Stalnaker's Epistemic Logic on Isabelle/HOL.[*]

Laura P. Gamboa Guzman[0000−0003−3133−5148] and Kristin Y.
Rozier[0000−0002−6718−2828]

Iowa State University, Ames IA 50011, USA
lpgamboa@iastate.edu, kyrozier@iastate.edu

Abstract. The foundations of formal models for epistemic and doxastic
logics often rely on certain logical aspects of modal logics such as S4
and S4.2 and their semantics; however, the corresponding mathematical
results are often stated in papers or books without including a detailed
proof, or a reference to it, that allows the reader to convince themselves
about them. We reinforce the foundations of the epistemic logic S4.2 for
countably many agents by formalizing its soundness and completeness
results for the class of all weakly-directed pre-orders in the proof assis-
tant Isabelle/HOL. This logic corresponds to the knowledge fragment,
i.e., the logic for formulas that may only include knowledge modalities
in Stalnaker's system for knowledge and belief. Additionally, we formal-
ize the equivalence between two axiomatizations for S4, which are used
depending on the type of semantics given to the modal operators, as one
is commonly used for the relational semantics, and the other one arises
naturally from the topological semantics.

1   Introduction

Epistemic logics are a family of logics that allow us to reason about knowledge
among a group of agents, as well as their knowledge about other's knowledge[12].
Reasoning about knowledge is useful for detecting and identifying faults during
the operation of complex critical systems [7,25], where important safety proper-
ties are formalized using a modal language that combines temporal, in particular,
LTL (Linear Temporal Logic), and epistemic modal operators, so to verify the
correctness of the system using model checking and related formal fault-detection
techniques [9,19,22].

   When it comes to modal logics for knowledge, most of these logics correspond
to normal logics between S4 and S5 [11,23]. In particular, we consider Stalnaker's
epistemic logic, which coincides with the logic S4.2. It is known that this logic
strictly stronger than S4, but still weaker than S5 [8]. This logic is known to be
sound and complete with respect to all weakly directed S4-frames, that is, all
frames consisting of reflexive and transitive binary relations that are confluent
[21], but this proof is often omitted in text books where most extensions to
system K (the weakest normal modal logic) are usually treated informally.

Additionally, we encode in Isabelle/HOL the axiomatization of S4 obtained from the study of the topological interpretation for modal languages, which was introduced prior to the relational one that is more commonly found in the literature. This topological interpretation is done by reading the modal necessity operator as an interior operator on a topological space, for which is known that the modal logic S4 is complete with respect to all topological spaces [1]. The preferred axiomatization for the logic of topological spaces differs from the one presented in [14], not only from the set of axioms, but also the deductive rules, since it captures the axioms for an interior operator instead of a reflexive and transitive binary relation. As a consequence, this makes the topological axiomatization not directly recognizable as a normal modal logic, since the deduction rules seem to be weaker at first glance. Since several authors have been recently developing topological semantics for notions of knowledge and belief [2,3,4], we decided to provide a formalization for this result which often times is briefly mentioned and applied without being proved in detail.

Contributions

We formalize Stalnaker's epistemic logic, which is expressively equivalent to S4.2, as well as some intermediate results for the underlying propositional logic and the modal logics K, .2 and S4 mainly regarding rewriting rules, properties for maximal consistent sets of formulas, and frame properties that are induces by the chosen set of axioms in the proof assistant Isabelle/HOL. Our main result is a formalization of the soundness and completeness of Stalnaker's epistemic logic (restricted to countably many agents), which is known to be given by the system S4.2 [23], with respect to all weakly-directed[1] S4 frames, this is, all frames consisting of a non-empty set W and a binary relation R_i on W, one for each agent label i, that is reflexive, transitive, and that satisfies the property described by the following condition

$$\forall x \, \forall y \, \forall z \, (xR_iy \wedge xR_iz \implies \exists w \, yR_iw \wedge zR_iw).$$

The proof uses a Henkin-style completeness method, which is commonly used for this kind of logics [5] and was already available on Isabelle's Archive of Formal Proofs [15].

As far as we know, all systems corresponding to some multi-agent epistemic logic already formalized in Isabelle/HOL, which are all contained in [15] (the ground base for our formalization), were complete with respect to a class of frames that is characterized by a universal formula, i.e. a property of frames that is given by a first order formula of the form $\forall \overline{x} \, \phi(\overline{x})$, where $\phi(\overline{x})$ is a quantifier free formula with variable symbols in $\overline{x} = (x_1, \ldots, x_n)$. However, the logic S4.2 isc complete with respect to a class of frames that cannot be characterized using a universal formula, but instead it is characterized by a universal-existential formula. This universal-existential characterization makes it harder to formalize

---

[1] Binary relations satisfying this condition are also called *confluent* and *convergent* in the literature [20,21].

its completeness result, since one has to show the existence of an object in the universe of the canonical model satisfying a condition on a union of consistent sets of formulas. For this, we followed an argument given by Stalnaker in [24] that includes a set of theorems that are consequences of the axiom (.2) in K, which imply the consistency of a set obtained by taking the union of all known facts for an agent in two different worlds that were accessible from a third one.

Nonetheless, our formalization also includes some intermediate results that are well known for all normal modal logics, and that are commonly used when dealing with formal proofs in Hilbert-style systems. Finally, we formalized the equivalence between two of the most used axiomatizations of S4, the one presented in [14] which is commonly used when dealing with the relational semantics for it [5], and the one introduced by McKinsey and Tarski for the topological semantics [1].

Related work

Our ground base is the Isabelle/HOL theory `EpistemicLogic.thy` [15], which contains not only the formalization of other epistemic logics such as S4 and S5, but also formalizes the definition of an abstract canonical model, as well as a simple and convenient way to work with any desired normal modal logic by adding necessary the axioms to the basic system K [14]. A related paper to this formalization is [26], which contains a broad and updated summary on formalizations of logical systems and correspondent important results using different theorem provers. Other ways to formalize logical systems and their completeness results on Isabelle have also been studied in [10] and [6], which include (but are not limited to) the use of natural deduction rules, sequent-style rules, and tableau rules for the formal systems, and coinductive methods for soundness and completeness results.

However, given the already existent formalization of LTL in Isabelle/HOL [22], as well as the prevalence of Isabelle as a tool for formal verification of safety requirements for critical systems, it becomes important to provide this formalization for this particular proof assistant. In addition to this, in [17] the authors defined and investigated the notion of KB_R-structures, which are used to represent a description of the epistemic status of a rational agent that is not necessarily aware of their ignorance, and provided a result that matches them with models of the epistemic logic S4.2. Modal logics between S4.2 and S5 are of special interest for applications in epistemic logic, since they allow formalizations of several degrees of ignorance for each one of the agents [20].

The paper is organized as follows: Section 2 introduces the necessary background on epistemic logic, including its relational and topological semantics, and how the syntax and the relational semantics were formalized in Isabelle/HOL in [15]. Section 3 explains our formalization of Stalnaker's epistemic logic, including the intermediate results necessary to prove the main results, and the limitations of these to only countably many agents. Section 4 explains our formalization of the equivalence between the two most common axiomatizations for S4, the one that arises from the topological semantics, and the one commonly used when

working with the relational semantics. Finally, in Section 5, we conclude with a discussion about the results, limitations, and future work.

2   Background

2.1   Stalnaker's Epistemic logic

We briefly present the axiomatic system developed by R. Stalnaker for both notions of knowledge and belief, as well as the main result for the "knowledge formulas" (i.e. for those formulas that do not contain any belief modal operators), which correspond to the multimodal system S4.2 [23]. We omit the proof for this result, as the Isabelle theory "Epistemic logic: Completeness of Modal Logics" [15] does not support belief formulas.

Consider the well-formed formulas obtained from the following grammar, where x ranges over the set propositional symbols and i ranges over the set of agent labels:

$$\phi, \psi ::= \bot \,|\, x \,|\, \phi \vee \psi \,|\, \phi \wedge \psi \,|\, \phi \rightarrow \psi \,|\, K_i \phi \,|\, B_i \phi.$$

The operators K_i and B_i mean "agent i knows" and "agent i believes," respectively. Although Stalnaker does not present his logic of knowledge and belief using this exact set of propositional connectives, but a proper subset of these, we added the remaining ones given that From's formalization includes all of them.

Stalnaker's principles (axioms) for knowledge and belief are presented in Table 1, as well as their interpretation in natural language. *Stalnaker's logic for knowledge and belief* corresponds to the formal system obtained by adding these axioms to the axioms and rules of the multi-modal logic S4, that is, the smallest logic containing the following axioms:

- all propositional tautologies,
- axiom K: $(K_i(\phi \rightarrow \psi) \wedge K_i\phi) \rightarrow K_i\psi$,
- axiom T: $K_i\phi \rightarrow \phi$, and
- axiom 4: $K_i\phi \rightarrow K_iK_i\phi$;

and that is closed under Modus Ponens and the Necesitation rule, "from ϕ infer $K_i\phi$", where i ranges over the set of agents.

Table 1. Axioms for knowledge and belief.

$B_i\phi \rightarrow K_iB_i\phi$	Positive introspection
$\neg B_i\phi \rightarrow K_i\neg B_i\phi$	Negative introspection
$K_i\phi \rightarrow B_i\phi$	Knowledge implies belief
$B_i\phi \rightarrow \neg B_i\neg\phi$	Consistency of belief
$B_i\phi \rightarrow B_iK_i\phi$	Strong belief

The following proposition summarizes some relevant properties of this logic.

Proposition 1. *The following are some key properties of Stalnaker's logic for knowledge and belief [23].*

1. *The following equivalences, one for each agent label i and formula ϕ, are theorems in this logic:*
$$B_i\phi \longleftrightarrow \neg K_i \neg K_i \phi.$$

2. *As a consequence of the previous property, by replacing 'B_i' with '$\neg K_i \neg K_i$' in the* Consistency of belief *axiom, we get that $\neg K_i \neg K_i \phi \rightarrow K_i \neg K_i \neg \phi$ (also known as axiom .2) is a theorem in this logic. This implies that the knowledge formulas of this logic correspond exactly to the logic given by the system S4.2, i.e. those that can be obtained from the rules and axioms of the multi-modal logic S4 in presence of the axiom .2.*

The above proposition allows us to interpret this logic by giving a semantics only for the propositional variables, boolean connectives and knowledge operators. Formally, we use structures $\mathfrak{M} = (\mathcal{F}, \pi)$ known as Kripke models, where the frame $\mathcal{F} = (W, (R_i)_i)$ is a pair consisting of a non-empty set of worlds W, a set of binary accessibility relations $R_i \subseteq W \times W$, one for each agent i, and $\pi : \mathrm{Var} \to 2^W$ is a valuation of propositional symbols. Formula satisfiability at a given world $w \in W$ is defined as follows:

$$
\begin{aligned}
&\mathfrak{M}, w \not\models \bot \\
&\mathfrak{M}, w \models x && \text{iff} && w \in \pi(x) \\
&\mathfrak{M}, w \models \phi \vee \psi && \text{iff} && \mathfrak{M}, w \models \phi \text{ or } \mathfrak{M}, w \models \psi \\
&\mathfrak{M}, w \models \phi \wedge \psi && \text{iff} && \mathfrak{M}, w \models \phi \text{ and } \mathfrak{M}, w \models \psi \\
&\mathfrak{M}, w \models \phi \to \psi && \text{iff} && \mathfrak{M}, w \not\models \phi \text{ or } \mathfrak{M}, w \models \psi \\
&\mathfrak{M}, w \models K_i\phi && \text{iff} && \forall v \in W \, (wR_iv \to \mathfrak{M}, v \models \phi) \\
&\mathfrak{M}, w \models B_i\phi && \text{iff} && \mathfrak{M}, w \models \neg K_i \neg K_i \phi.
\end{aligned}
$$

One can use functions $\mathcal{K} : W \to 2^W$ instead of sets of ordered pairs $R \subseteq W \times W$, as there is a correspondence between these objects by setting

$$wRv \iff v \in \mathcal{K}(w),$$

for all $w, v \in W$.

2.2   Epistemic Logic: Completeness of Modal Logics

The "Epistemic Logic: Completeness of Modal Logics" entry on Isabelle's AFP [15] contains not only a formalization for the completeness results for some epistemic logics, but it also formalizes the general strategy for Henkin-style proofs for completeness. This is what allowed us to formalize a proof for the completeness result for S4.2. We show here the formalization of the Kripke models from [15], which uses functions instead of sets of ordered pairs.

datatype $('i, \, 'w) \, kripke =$
  Kripke ($\mathcal{W}$: ‹$'w \, set$›) (π: ‹$'w \Rightarrow id \Rightarrow bool$›) ($\mathcal{K}$: ‹$'i \Rightarrow 'w \Rightarrow 'w \, set$›)

Consequently, given a Kripke model $\mathfrak{M} = (W, (\mathcal{K}_i)_{i \in I}, \pi)$ with accessibility functions $\mathcal{K}_i$ for each agent i, formula satisfiability is defined by setting

$$\mathfrak{M}, w \models K_i \phi \quad \text{iff} \quad \forall v \in \mathcal{K}_i(w) \, (\mathfrak{M}, v \models \phi).$$

Additionally, the *dual* operator for each knowledge operator K_i is denoted in this formalization as L_i and is defined as a short hand for 'agent i does not know if something is false'. In other words, $L_i \phi := \neg K_i(\neg \phi)$, for all formulas ϕ. The Kripke semantics for this operator corresponds to [5,8]

$$\mathfrak{M}, w \models L_i \phi \quad \text{iff} \quad \exists v \in \mathcal{K}_i(w) \, (\mathfrak{M}, v \models \phi).$$

We show here the corresponding formalization presented in [14] for the Kripke semantics.

primrec *semantics* :: ‹($'i$, $'w$) *kripke* $\Rightarrow$ $'w$ $\Rightarrow$ $'i$ *fm* $\Rightarrow$ *bool*›
  (-, - $\models$ - [*50, 50*] *50*) **where**
  ‹(M, $w \models \bot$) = *False*›
| ‹(M, $w \models$ *Pro x*) = π M w x›
| ‹(M, $w \models (p \vee q)$) = ((M, $w \models p$) $\vee$ (M, $w \models q$))›
| ‹(M, $w \models (p \wedge q)$) = ((M, $w \models p$) $\wedge$ (M, $w \models q$))›
| ‹(M, $w \models (p \longrightarrow q)$) = (($M$, $w \models p$) $\longrightarrow$ (M, $w \models q$))›
| ‹(M, $w \models K$ i p) = ($\forall v \in \mathcal{W}$ $M \cap \mathcal{K}$ M i w. M, $v \models p$)›

From's formalization then focuses on proving the soundness and completeness results for each of the most commonly found *normal modal logics* in the literature concerning certain *classes of frames* [5,8,14]. We now summarize the relevant ones for our formalization.

1. The basic logic, K, whose corresponding axiomatic system consists of all propositional tautologies and the axiom K, and is closed under Modus Ponens and the Necessitation Rule, is sound and complete with respect to the class of all frames.
2. The logic S4 is sound and complete with respect to the class of all transitive and reflexive frames.

Notice that From's formalization does not include modal operators for belief, this restricts ourselves to the knowledge fragment of the language. However, Proposition 1 tells us that belief is equivalent to knowledge, thus we do not lose any information by restricting to the knowledge fragment.

2.3 Topological semantics and its axioms

The topological semantics for modal logics was introduced before the relational semantics that are now a days dominant in this field [1], and it was indeed there from where the first semantics completeness proof for S4 was produced. Let us recall now the notion of this topological semantics for a language with a single

modal operator $\Box$. Let $\mathcal{L}$ be the language composed of all formulas given by the following grammar:

$$\phi, \psi ::= x \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \phi \rightarrow \psi \mid \Box\phi,$$

where x ranges over the set of propositional symbols Var. Formulas in $\mathcal{L}$ are interpreted in a topological model $M = \langle W, \tau, v \rangle$, consisting of a non-empty set W, a topology τ over W, and a valuation $v : \text{Var} \rightarrow 2^W$ in the following way:

- $M, w \models x$ iff $x \in v(x)$;
- $M, w \models \neg\phi$ iff $M, x \not\models \phi$;
- $M, w \models \phi \wedge \psi$ iff $M, x \models \phi$ and $M, x \models \psi$;
- $M, w \models \phi \vee \psi$ iff $M, x \models \phi$ or $M, x \models \psi$;
- $M, w \models \phi \rightarrow \psi$ iff $M, x \not\models \phi$ or $M, x \models \psi$;
- $M, w \models \Box\phi$ iff there exists $U \in \tau$ such that $w \in U$ and $M, y \models \phi$ for all $y \in U$.

Although there is nothing inherently wrong with using deductive system presented in [14] for the logic S4, the following axiomatization to one most commonly known is often preferred when working with the topological semantics, as the meaning of the axioms and rules under this semantics resembles some well known properties of topological spaces [1].

Table 2. Topological S4 axioms and rules.

Axiom	Formula	Rule	Formula
N	$\Box\top$	MP	$\dfrac{\phi \rightarrow \psi \quad \phi}{\psi}$
R	$\Box(\phi \wedge \psi) \leftrightarrow (\Box\phi \wedge \Box\psi)$		
T	$\Box\phi \rightarrow \psi$	M	$\dfrac{\phi \rightarrow \psi}{\Box\phi \rightarrow \Box\psi}$
4	$\Box\phi \rightarrow \Box\Box\phi$		

Notice that at first it is not obvious weather or not the logic obtained from this axiomatization is a normal modal logic, often defined as a logic which *extends* system K [5], as neither axiom K nor the Necessitation Rule are present in the list of axioms and rules. However, we formalized a proof for the equivalence between both axiomatizations in the context of a multi-agent epistemic logic, as in recent years several authors have been developing topological semantics for notions of knowledge and belief [2,3,4], where often times this result is briefly mentioned and applied, but not proved in detail. Nonetheless, it is also worth noticing here that the relational semantics of S4 is no more than a particular case for the topological semantics, as one can assign a binary relation to each topological space by defining what is known as the *specialization order* [1].

3 Formalization

We now consider the epistemic logic based on the axioms in Table 1 and the results in Proposition 1 for the knowledge fragment of the language, and we

prove the soundness and completeness of the pure epistemic logic obtained from such system with respect to all frames consisting of weakly directed pre-orders. This is done by combining and applying the results formalized in [14] with some auxiliary results provided in the *Utility* section of our Isabelle theory, which allows us to utilize the canonical model strategy to prove completeness of the obtained system. We do not formalize a logic for both knowledge and belief, since we aimed to work on top of the formalization in [15], which considers modalities only for knowledge. Formalizing the whole logic for both knowledge and belief will then require to develop a new theory almost from scratch that includes modalities for both notions.

In order to do this, we prove first some intermediate results towards the completeness of the system obtained by adding axiom .2 to the system K^2, including some results about the underlying propositional logic. This system is known to be complete with respect to the class of weakly directed frames, and, although we do not formalize this result completely, we do formalize a version of the Truth lemma for this system, which is needed so that we can combine it with the results for system S4 formalized in [14] to achieve our goal of formalizing the completeness result for the logic S4.2 with respect to all weakly directed pre-orders.

3.1   Rewriting propositional and modal formulas

In the deductive system formalized in [15], deduction from a set of premises is defined as follows: given a set of formulas $\Gamma \cup \{\phi\}$, we say that "ϕ is derived from Γ" (denoted $\Gamma \vdash \phi$) if there are formulas $\psi_1, \ldots, \psi_k$ in Γ such that the formula $\psi_1 \rightarrow (\psi_2 \rightarrow \ldots (\psi_k \rightarrow \phi))$ is a theorem in the system, where k is a non-negative integer. It is well known that this formula is logically equivalent to $(\psi_1 \wedge \ldots \wedge \psi_k) \rightarrow \phi$ in classical propositional logic, thus the notion can be defined by requiring the latter to be a theorem in the system instead. Being able to translate between this two equivalent formulas in our formal deductive system played an important role for the proof of our main result, thus we provided a formalization of several results of this kind in the *Utility* section of our Isabelle theory, which includes some results that were not used later but that might become handy for the development of the formalization of other related theories in the future.

Similarly to the **imply** function in [15], which produces, from a list of formulas $[\psi_1, \ldots, \psi_k]$ and a formula ϕ, the formula $\psi_1 \rightarrow (\psi_2 \rightarrow \ldots (\psi_k \rightarrow \phi))$, we introduced the function **conjunct**, which takes a list of formulas $[\psi_1, \ldots, \psi_k]$ and produces the formula $\psi_1 \wedge \ldots \wedge \psi_k \wedge \top$ (notice that the input may be an empty list, in which case the output is $\top$). We formalized some results about logical equivalences, derived rules and maximal consistent sets regarding the **imply** and **conjunct** functions that are well know for the logic K, some of which are presented in the following lemmas, which are required to prove in section 3.2 that the axiom .2 induces the weakly directed property on all frames that

---

[2] This is also known as system G in the literature [8].

satisfies it, following [24]. We include the proofs for those lemmas that require elaborated arguments.

Lemma 1 (Derived rules). *For all formulas $\psi_1, \ldots, \psi_k, \phi, \vdash (\psi_1 \wedge \ldots \wedge \psi_k) \to \phi$ if and only if $\vdash \psi_1 \to (\psi_2 \to \ldots (\psi_k \to \phi))$.*

lemma *K-imply-conjunct*:
  assumes ‹$A \vdash imply\ G\ p$›
  shows ‹$A \vdash ((conjunct\ G) \longrightarrow p)$›

lemma *K-conjunct-imply*:
  assumes ‹$A \vdash ((conjunct\ G) \longrightarrow p)$›
  shows ‹$A \vdash imply\ G\ p$›

Lemma 2 (Logical equivalences). *The following two lemmas capture the fact that in system K, hence in any normal modal logic, the formulas $(K_i\psi_1 \wedge \ldots \wedge K_i\psi_k)$ and $K_i(\psi_1 \wedge \ldots \wedge \psi_k)$ are equivalent, for any finite set of formulas $\psi_1, \ldots, \psi_k$ and any agent i.*

lemma *K-conjunction-in-mult*: ‹$A \vdash ((K\ i\ (conjunct\ G)) \longrightarrow conjunct\ (map\ (K\ i)\ G))$›

lemma *K-conjunction-out-mult*: ‹$A \vdash (conjunct\ (map\ (K\ i)\ G) \longrightarrow (K\ i\ (conjunct\ G)))$›

Lemma 3 (Closure under conjunctions for MCSs). *The following lemma proves that maximal consistent sets of formulas are closed under conjunctions, that is, if Γ is a maximal consistent set of formulas and $\psi_1, \ldots, \psi_k$ are some formulas in Γ, then $\psi_1 \wedge \ldots \wedge \psi_k \in \Gamma$.*

lemma *mcs-conjunction-mult*:
  assumes ‹$consistent\ A\ V$› **and** ‹$maximal\ A\ V$›
  shows ‹$(set\ (S :: ('i\ fm\ list)) \subseteq V \wedge finite\ (set\ S)) \longrightarrow (conjunct\ S) \in V$›

Lemma 4. *For all formulas ϕ, ψ, θ, it is the case that*

$$\vdash ((K_i\phi \wedge K_i\psi) \to \theta) \to (K_i(\phi \wedge \psi) \to \theta)$$

for any agent label i, as long as the type of i is countable.

lemma *K-conj-imply-factor*:
  fixes $A :: $ ‹$(('i :: countable)\ fm \Rightarrow bool)$›
  shows ‹$A \vdash ((((K\ i\ p) \wedge (K\ i\ q)) \longrightarrow r) \longrightarrow ((K\ i\ (p \wedge q)) \longrightarrow r))$›

The assumption over the set of agents labels for the previous lemma is imposed by the proof that was formalized for it, as it relies on the proof for the completeness of K in [15] which requires this condition, as depicted in Figure 1.

Additionally, we formalized the following lemma, which plays a significant role in the proof of the completeness result for Stalnaker's epistemic logic in this formalization, which follows [24].
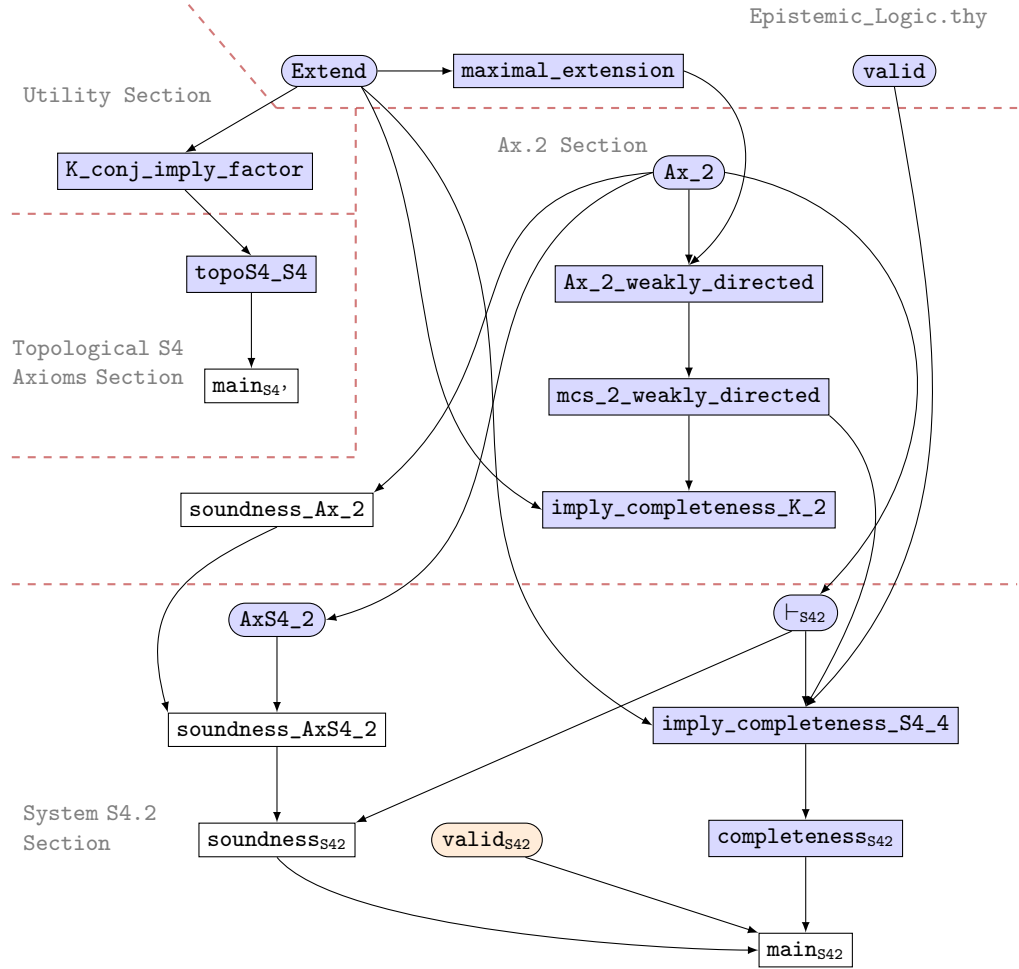
Fig. 1. Dependency graph showing the main results and the definitions, abbreviations, and interpediate results from their proofs that require the countable type condition. The dotted lines and the grey text show the files or sections of the Isabelle theory corresponding to our formalization where these can be found. Definitions and abbreviations are represented by rounded rectangles, whereas lemmas and theorems are represented by rectangles. Those that explicitly mention the countability condition are colored in blue, and the color orange means that this result is applied using the set of natural numbers to label the agents.

Lemma 5. *Given any pair of formulas* ϕ, ψ, $(K_i\phi \wedge L_i\psi) \rightarrow L_i(\phi \wedge \psi)$ *is a theorem in system K, for every agent label* i.

lemma *K-thm*: ‹$A \vdash ((K\ i\ p) \wedge (L\ i\ q) \longrightarrow L\ i\ (p \wedge q))$›

Proof. First, notice that $\vdash \phi \rightarrow (\neg(\phi \wedge \psi) \rightarrow \neg\psi)$, hence $\vdash K_i\phi \rightarrow K_i(\neg(\phi \wedge \psi) \rightarrow \neg\psi)$. On the other hand, we have that $\vdash K_i(\neg(\phi \wedge \psi) \rightarrow \neg\psi) \rightarrow ((K_i\neg(\phi \wedge \psi)) \rightarrow K_i\neg\psi)$, thus $\vdash K_i\phi \rightarrow ((K_i\neg(\phi \wedge \psi)) \rightarrow K_i\neg\psi)$. This implies that $\vdash K_i\phi \rightarrow (L_i\psi \rightarrow L_i(\phi \wedge \psi))$, which is equivalent to $\vdash (K_i\phi \wedge L_i\psi) \rightarrow L_i(\phi \wedge \psi)$.

3.2   Axiom .2

We formalize axiom schema .2, which when added to the axioms and rules of system K imposes a structure on the canonical model, namely, we obtain a weakly directed frame. For this, the **inductive** command let us define the axiom schema in such a way that i and p can be instantiated at will, as long as the type for the agents labels is countable.

inductive *Ax-2* :: ‹($'i$ :: *countable*) *fm* $\Rightarrow$ *bool*› **where**
   ‹*Ax-2* ($\neg$ *K i* ($\neg$ *K i p*) $\longrightarrow$ *K i* ($\neg$ *K i* ($\neg$ *p*)))›

A frame $\mathcal{F} = (W, (R)_{i \in I})$ is said to be *weakly directed* if whenever vR_iw and vR_iu, there exists $x \in W$ such that wR_ix and uR_ix, for each $i \in I$. For this, we formalize this property for Kripke frames as follows:

definition *weakly-directed* :: ‹($'i$, $'s$) *kripke* $\Rightarrow$ *bool*› **where**
   ‹*weakly-directed* $M \equiv \forall i.\ \forall s \in \mathcal{W}\ M.\ \forall t \in \mathcal{W}\ M.\ \forall r \in \mathcal{W}\ M.$
     $(r \in \mathcal{K}\ M\ i\ s \wedge t \in \mathcal{K}\ M\ i\ s) \longrightarrow (\exists\ u \in \mathcal{W}\ M.\ (u \in \mathcal{K}\ M\ i\ r \wedge u \in \mathcal{K}\ M\ i\ t))$›

The soundness of axiom schema .2 with respect to weakly directed frames is formalized in our Isabelle theory, and it follows from the definitions for the semantics and the weakly directed property. However, proving that the property holds for the canonical model when adding the axiom to a normal modal logic is non trivial. Unlike the frame properties imposed by the axioms considered in the Epistemic Logics formalized in [14], which are all universal properties, this property is universal-existential, so to prove that the canonical model has this property means that one has to show the existence of a possible world satisfying a property under some assumptions.

Recall that the *canonical frame*, $\mathcal{F}^{\mathrm{can}} = (W^{\mathrm{can}}, (R_i^{\mathrm{can}})_{i \in I})$, consists of the set all maximal consistent sets of formulas (with respect to K+.2) as the set of possible worlds, W^{can}, and the accessibility relations R_i^{can} are defined as follows:

$$\Gamma R_i^{\mathrm{can}} \Delta \quad \text{iff} \quad \{\phi :\ K_i\phi \in \Gamma\} \subseteq \Delta,$$

for each agent i. Thus, showing that the canonical frame for a system including axiom .2 is weakly directed means that one has to verify that if we have that $\{\phi :\ K_i\phi \in \Gamma\} \subseteq \Delta_1$ and $\{\phi :\ K_i\phi \in \Gamma\} \subseteq \Delta_2$ for some maximal consistent sets Γ, Δ_1 and Δ_2, then there exists a maximal consistent set Θ such that $\{\phi :\ K_i\phi \in \Delta_1\} \cup \{\phi :\ K_i\phi \in \Delta_2\} \subseteq \Theta$. This is captured in our formalization by the following lemma:

Lemma 6 (Weakly directed property and the axiom .2). *Suppose that* V, U, W *are three maximal consistent sets with respect to a normal modal logic containing the axiom .2. If* $V R_i^{\mathrm{can}} U$ *and* $V R_i^{\mathrm{can}} W$, *then there exists a maximal consistent set* X *such that* $U R_i^{\mathrm{can}} X$ *and* $W R_i^{\mathrm{can}} X$.

lemma *Ax-2-weakly-directed*:
  fixes $A :: \langle(('i :: countable)\ fm \Rightarrow bool)\rangle$
  assumes $\langle\forall p.\ Ax\text{-}2\ p \longrightarrow A\ p\rangle$ $\langle consistent\ A\ V\rangle$ $\langle maximal\ A\ V\rangle$
    and $\langle consistent\ A\ W\rangle$ $\langle maximal\ A\ W\rangle$ $\langle consistent\ A\ U\rangle$ $\langle maximal\ A\ U\rangle$
    and $\langle W \in reach\ A\ i\ V\rangle$ $\langle U \in reach\ A\ i\ V\rangle$
  shows $\langle\exists X.\ (consistent\ A\ X) \wedge (maximal\ A\ X) \wedge X \in (reach\ A\ i\ W) \cap (reach\ A\ i\ U)\rangle$

lemma *mcs-2-weakly-directed*:
  fixes $A :: \langle(('i :: countable)\ fm \Rightarrow bool)\rangle$
  assumes $\langle\forall p.\ Ax\text{-}2\ p \longrightarrow A\ p\rangle$
  shows $\langle weakly\text{-}directed\ (Kripke\ (mcss\ A)\ pi\ (reach\ A))\rangle$
  unfolding *weakly-directed-def*

Proof. For the first lemma in our formalization, fix a set of formulas A and three maximal consistent sets of formulas V, U, W (with respect to A) satisfying the lemma assumptions for some agent label i of a countable type. Assume towards contradiction that such set X does not exist, then

$$S := \{\phi :\ K_i\phi \in W\} \cup \{\phi :\ K_i\phi \in U\}$$

has to be inconsistent with respect to A, hence there are formulas $\theta_1, \ldots, \theta_k \in \{\phi :\ K_i\phi \in U\}$ and $\psi_1, \ldots, \psi_m \in \{\phi :\ K_i\phi \in W\}$, for some $k, m \in \mathbb{N}$, such that

$$A \vdash (\alpha \wedge \beta) \rightarrow \bot,$$

where $\alpha = \theta_1 \wedge \ldots \wedge \theta_k$ and $\beta = \psi_1 \wedge \ldots \wedge \psi_m$. This implies that $A \vdash K_i K_i(\neg(\alpha \wedge \beta))$, since $\phi \rightarrow \bot$ is equivalent to $\neg\phi$ for every formula ϕ, by applying the Necessitation rule twice. By definition, we have that $K_i\theta_1, \ldots, K_i\theta_k \in U$ and $K_i\psi_1, \ldots, K_i\psi_m \in W$, thus

$$K_i\theta_1 \wedge \ldots \wedge K_i\theta_k \in U \quad \text{and} \quad K_i\psi_1 \wedge \ldots \wedge K_i\psi_m \in W,$$

since these sets are closed under logical consequences. We then use the logical equivalences and properties for maximal consistent sets from the *Utility section* (see section 3.1) to obtain that $K_i\alpha \in U$ and $K_i\beta \in W$. This implies that the formulas $L_i K_i\alpha$ and $L_i K_i\beta$ are elements of V, and so is the formula $K_i L_i\alpha$, since V contains every instance of axiom .2 and is closed under logical consequences. This implies that $(L_i K_i\beta) \wedge (K_i L_i\alpha) \in V$, thus $L_i(K_i\beta \wedge L_i\alpha) \in V$, so there exists a maximal consistent set Z such that

$$V R_i^{\mathrm{can}} Z \quad \text{and} \quad K_i\beta \wedge L_i\alpha \in Z.$$

Applying the lemma *K-thm* we get that $L_i(\beta \wedge \alpha) \in Z$, but notice that $K_i\neg(\alpha \wedge \beta) \in Z$, thus $K_i\neg(\beta \wedge \alpha) \in Z$, which is a contradiction because we have found a formula ϕ such that $\phi, \neg\phi \in Z$.
The second lemma follows directly from the previous one.

    Note that we have restricted ourselves to countable types for the set of agent labels in formalization of the previous two lemmas, as we are only allowed to

extend a consistent set into a maximal one when the language is countable, because of a dependency shown in Figure 1. Unlike in the respective result for each normal modal logic formalized in [14], this restriction to a countable type was necessary as we were dealing with a universal-existence property and not with a purely universal one. With that being done, we prove a version of the Truth lemma for the minimal normal modal logic that includes axiom .2, which is the relevant result about this system that will allow us to prove the completeness result for system S4.2 in the next section.

Lemma 7 (Imply completeness for Axiom .2). *Let $\Gamma \cup \{\phi\}$ be a set of formulas. Suppose that, for all weakly directed Kripke structures M, $M, w \models \Gamma$ implies $M, w \models \phi$, for each world $w \in M$. Then there are formulas $\gamma_1, \gamma_2, \ldots, \gamma_m \in \Gamma$ such that*

$$\vdash_{.2} \gamma_1 \to (\gamma_2 \to \ldots (\gamma_m \to \phi) \ldots).$$

lemma *imply-completeness-K-2*:
  assumes *valid*: ‹∀ (*M* :: (′*i* :: *countable*, ′*i* *fm* *set*) *kripke*). ∀ *w* ∈ 𝒲 *M*.
    weakly-directed *M* ⟶ (∀ *q* ∈ *G*. *M*, *w* ⊨ *q*) ⟶ *M*, *w* ⊨ *p*›
  shows ‹∃ *qs*. *set* *qs* ⊆ *G* ∧ (*Ax-2* ⊢ *imply* *qs* *p*)›

We omit the proof for this lemma, since it follows the same strategy as the correspondent ones for the systems formalized in [15].

3.3   System S4.2

We define system S4.2 as the one obtained by adding to system K the axioms T, 4 and .2, by making use of the abbreviation $\oplus$ introduced in [15] that allows to combine axiom predicates:

abbreviation *SystemS4-2* :: ‹(′*i* :: *countable*) *fm* ⇒ *bool*› (⊢$_{S42}$ - [50] 50) **where**
  ‹⊢$_{S42}$ *p* ≡ *AxT* ⊕ *Ax4* ⊕ *Ax-2* ⊢ *p*›

Recall that axioms T and 4 impose reflexivity and transitivity on the canonical frame, respectively [5], which was formalized in [15]. This implies that the composition of these two with axiom .2 imposes all three conditions on the canonical frame, which leads to the soundness and completeness of S4.2 with respect to all weakly directed pre-orders. To prove the completeness result, we prove first the analog results to Lemmas 6 and 7 but for S4.2 and Kripke models based on weakly directed preorders.

Lemma 8 (S4.2 and Weakly directed preorders). *Let $\Gamma \cup \{\phi\}$ be a set of formulas. Suppose that, for all countable Kripke structures M based on weakly directed preorders, $M, w \models \Gamma$ implies $M, w \models \phi$, for each world $w \in M$. Then, there are formulas $\gamma_1, \ldots, \gamma_m \in \Gamma$ such that*

$$\vdash_{S42} \gamma_1 \to (\ldots \to (\gamma_m \to \phi) \ldots).$$

lemma *imply-completeness-S4-2*:
  assumes *valid*: ‹∀ (*M* :: (′*i* :: *countable*, ′*i* *fm* *set*) *kripke*). ∀ *w* ∈ 𝒲 *M*.
    w-directed-preorder *M* ⟶ (∀ *q* ∈ *G*. *M*, *w* ⊨ *q*) ⟶ *M*, *w* ⊨ *p*›
  shows ‹∃ *qs*. *set* *qs* ⊆ *G* ∧ (*AxS4-2* ⊢ *imply* *qs* *p*)›

This implies that if a formula is valid in all countable Kripke structures based on weakly directed preorders, then it is a theorem in S4.2.

lemma *completeness$_{S42}$*:
  assumes ‹∀ (*M* :: (′*i* :: *countable*, ′*i fm set*) *kripke*). ∀ *w* ∈ 𝒲 *M*. *w*-*directed*-*preorder*
M ⟶ *M*, *w* ⊨ *p*›
    shows ‹⊢$_{S42}$ *p*›

From this it follows our main result for this logic, namely the completeness of it with respect to all frames consisting of weakly directed pre-orders.

Theorem 1 (Completeness of S4.2). *A formula is valid in all countable Kripke structures based on weakly directed preorders if and only if it is a theorem in S4.2.*

theorem *main$_{S42}$*: ‹*valid$_{S42}$* *p* ⟷ ⊢$_{S42}$ *p*›

4   An alternative axiomatization for System S4

Inspired by the last section of [15], we formalized an alternative axiomatization for System S4 that is often used when dealing with the *topological semantics* [1] for modal operators and we show its equivalence to the one considered in [14]. We formalize the system corresponding to the axioms and rules in Table 2 on Isabelle, which we call topoS4, and, if a formula ϕ is a theorem in this system, we denote this by $\vdash_{Top} \phi$.

inductive *System-topoS4* :: ‹′*i fm* ⇒ *bool*› (⊢$_{Top}$ - [50] 50) **where**
    A1′: ‹*tautology p* ⟹ ⊢$_{Top}$ *p*›
| *AR*: ‹⊢$_{Top}$ ((*K i* (*p* ∧ *q*)) ⟷ ((*K i p*) ∧ *K i q*))›
| *AT′*: ‹⊢$_{Top}$ (*K i p* ⟶ *p*)›
| *A4′*: ‹⊢$_{Top}$ (*K i p* ⟶ *K i* (*K i p*))›
| *AN*: ‹⊢$_{Top}$ *K i* ⊤›
| *R1′*: ‹⊢$_{Top}$ *p* ⟹ ⊢$_{Top}$ (*p* ⟶ *q*) ⟹ ⊢$_{Top}$ *q*›
| *RM*: ‹⊢$_{Top}$ (*p* ⟶ *q*) ⟹ ⊢$_{Top}$ ((*K i p*) ⟶ *K i q*)›

To show that this formulation is equivalent to the one in [14] (which is based on [5]), we provide derivations of axiom K and the Necessitation Rule (from ϕ deduct $\Box\phi$). This is enough as our system already includes axioms T and 4 in the same fashion as in [14], and is based over the same propositional logic.

Lemma 9. *For all formulas ϕ and ψ, $\vdash_{top} (K_i\phi \land K_i(\phi \to \psi)) \to K_i\psi$ and, if $\vdash_{top} \phi$, then $\vdash_{top} K_i\phi$.*

lemma *topoS4-AxK*: ‹⊢$_{Top}$ (*K i p* ∧ *K i* (*p* ⟶ *q*) ⟶ *K i q*)›

lemma *topoS4-NecR*:
  assumes ‹⊢$_{Top}$ *p*›
    shows‹⊢$_{Top}$ *K i p*›

Proof. For the first part, notice that $\vdash_{top} (\phi \wedge (\phi \to \psi)) \to \psi$, since it is an instance of a propositional tautology. Then, we apply RM to obtain that $\vdash_{top} K_i(\phi \wedge (\phi \to \psi)) \to K_i\psi$, which implies that $\vdash_{top} (K_i\phi \wedge K_i(\phi \to \psi)) \to K_i\psi$. For the second one, suppose that $\vdash_{top} \phi$ and notice that $\vdash_{top} \phi \to (\top \to \phi)$, hence $\vdash_{top} \top \to \phi$. Applying RM we get that $\vdash_{top} K_i\top \to K_i\phi$, thus $\vdash_{top} K_i\phi$.

From this it then follows that any formula derivable in the classical S4 system (denoted $\vdash_{S4}$) can be derived in our system as well.

Lemma 10. *All S4 theorems are theorems in topoS4.*

lemma *S4-topoS4*: ‹$\vdash_{S4} p \implies \vdash_{Top} p$›

The converse follows by a similar argument, we show that axioms and rules from our system are all derivable in $\vdash_{S4}$, under the condition that there are only countably many agents.

Lemma 11. *All theorems in topoS4 are theorems in S4, assuming that there are only countably many agents.*

lemma *topoS4-S4*:
  fixes p :: ‹$('i :: countable)\ fm$›
  shows ‹$\vdash_{Top} p \implies \vdash_{S4} p$›

By combining the last two results with the main result for S4 in [14], we obtain formalized soundness and completeness for this alternative axiomatization of S4 over the class of S4 frames, namely, all reflexive and transitive frames.

Theorem 2 (Soundness and Completeness of topoS4). *A formula is valid in all S4 Kripke models if and only if it is a theorem in topoS4.*

theorem $main_{S4}'$: ‹$valid_{S4}\ p \longleftrightarrow (\vdash_{Top} p)$›

5   Results, Discussion and Future work

We have formalized the soundness and completeness for Stalnaker's Epistemic Logic S4.2 with respect to the class of Kripke frames consisting of weakly-directed pre-orders for countably many agents, which has not been formalized before neither in Isabelle, nor in any other publicly available proof assistant. Additionally, the equivalence between the topological axiomatization of S4 and the in [15] is also described in this document. The proofs for the main result, as well as for many of the intermediate results, have been sketched before in multiple sources, but we were not able to find a unique source that includes a fully detail proof of this theorem, making this the first work of its kind. Additionally, given the recent interest in applications of the topological semantics for epistemic modal operators [2,3,4], some of which coincide with Stalnaker's epistemic logic, this provides a reinforcement for the foundations of these works.

We emphasize on the assumption of the cardinality of the set of agent labels, as it was necessary to impose such restriction even in some definitions in our

formalization, thus creating a discrepancy with the definitions commonly found in the literature. We present in Figure 1 a summary of the definitions and results of our formalization and the one in [15] that rely on this condition, since the formalization for the general strategy applied to obtain the completeness results in [15] requires it to be able to obtain maximal consistent sets. Although, in theory, it is possible to provide an argument for the general case using Zorn's lemma (this was also later noted in [14]), which is available in [13].

Further work in formalizing in Isabelle/HOL of different formal aspects of modal logics that include S4 operators, as is the case with many temporal logics like LTL with its *always* operator, for which a complete axiomatization is already known [16]; as well as concrete examples of epistemic scenarios based on Stalnaker's principles, like the example detailed in [18]. We hope that this work will facilitate further work in formalizing different logical systems in Isabelle/HOL.

References

1. Aiello, M.: Reasoning About Space: The Modal Way. Journal of Logic and Computation **13**(6), 889–920 (dec 2003). `https://doi.org/10.1093/logcom/13.6.889`, `https://academic.oup.com/logcom/article-lookup/doi/10.1093/logcom/13.6.889`
2. Baltag, A., Bezhanishvili, N., González, S.F.: Topological evidence logics: Multi-agent setting. presented at TbiLLC 2019 (9 2019), `https://staff.fnwi.uva.nl/n.bezhanishvili/Papers/TbiLLC2019.pdf`
3. Baltag, A., Bezhanishvili, N., Özgün, A., Smets, S.: Justified belief and the topology of evidence. In: Proceedings of the 23rd International Workshop on Logic, Language, Information, and Computation - Volume 9803. p. 83–103. Springer-Verlag, Berlin, Heidelberg (2016). `https://doi.org/10.1007/978-3-662-52921-8_6`, `https://doi.org/10.1007/978-3-662-52921-8_6`
4. Baltag, A., Bezhanishvili, N., Özgün, A., Smets, S.: The topology of belief, belief revision and defeasible knowledge. In: Grossi, D., Roy, O., Huang, H. (eds.) Logic, Rationality, and Interaction. pp. 27–40. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
5. Blackburn, P., de Rijke, M., Venema, Y.: Modal logic. Modal Logic (6 2001). `https://doi.org/10.1017/CBO9781107050884`
6. Blanchette, J.C., Popescu, A., Traytel, D.: Soundness and Completeness Proofs by Coinductive Methods. Journal of Automated Reasoning **58**(1), 149 – 179 (Jan 2017). `https://doi.org/10.1007/s10817-016-9391-3`, `https://hal.inria.fr/hal-01643157`
7. Bozzano, M., Cimatti, A., Gario, M., Tonetta, S.: Formal design of fault detection and identification components using temporal epistemic logic. In: International Conference on Tools and Algorithms for the Construction and Analysis of Systems. pp. 326–340. Springer (2014)
8. Chagrov, A., Chagrov, P., Zakharyaschev, M.: Modal Logic. Oxford logic guides, Clarendon Press (1997), `https://books.google.com/books?id=dhgi5NF4RtcC`
9. Cimatti, A., Gario, M., Tonetta, S.: A lazy approach to temporal epistemic logic model checking. In: Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems. p. 1218–1226. AAMAS '16, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC (2016)

10. Dionísio, F.M., Gouveia, P., Marcos, J.: Defining and using deductive systems with isabelle. Computing, Philosophy, and Cognition pp. 271–293 (2005), `http://temporallogic.org/courses/AppliedFormalMethods/DefiningAndUsingDeductiveSystemsWithIsabelle.pdf`
11. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer Netherlands (2008). `https://doi.org/10.1007/978-1-4020-5839-4`
12. Fagin, R., Halpern, J.Y., Moses, Y., Vardi, M.: Reasoning About Knowledge. MIT Press (1995)
13. Fleuriot, J.D., Nipkow, T., Sternagel, C.: Zorn's lemma (ported from larry paulson's zorn.thy in zf)., `https://isabelle.in.tum.de/dist/library/HOL/HOL/Zorn.html`
14. From, A.H.: Formalized soundness and completeness of epistemic logic. In: International Workshop on Logic, Language, Information, and Computation. pp. 1–15. Springer (2021)
15. From, A.H.: Epistemic logic: Completeness of modal logics. Archive of Formal Proofs (Oct 2018), `https://isa-afp.org/entries/Epistemic_Logic.html`, Formal proof development
16. Goldblatt, R.: Logics of Time and Computation. CSLI Publications (1992)
17. Koutras, C.D., Zikos, Y.: Relating truth, knowledge and belief in epistemic states. In: Liu, W. (ed.) Symbolic and Quantitative Approaches to Reasoning with Uncertainty. pp. 374–385. Springer Berlin Heidelberg, Berlin, Heidelberg (2011)
18. Koutras, C.D., Zikos, Y.: Relating truth, knowledge and belief in epistemic states (2015)
19. Merz, S.: Tla: Lamport's temporal logic of actions. Online: `https://isabelle.in.tum.de/library/HOL/HOL-TLA/README.html` (March 2005), directory in the Isabelle/HOL Library
20. Pacheco, L., Tanaka, K.: On the degrees of ignorance: via epistemic logic and μ-calculus. In: International Workshop on Philosophy and Logic of Social Reality (SOCREAL) 2022. pp. 74–78 (2022), `http://hdl.handle.net/2115/84806`
21. Rendsvig, R., Symons, J.: Epistemic Logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Summer 2021 edn. (2021)
22. Sickert, S.: Linear temporal logic. Archive of Formal Proofs (Mar 2016), `https://isa-afp.org/entries/LTL.html`, Formal proof development
23. Stalnaker, R.: On logics of knowledge and belief. Philosophical Studies **128**, 169–199 (3 2006). `https://doi.org/10.1007/S11098-005-4062-Y`
24. Stalnaker, R.: Lecture Notes | Modal Logic | Linguistics and Philosophy | MIT OpenCourseWare (2009), `https://dspace.mit.edu/bitstream/handle/1721.1/100157/24-244-fall-2009/contents/lecture-notes/index.htm`
25. Tonetta, S., Gario, M., Cimatti, A., Bozzano, M.: Formal design of asynchronous fault detection and identification components using temporal epistemic logic. Logical Methods in Computer Science **11** (2015)
26. Villadsen, J., From, A.H., Jensen, A.B., Schlichtkrull, A.: Interactive Theorem Proving for Logic and Information, pp. 25–48. Springer International Publishing, Cham (2022). `https://doi.org/10.1007/978-3-030-90138-7_2`, `https://doi.org/10.1007/978-3-030-90138-7_2`