathrm{D}}^{\emptyset}$ trivially satisfies (NC), since $\mathbf{L}_{\mathrm{D}}^{\emptyset}$ describes agents as being incapable of applying any inference rules. This means that knowledge is not described as closed under any logical law. $\mathbf{L}_{\mathrm{D}}^{\emptyset}$ also trivially satisfies (NI), since (NI) places no requirement on an agent's ability to perform inferences, if the agent does not know any inference rules.

Case 2: $\Lambda \neq \emptyset$

To see that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ satisfies (NC), consider an arbitrary inference rule R with premises $\varphi_1, \ldots, \varphi_n$ and conclusion ψ. We must show that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ does not prove the following theorem:

R-closure: $\mathsf{K}\,\varphi_1 \wedge \cdots \wedge \mathsf{K}\,\varphi_n \rightarrow \mathsf{K}\,\psi$

The reason why we cannot allow $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ to prove R-closure is that R-closure describes an agent's knowledge as closed under the logical law R which violates (NC). To see that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ does indeed *not* prove R-closure, consider the two cases in which $R \notin \Lambda$ and $R \in \Lambda$. If $R \notin \Lambda$, $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ clearly does not prove R-closure, since agents will be described as being incapable of applying R. If $R \in \Lambda$, $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ will have an axiom that equips agents with R. This axiom will, as we have seen, take following form:

$$(\mathrm{R_D})\quad \langle \ddagger \rangle^i (\Delta \wedge \mathsf{K}\,\varphi_1 \wedge \cdots \wedge \mathsf{K}\,\varphi_n) \rightarrow \langle \ddagger \rangle^i \langle R \rangle^{\lambda} (\Delta \wedge \mathsf{K}\,\varphi_1 \wedge \cdots \wedge \mathsf{K}\,\varphi_n \wedge \mathsf{K}\,\psi)$$

($\mathrm{R_D}$) does not take the form of R-closure, since ($\mathrm{R_D}$) requires an agent to *reason* in order to come to know ψ. By contrast, R-closure says that an agent who knows $\varphi_1, \ldots, \varphi_n$ *automatically* knows ψ. In other words, (R_{D}) — as opposed to R-closure — treats reasoning as a necessary condition for *a priori* knowledge expansion. We conclude that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ satisfies (NC).

To show that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ satisfies (NI), we must show that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ describes agents as being capable of applying the rules in Λ as prescribed by (NI). Consider therefore an arbitrary inference rule R such that $R \in \Lambda$. Since $R \in \Lambda$, $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ will have the axiom (R_{D}). This axiom says that if an agent knows the premises of an instance of R then, given sufficient resources, the agent can derive the relevant conclusion. And since this is exactly what (NI) prescribes, we conclude that $\mathbf{R}_{\mathrm{D}}$ satisfies (NI).

This completes the proof that $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ jointly satisfies (NC) and (NI) regardless of the specification of Λ. In other words, $\mathbf{L}_{\mathrm{D}}^{\Lambda}$ describes agents as being logically non-omniscient and logically non-ignorant in the right way. As such, $\mathbf{L}_{\mathrm{D}}$ is on the table as a novel solution to the problem of logical omniscience.

4. Conclusions and Future Work

I have argued that a proper solution to the problem of logical omniscience should treat agents as logically non-omniscient and logically non-ignorant at the same time. I also argued that many of the most popular responses to logical omniscience falter in this respect, because the concepts of knowledge and reasoning are not formally separated. I went on to axiomatize a dynamic epistemic logic that models an agent's epistemic state as it evolves over the course of reasoning. As I showed, this logic treats agents as logically non-omniscient and logically non-ignorant in the right way. As such, the presented dynamic framework is on the table as a novel solution to the problem of logical omniscience.

For future work, I plan to develop a model theory for which the logic presented in this paper is sound and complete. While such a model theory cannot be based on possible worlds on pain of logical omniscience, it is my hope to retain the core of the impossible worlds framework and model knowledge as "truth in all epistemically possible worlds". In light of the arguments presented in Section 2, we must allow impossible worlds to be arbitrarily logically ill-behaved in order to avoid all traits of logical omniscience. To ensure that agents are still characterized as logically competent, we then need to provide an interesting semantics for sentences of the form $\langle \mathsf{R} \rangle^{\lambda} \varphi$ (and derivatively, $[\mathsf{R}]^{\lambda} \varphi$). While a full exposition of such a semantics must wait for future work, let me here outline how I intend to proceed.

The key starting-point is that reasoning issues a change in an agent's epistemic state. And since an epistemic model (consisting of a set of worlds, an accessibility relation, and a valuation function) can be thought of as a complete description of an agent's epistemic state in terms of indistinguishability between worlds, we need to develop a semantics of $\langle \mathsf{R} \rangle^{\lambda} \varphi$ that appeals to a suitable relation between epistemic models. This relation should then be thought of as a *transition* from one epistemic model to another epistemic model, or, equivalently, from one epistemic state to another epistemic state. Roughly, the relevant transition between epistemic states S and S' should ensure that whenever an agent has performed an inference in state S, the agent knows the conclusion of this inference in state S'. This will allow us to capture a sufficiently strong notion of logical competence, despite that epistemic states themselves are not closed under any weak or strong notion of logical consequence.

To the best of the author's knowledge, this kind of dynamic model theory has not yet been developed to deal with logical omniscience. However, the general idea of modelling epistemic actions in terms of relations between models has been studied in quite some detail in the literature. For instance, Ditmarsch *et al.* [13] show how epistemic actions such as public announcement can be modelled using relations between models. There is reason to be optimistic that the tools developed in these studies can be fruitfully utilized in the context of developing a dynamic model theory of knowledge that solves the problem of logical omniscience.

Acknowledgements. Earlier versions of this paper were presented at the Annual Meeting of Danish Philosophy Society 2015 and at the BA philosophy student conference 2014 at Aarhus University. I would like to thank the audiences at both occasions for helpful feedback. A special thanks to Samantha Bennett, Jørgen Berthelsen, Jens Christian Bjerring, Casper Storm Hansen, Thomas Mikaelsen, Rasmus Pedersen, and an anonymous referee for very helpful comments and suggestions.

References

[1] Anderson, A., and N. Belnap, *Entailment. The Logic of Relevance and Necessity*, Princeton University Press, 1975.

[2] Ågotnes, T., "A logic of finite syntactic epistemic states", 2004.

[3] Ågotnes, T., and N. Alechina, "The dynamics of syntactic knowledge", *Journal of Logic and Computation*, (2006): 1–34. DOI: 10.1093/logcom/exl019
[4] Belnap, N., *A Useful Four-Valued Logic*, D. Reichel, Dordrecht, 1977.
[5] Bjerring, J. C., "Non-ideal epistemic spaces", 2010.
[6] Bjerring, J. C., "Impossible worlds and logical omniscience: An impossibility result", *Synthese*, 190, 13 (2013): 2505–2524. DOI: 10.1007/s11229-011-0038-y
[7] Bjerring, J. C., "Problems in epistemic space", *Journal of Philosophical Logic*, 43 (2014): 153–170. DOI: 10.1007/s10992-012-9257-z
[8] Bostock, D., *Intermediate Logic*, Oxford, Oxford University Press, 1997.
[9] Chalmers, D., "The nature of epistemic space", pages 60–107 in *Epistemic Modality*, New York, Oxford University Press, 2011. DOI: 10.1093/acprof:oso/9780199591596.003.0003
[10] Cresswell, M. J, "Classical intensional logic", *Theoria*, 36, 3 (1970): 347–372. DOI: 10.1111/j.1755-2567.1970.tb00433.x
[11] Cresswell, M. J, "Intensional logics and logical truth", *Journal of Philosophical Logic*, 1, 1 (1972): 2–15. DOI: 10.1007/BF00649986
[12] Cresswell, M. J, *Logics and Languages*, Methuen and Co., 1973.
[13] Ditmarsch, W., W. Hoek, and B. Kooi, *Dynamic Epistemic Logic*, Springer, 2008.
[14] Drapkin, J., and D. Perlis, "A preliminary excursion into step-logics", pages 262–269 in *Proceedings of the SIGART International Symposium on Methodologies for Intelligent Systems*, C. Ghidini, P. Giodini, and W. van der Hoek (eds.), 1986. DOI: 10.1145/12808.12837
[15] Drapkin, J., and D. Perlis, 'Reasoning situated in time I: Basic concepts". *Journal of Experimental and Theoretical Artificial Intelligence*, 2 (1990): 75–98.
[16] Duc, H. N., "Logical omniscience vs. logical ignorance on a dilemma of epistemic logic", Chapter 20 in *Progress in Artificial Intelligence. 7th Portuguese Conference on Artificial Intelligence, EPIA'95, Funchal, Madeira Island, Portugal*, C. A. Pinto-Ferreira and N. J. Mamede (eds.), 1995. DOI: 10.1007/3-540-60428-6_20
[17] Duc, H. N., "Reasoning about rational, but not logically omniscient, agents", *Journal of Logic and Computation*, 7, 5 (1997): 633–648. DOI: 10.1093/logcom/7.5.633
[18] Duc, H. N., "Resource-bounded reasoning about knowledge", 2001.
[19] Elgot-Drapkin, J., "Active logics: A unified formal approach to episodic reasoning", Technical report, University of Maryland, 1999.
[20] Fagin, R., and J. Halpern, "Belief, awareness and limited reasoning", *Artificial Intelligence*, 34 (1988): 39–76. DOI: 10.1016/0004-3702(87)90003-8

[21] Fagin, R., J. Halpern, and M. Vardi, "A nonstandard approach to the logical omniscience problem", *Artificial Intelligence*, 79 (1995): 203–240. DOI: 10.1016/0004-3702(94)00060-3

[22] Fagin, R., J. Halpern, M. Vardi, and Y. Moses, *Reasoning About Knowledge*, MIT Press, 1995.

[23] Heyting, A., *Intuitionism: An Introduction*, Amsterdam, North Holland, 1956.

[24] Hintikka, J., *Knowledge and Belief: An Introduction to the Two Notions*, Cornell University Press, 1962.

[25] Hintikka, J., "Impossible possible worlds vindicated", *Journal of Philosophical Logic*, 4, 4 (1975): 475–484. DOI: 10.1007/BF00558761

[26] Jago, M., "Logics for resource-bounded agents", 2006.

[27] Jago, M., "The problem of rational knowledge", *Erkenntnis*, 79, S6 (2014): 1151. DOI: 10.1007/s10670-013-9545-1

[28] Jago, M., *The Impossible: An Essay on Hyperintensionality*, Oxford University Press, 2014.

[29] Konolige, K., *A Deduction Model of Belief*, Morgan Kaufman, 1986.

[30] Lakemeyer, G., "Tractable meta-reasoning in propositional logics of belief", pages 198–202 in *Tenth International Joint Conference on Artificial Intelligence*, 1987.

[31] Lemmon, E. J., *Beginning Logic*, Hacker Publishing Company, 1998.

[32] Levesque, H. J., "A logic of implicit and explicit belief", pages 198–202 in *National Conference on Artificial Intelligence*, 1984.

[33] Priest, G., *An Introduction to Non-Classical Logic*, second edition, Cambridge University Press, 2008.

[34] Rantala, V., "Impossible worlds semantics and logical omniscience", *Acta Philosophica Fennica*, 35 (1982): 106–115.

[35] Routley, R., and R. Meyer, "The semantics of entailment II", *Journal of Philosophical Logic*, 1 (1972): 53–73. DOI: 10.1007/BF00649991

[36] Routley, R., and R. Meyer, "The semantics of entailment III", *Journal of Philosophical Logic*, 1 (1972): 192–208. DOI: 10.1007/BF00650498

[37] Routley, R., and R. Meyer, *The Semantics of Entailment I*, pages 194–243 in *Truth, Syntax and Modality*, series "Studies in Logic and the Foundations of Mathematics", 1973. DOI: 10.1016/S0049-237X(08)71541-6

[38] Wansing, H., "A general possible worlds framework for reasoning about knowledge and belief", *Studia Logica*, 49 (1990): 523–539. DOI: 10.1007/BF00370163

Mattias Skipper Rasmussen
University of Aarhus
Jens Chr. Skous Vej 7, DK-8000
`mattias.skipper.rasmussen@post.au.dk`