The Topology of Belief, Belief Revision and Defeasible Knowledge

Alexandru Baltag[1], Nick Bezhanishvili[2⋆], Aybüke Özgün[1], Sonja Smets[1⋆⋆]

[1] University of Amsterdam, The Netherlands
[2] Utrecht University, The Netherlands

Abstract. We present a new topological semantics for doxastic logic, in which the belief modality is interpreted as the closure of the interior operator. We show that this semantics is the most general (extensional) semantics validating Stalnaker's epistemic-doxastic axioms [22] for "strong belief", understood as *subjective certainty*. We prove two completeness results, and we also give a topological semantics for update (dynamic conditioning), i.e. the operation of revising with "hard information" (modeled by restricting the topology to a subspace). Using this, we show that our setting fits well with the defeasibility analysis of knowledge [18]: topological knowledge coincides with undefeated true belief. Finally, we compare our semantics to the older topological interpretation of belief in terms of Cantor derivative [23].

1 Introduction

Ever since Edmund Gettier published his famous counterexamples [14], formal epistemologists have been concerned with understanding the relation between belief and knowledge, and in particular with finding the conditions that distinguish an item of belief (no matter how true and justified) from an item of knowledge. This question can be approached from two sides: 1) we start with the weakest notion of true justified (or justifiable) belief and add conditions in order to argue that they establish a "good" (e.g. factive, correctly-justified, unrevisable, coherent, stable, truth-sensitive) notion of knowledge; or 2) we start from a chosen notion of knowledge and weaken it to obtain a "good" (e.g. consistent, introspective, possibly false) notion of belief. Most research in formal epistemology follows the first approach. In particular, the standard *topological semantics for knowledge* (in terms of the interior operator) can be included within this first approach, as based on a notion of knowledge as "correctly justified belief": according to the interior semantics, a proposition (set of possible worlds) P is known if there exists some "true evidence" (i.e. an open set A containing the real world s) that entails P (i.e. $A \subseteq P$). Another example of the first approach is

⋆ Supported by the Dutch NWO grant 639.032.918 and the Rustaveli Science Foundation of Georgia grant FR/489/5-105/11.

⋆⋆ S. Smets contribution to this paper has received funding from the ERC under the European Community's 7th Framework Programme/ERC Grant agreement no. 283963.

the so-called *defeasibility analysis of knowledge* proposed by Lehrer and Paxson [18], Klein [16] and other authors: knowledge is defined as *undefeated (justified) true belief*; i.e. true belief that cannot be defeated by any new (true) evidence.

While most research in formal epistemology follows the first approach, the second approach has to date received much less attention from formal logicians. This is rather surprising, since such a "knowledge-first" approach has been persuasively defended by one of the most influential contemporary epistemologists (Williamson [26]). The only formal account following this second approach that we are aware of (prior to our own work) is the one given by Stalnaker [22], using a relational semantics for knowledge, based on Kripke models in which the accessibility relation is a directed preorder. In this setting, Stalnaker argues that the "true" logic of knowledge is the modal logic $S4.2$ and that belief can be defined as the *epistemic possibility of knowledge*. In other words, *believing p is equivalent to "not knowing that you don't know" p*:

$$Bp = \neg K \neg Kp.$$

Stalnaker justifies this identity from first principles based on a particular notion of belief, namely *belief as "subjective certainty"*. Stalnaker refers to this concept as "strong belief", but we prefer to call it *full belief*[3]. What is important about this type of belief is that it is *subjectively indistinguishable from knowledge*: an agent "fully believes" p iff in fact she "believes that she knows" p.

The resulting conception of knowledge is clearly different from Williamson's (who rejects the KK principle[4]), but it is closely related to the above-mentioned *defeasibility analysis* [17]. Indeed, Stalnaker proceeds to formalize AGM belief revision, based on a special case of the above semantics, in which the accessibility relation is assumed to be a weakly connected preorder, and (conditional) beliefs are defined by minimization. This validates the AGM principles for belief revision. Stalnaker shows that in this special case his notion of knowledge coincides with (a simplified and idealized version of) Lehrer's concept of *undefeated (justified) true belief*: i.e. true belief that cannot be defeated by (revising with) any new (true) evidence. However, this special case supports a stronger logic of knowledge (the system $S4.3$). Since Stalnaker defends the weaker $S4.2$ as the "true" logic of knowledge, he is lead to argue *against* the defeasibility theory.

[3] We adopt this terminology both because we want to avoid the clash with the very different notion of strong belief (due to Battigalli and Siniscalchi [2]) that is standard in epistemic game theory, and because we think that the intuitions behind Stalnaker's notion are very similar to the ones behind Van Fraassen's probabilistic concept of full belief [13].

[4] In formal epistemology, the "KK principle" is one of the names given to the axiom of Positive Introspection $K\varphi \to KK\varphi$, also known as the modal axiom 4. The well-known modal system $S4$ consists of this Positive Introspection axiom 4 together with the axiom T of Factivity, or Truthfulness ($K\varphi \to \varphi$), as well as Kripke's axiom K ($K(\varphi \to \psi) \to (K\varphi \to K\psi)$) and the rules of Modus Ponens and Necessitation. Stalnaker's system $S4.2$ includes the system $S4$, and hence contains the KK principle, contradicting Williamson's conception.

In this paper, we aim to generalize Stalnaker's formalization, making it independent from the concept of plausibility order and from relational semantics. In fact, we are looking for *the most general extensional semantics for "full belief"* (in the above-mentioned sense). By an "extensional" semantics we mean here any semantics that assigns the same meaning to sentences having the same extension. Essentially, an extensional semantics takes the meaning of a sentence to be given by a "U.C.L.A. proposition" in the sense of Anderson-Belnap-Dunn[5]: a set of possible worlds (intuitively thought of as the set of worlds at which the proposition is true). We prove that the most general extensional semantics validating Stalnaker's axioms is a *topological* one, that extends the standard topological interpretation of knowledge (as *interior* operator) with a new topological semantics for belief, given by the *closure of the interior* operator (with respect to an extremally disconnected topology). We compare our new semantics with the older topological interpretation of belief in terms of Cantor derivative, giving several arguments in favor of our semantics.

We prove that the logic of knowledge and belief with respect to our semantics is completely axiomatized by Stalnaker's epistemic-doxastic principles. Furthermore, we show that the complete logic of knowledge in this setting is indeed the system $S4.2$, while the complete logic of belief is the standard system $KD45$. We formalize the action of learning (conditioning with) new "hard" (true) information P as a *topological update* operator, using the *relativization* of the original topology to (the subspace corresponding to) the set P. This allows us to model belief revision of a more general type than the one axiomatized by the AGM theory. We show that, in this generalized setting, *Stalnaker's objections to the defeasibility theory of knowledge do not apply*: when interpreted (as interior) over topological spaces, Stalnaker's notion of knowledge (having $S4.2$ as its complete logic) *coincides with undefeated (justified) true belief.*

2 Background: Topological Interpretation of Knowledge

2.1 Topological Preliminaries

For the basic definitions of general topology we refer to [11] or any other textbook in General Topology. Here we just recall that a *topological space* is a pair (X, τ), where X is a non-empty set and τ is a *topology* on X, i.e. a family $\tau \subseteq \mathcal{P}(X)$ containing X and $\emptyset$ and closed under finite intersections and arbitrary unions. Elements of τ are called *open sets*. Complements of open sets are called *closed sets*. An open set containing $x \in X$ is called an *open neighbourhood* of x. The *interior* $\mathrm{Int}(A)$ of a set $A \subseteq X$ is the largest open set contained in A. The *closure* $\mathrm{Cl}(A)$ of A is the least closed set containing A. It is easy to check that $\mathrm{Int}(A) = X \setminus \mathrm{Cl}(X \setminus A)$.

[5] Dunn [10] explains this name as follows: 'The name honors the university that has had both R. Carnap and R. Montague in its faculty, since in modern times they (together with others, e.g. S. Kripke and R. Stalnaker) have been proponents of this construction. But the idea actually originates with Boole, who suggested thinking of propositions as "sets of cases" (...).'

An *alternative definition of topological spaces* (due to Kuratowski) takes the closure operation (or equivalently, the interior) as the basic notion. Stated in terms of interior, a topological space is a pair (X, Int), where X is a non-empty set and $\mathrm{Int} : \mathcal{P}(X) \to \mathcal{P}(X)$ is an operation satisfying the (dual of the) so-called *Kuratowski axioms*: $\mathrm{Int}(X) = X$, $\mathrm{Int}(A) \subseteq A$, $\mathrm{Int}(\mathrm{Int}(A)) = \mathrm{Int}(A)$, $\mathrm{Int}(A \cap B) = \mathrm{Int}(A) \cap \mathrm{Int}(B)$. In this setting, the family of open sets is defined by putting $\tau = \{A \subseteq X : \mathrm{Int}(A) = A\}$. It is easy to see that this is a topology. Indeed, the two definitions of topological spaces are equivalent.

2.2 The Interior Semantics for Modal Logic

We start by recalling the standard topological semantics of modal (epistemic) logic, originating in the work of Tarski and McKinsey [19]. We consider the standard unimodal language $\mathcal{L}_K$ with a countable set of propositional letters Prop, and a modal operator K. Formulas are defined as usual by

$$\varphi ::= \bot \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi,$$

where $p \in \mathsf{Prop}$. Abbreviations for the connectives $\vee$, $\to$ and $\leftrightarrow$ are standard. The possibility operator $\langle K\rangle\varphi$ is defined as $\langle K\rangle\varphi := \neg K\neg\varphi$.

Definition 1. *A* ***topological model*** $\mathcal{M} = (X, \tau, \nu)$ *is a tuple where* (X, τ) *is a topological space and* ν *is a* valuation, *i.e., a map* $\nu : \mathsf{Prop} \to \mathcal{P}(X)$*. We let* Cl *and* Int *denote the closure and interior operators, respectively. The* ***topological semantics for modal formulas*** *is defined by the following inductive definition, where* $\mathcal{M} = (X, \tau, \nu)$ *is a topological model and* $p \in \mathsf{Prop}$*:*

$$\begin{array}{ll} [\![\bot]\!]^{\mathcal{M}} = \emptyset, & [\![p]\!]^{\mathcal{M}} = \nu(p) \\ [\![\varphi \wedge \psi]\!]^{\mathcal{M}} = [\![\varphi]\!]^{\mathcal{M}} \cap [\![\psi]\!]^{\mathcal{M}} & [\![\neg\varphi]\!]^{\mathcal{M}} = X \setminus [\![\varphi]\!]^{\mathcal{M}} \\ [\![K\varphi]\!]^{\mathcal{M}} = \mathrm{Int}[\![\varphi]\!]^{\mathcal{M}} & \end{array}$$

As $\langle K\rangle\varphi$ is equivalent to $\neg K\neg\varphi$, it is easy to see that $[\![\langle K\rangle\varphi]\!]^{\mathcal{M}} = \mathrm{Cl}[\![\varphi]\!]^{\mathcal{M}}$. We skip the index $\mathcal{M}$ if it is clear from the context. Truth, validity, soundness and completeness wrt topological semantics are defined as usual.

Proposition 1. *(see e.g. [3], [20] and [7]) The modal logic* $S4$ *is sound and complete wrt all topological spaces.*

2.3 Epistemic Interpretation: open sets as pieces of evidence

The original reason for interpreting interior as knowledge was that the Kuratowski axioms match exactly the $S4$ axioms, and in particular the principles

$$(T) \quad Kp \to p$$

of Truthfulness of Knowledge ("factivity") and

$$(KK) \quad Kp \to KKp$$

of Positive Introspection of Knowledge (known as axiom 4 in modal logic).

Philosophically, one of the best arguments in favor of the topological semantics is negative: namely, the fact that it does *not* validate the principle

$$\neg Kp \to K\neg Kp$$

This principle, known as (5) or Negative Introspection, is rejected by essentially all philosophers. One of its undesirable consequences is that it makes it impossible for a rational agent to have wrong beliefs about her knowledge: *she always knows whatever she believes that she knows.* This is known in the literature as Voorbraak's paradox [25]: it contradicts the day-to-day experience of encountering agents who believe they know things that they do not actually know[6].

But, even beyond the issue of negative introspection, the topological semantics can arguably give us a deeper insight into the nature of knowledge and its evidential basis than the usual Kripke semantics. From an extensional point of view, the properties U that are directly observable by an agent naturally form an *open basis* for a topology: closure under finite intersections captures an agent's ability to combine finitely many pieces of evidence into a single piece[7]. A proposition P is true at world w if $w \in P$. If an open U is included in a set P, then we can say that *proposition P is entailed (supported, justified) by evidence U.* Open neighbourhoods U of the actual world w play the role of *sound (correct, truthful) evidence.* The actual world w is in the interior of P iff there exists such a sound piece of evidence U that supports P. So *the agent "knows" P if she has a correct justification for P* (based on a sound piece of evidence supporting P). Moreover, open sets will then correspond to properties that are in principle *verifiable* by the agent: whenever they are true they can be known. Dually, closed sets will correspond to *falsifiable* properties. See Vickers [24] and Kelly [15] for more on this interpretation and its connections to Epistemology, Logic and Learning Theory.

So the knowledge-as-interior conception can be seen as an implementation of one of the most widespread intuitive responses to Gettier's challenge: knowledge is "correctly justified belief" (rather than being simply true justified belief). To qualify as knowledge, not only the content of one's belief has to be truthful, but its evidential justification has to be sound.

2.4 Extensions and Improvements

The interior-based semantics for knowledge has been extended to multiple agents [4], to common knowledge [1, 6], to logics of learning ("topo-logic", see [20]), to topological versions of dynamic-epistemic logic [27]. See [3] for a comprehensive survey of the field.

---

[6] This common experience can be considered the starting point of all epistemological reflection, and historically played such a role, see e.g. in Platonic dialogues.

[7] But see van Benthem and Pacuit [5] for a more general logical account of evidence-management which relaxes this assumption: by using instead a neighbourhood semantics, this account can deal with agents who have not yet managed to combine all their pieces of evidence.

But there are two other topologically-based logics that are of particular interest in this paper. The first is an alternative semantics for modal logic, in terms of Cantor's derivative operation, which has been proposed as a semantics for belief. We will give a critical presentation of this alternative in Section 6. The second is a logic that *strengthens* $S4$, namely:

$$S4.2 = S4 + (\langle K \rangle Kp \rightarrow K \langle K \rangle p).$$

By $L + \varphi$ we denote the smallest normal modal logic containing L and φ.

Recall that a topological space (X, τ) is called *extremally disconnected* if the closure of every open subset of X is open.

Proposition 2. *[3, p. 253] $S4.2$ is sound and complete wrt all extremally disconnected topological spaces.*

We also recall that a topological space (X, τ) is called an *Alexandroff space* if the intersection of open sets of X is open. It is well known that Alexandroff spaces correspond to reflexive and transitive Kripke frames, see e.g., [3], [20] or [7]. Moreover, the evaluation of modal formulas in an Alexandroff space coincides with their evaluation in the corresponding Kripke frame.

A Kripke frame (X, R) is called *directed*[8] if

$$(\forall x, y, z)(xRy \wedge xRz) \rightarrow (\exists u)(yRu \wedge zRu).$$

It is well known, see e.g., [8] or [9] that $S4.2$ is sound and complete wrt reflexive, transitive and directed Kripke frames.

We give a few examples of extremally disconnected spaces. Alexandroff spaces corresponding to reflexive, transitive and directed Kripke frames are extremally disconnected. Another classical example of an extremally disconnected space is the Stone-Čech compactification $\beta(\mathbb{N})$ of the set of natural numbers with a discrete topology. Also it is well known that topological spaces that are Stone-dual to complete Boolean algebras are extremally disconnected [21].

3 The topology of full belief

3.1 Stalnaker's Epistemic-Doxastic Axioms

In his paper [22], Stalnaker proposes a very interesting analysis of the relationship between knowledge and (justified or justifiable) belief. This is based on a conception of belief as "subjective certainty": from the point of the agent in question, her belief is subjectively indistinguishable from her knowledge. In this paper, we will refer to Stalnaker's notion as "full belief".

The *bimodal language $\mathcal{L}_{KB}$ of knowledge and (full) belief* is given recursively:

$$\varphi ::= \perp \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi \mid B\varphi,$$

[8] Directedness is also called *confluence* or the *Church-Rosser property.*

where $p \in \mathsf{Prop}$. We will also consider two unimodal fragments of this language $\mathcal{L}_K$ (having K as its only modality) and $\mathcal{L}_B$ (having only B).

Stalnaker's epistemic-doxastic axioms for the logic KB are given in the Table below.

	Stalnaker's Epistemic-Doxastic Axioms	
(K)	$K(\varphi \to \psi) \to (K\varphi \to K\psi)$	Knowledge is additive
(T)	$K\varphi \to \varphi$	Knowledge implies truth
(KK)	$K\varphi \to KK\varphi$	Positive introspection for K
(CB)	$B\varphi \to \neg B\neg\varphi$	Consistency of belief
(PI)	$B\varphi \to KB\varphi$	(Strong) positive introspection of B
(NI)	$\neg B\varphi \to K\neg B\varphi$	(Strong) negative introspection of B
(KB)	$K\varphi \to B\varphi$	Knowledge implies Belief
(FB)	$B\varphi \to BK\varphi$	Full Belief
	Inference Rules	
(MP)	From φ and $\varphi \to \psi$ infer ψ.	Modus Ponens
(K-Nec)	From φ infer $K\varphi$.	Necessitation

We will refer to this axiomatic system as KB. The axioms seem very natural and uncontroversial: the first three are the $S4$ axioms for knowledge; (CB) captures the consistency of beliefs, and in the context of the other axioms will be equivalent to the modal axiom (D) for beliefs: $\neg B\bot$; (PI) and (NI) capture strong versions of introspection of beliefs: the agent knows what she believes and what not; (KB) means that agents believe what they know; and finally, (FB) captures the essence of "full belief" as subjective certainty (the agent believes that she knows all the things that she believes). Finally, the rules of Modus Ponens and Necessitation seem uncontroversial (for implicit knowledge, if not for explicit knowledge) and are accepted by a majority of authors (and in particular, they are implicitly used by Stalnaker).

The above axioms imply that *belief can be defined in terms of knowledge*:

Proposition 3. *(Stalnaker) The equivalence*

$$B\varphi \leftrightarrow \neg K\neg K\varphi$$

is provable in the system KB. Moreover, all the axioms of the standard system $KD45$ for belief are provable in the system KB, and in particular: Kripke's axiom for belief ($B(\varphi \to \psi) \to (B\varphi \to B\psi)$); the so-called axiom (D) ($\neg B\bot$); axiom 4 (positive introspection) for belief ($B\varphi \to BB\varphi$); the axiom 5 (negative introspection) for belief ($\neg B\varphi \to B\neg B\varphi$).

Finally, the formula $\langle K\rangle K\varphi \to K\langle K\rangle\varphi$ is also provable in KB: i.e. all the axioms of the system $S4.2$ hold for knowledge in the system KB.

3.2 Our topological semantics for full belief

Definition 2. *An **extensional (and compositional) semantics** for the language $\mathcal{L}_{KB}$ of knowledge and full belief is a triplet (X, B, K), where X is a set of possible worlds, and $B : \mathcal{P}(X) \to \mathcal{P}(X)$ and $K : \mathcal{P}(X) \to \mathcal{P}(X)$ are unary operations on (sub)sets of worlds.*

Any extensional semantics (X, B, K), together with a valuation $\nu : Prop \to \mathcal{P}(X)$, gives us an ***extensional model*** $M = (X, B, K, \nu)$*, in which we can interpret the formulas φ of $\mathcal{L}_{KB}$ in the obvious way: the clauses for propositional connectives are the same as in the topological semantics above, and in rest we put*

$$[\![K\varphi]\!]^{\mathcal{M}} = K[\![\varphi]\!]^{\mathcal{M}} \qquad [\![B\varphi]\!]^{\mathcal{M}} = B[\![\varphi]\!]^{\mathcal{M}}.$$

As usual, a formula is valid in an extensional semantics (X, B, K) if it is true at all worlds of all models $M = (X, B, K, \nu)$ based on it. An inference rule is valid if it preserves validity of formulas.

A special case of extensional semantics for the language $\mathcal{L}_{KB}$ is *our proposed topological semantics*:

Definition 3. *A* ***topological semantics for the language*** $\mathcal{L}_{KB}$ *is an extensional semantics (X, K^τ, B^τ), where (X, τ) is a topological space, $K^\tau = Int^\tau$ is the interior operator with respect to the topology τ, and $B^\tau = \mathrm{Cl}^\tau(\mathrm{Int}^\tau)$ is the closure of the interior with respect to τ.*

Proposition 4. *A topological space validates all the axioms and rules of the system $KD45$ for belief (with the semantics given above) iff it is extremally disconnected.*

The proof of this result is rather long and intricate, and is left for a future journal publication.

Proposition 5. *A topological space validates all the axioms and rules of Stalnaker's system KB (with the semantics given above) iff it is extremally disconnected.*

Proof. It is easy to check that extremally disconnected spaces validate all the the axioms of KB. The other direction follows from Propositions 3 and 4.

Now we can give a Topological Representation Theorem for extensional models of KB:

Theorem 1. *An extensional semantics (X, B, K) validates all the axioms and rules of Stalnaker's system KB iff it is a topological semantics given by an extremally disconnected topology τ on X (such that $K = K^\tau = \mathrm{Int}^\tau$ and $B = B^\tau = \mathrm{Cl}^\tau(\mathrm{Int}^\tau)$).*

Proof. One direction is proved in the previous Proposition, so let us look at the other direction. Suppose an extensional semantics (X, B, K) validates the axioms and rules of KB. Then the validity of the $S4$ axioms implies that K satisfies the Kuratowski conditions for topological interior, and so it gives rise to a topology τ in which $K = \mathrm{Int}^\tau$. By the Proposition in the previous section, the KB axioms imply that $B = \neg K \neg K$, i.e. $B = \neg \mathrm{Int}^\tau \neg \mathrm{Int}^\tau = \mathrm{Cl}^\tau \mathrm{Int}^\tau$. Hence, (X, B, K) *is* a topological semantics, in the sense above, for a topology τ. By the previous Proposition, the validity of KB implies that τ is extremally disconnected.

This last result shows that Stalnaker's axioms are just an alternative axiomatization of extremally disconnected topological spaces, in which both interior and the closure of interior are taken as primitive operations. The conclusion is that *our topological semantics is indeed the most general (extensional compositional) semantics validating Stalnaker's axioms.*

3.3 Completeness Results

Let us first look at the bimodal logic KB of knowledge and (full) belief.

Theorem 2. *The (sound and) complete logic of knowledge and belief on extremally disconnected spaces is given by Stalnaker's system KB.*

Proof. This follows trivially from our Topological Representation Theorem for extensional models of KB (Theorem 1 in Section 3.2).

Next, we look at the unimodal fragment $\mathcal{L}_K$ having K as the only modality. In fact, this language has exactly the same expressivity as KB (since the belief operator can be eliminated via the identity $B\varphi = \neg K \neg K\varphi$). Moreover, we already know (by Proposition 2 in section 2.3) that *the sound and complete logic of knowledge on extremally disconnected spaces is $S4.2$.*

Further, we look at the unimodal fragment $\mathcal{L}_K$ having B as the only modality: this logic is less expressive than the bimodal language KB, since knowledge is not reducible to belief.

Theorem 3. *The complete logic of belief on extremally disconnected spaces is $KD45$.*

This result, though unsurprising, is technically the hardest result in this paper . The proof is long and intricate, and is left for a future journal publication.

4 From updates to defeasible knowledge

Conditioning (with respect to some qualitative plausibility order or to a probability measure) is the most widespread way to model the learning of "hard" information[9]. The prior plausibility/probability assignment (encoding the agent's original beliefs before the learning) is changed to a new such assignment, obtained from the first one by conditioning with the new information P. In the qualitative case, this means just restricting the original order to P-worlds; while in the probabilistic case, restriction has to be followed by re-normalization (to ensure that the probabilities newly assigned to the remaining worlds add up to 1). In Dynamic Epistemic Logic, one makes also a distinction between *simple ("static") conditioning* and *dynamic conditioning* (also known as "update").

[9] This term is used to denote information that comes with an inherent warranty of veracity, e.g. because of originating from an infallibly truthful source.

The first essentially corresponds to conditional beliefs: the change is made only locally, affecting only one occurrence of the belief operator $B\varphi$ (which is thus locally replaced by conditional belief $B^P\varphi$) or of the probability measure (which is locally replaced by conditional probability). In contrast, an update is a global change, at the level of the whole model (thus recursively affecting the meaning of all occurrences of doxastic operators). In this paper, due to space restrictions, we only investigate the natural topological analogue of *dynamic conditioning*.

Topological Updates. As recognized already in [27] (among others), the natural topological analogue of dynamic conditioning (update) is the operation of taking the *restriction (or "relativization") of a topology* τ on X to a subset $P \subseteq X$. What we obtain in this way is a *subspace* of the original topological space. Given a topological space (X, τ) and a set $P \subseteq X$, a space (P, τ_P) is called a *subspace* of (X, τ) if $\tau_P = \{U \cap P : U \in \tau\}$. It is well-known that the closure and interior operators in the relativized topology (P, τ_P), denoted by Cl_{τ_P} and Int_{τ_P} respectively, satisfy the following equations for every $A \subseteq P$:

$$\mathrm{Cl}_{\tau_P}(A) = \mathrm{Cl}(A) \cap P, \qquad \mathrm{Int}_{\tau_P}(A) = \mathrm{Int}((X \setminus P) \cup A) \cap P.$$

Update Modalities. The *dynamic language* $\mathcal{L}_{\mathbf{KB!}}$ is obtained by extending $\mathcal{L}_{KB}$ with *(existential) dynamic update modalities* $\langle !\varphi\rangle\psi$, meaning that: φ is true and after the agent learns this, ψ becomes true. The corresponding universal modality is defined by putting $[!\varphi]\psi := \neg\langle !\varphi\rangle\neg\psi$.

Definition 4. *(**Semantics of Updates**) Let $\mathcal{M} = (X, \tau, \nu)$ be a topological model. Given a formula φ we will denote by $\mathcal{M}_\varphi$ the **relativized model***

$$\mathcal{M}_\varphi = (\llbracket\varphi\rrbracket, \tau_{\llbracket\varphi\rrbracket}, \nu_{\llbracket\varphi\rrbracket}),$$

*where $\llbracket\varphi\rrbracket = \llbracket\varphi\rrbracket^{\mathcal{M}}$ is the interpretation of φ in $\mathcal{M}$, $\tau_{\llbracket\varphi\rrbracket}$ is the relativized topology and $\nu_P(p) = \nu(p) \cap P$, for each $p \in \mathsf{Prop}$. The **semantics of** $\mathcal{L}_{KB!}$ is obtained by extending the semantics of $\mathcal{L}_{KB}$ with the following clause:*

$$\llbracket\langle !\varphi\rangle\psi\rrbracket^{\mathcal{M}} = \llbracket\psi\rrbracket^{\mathcal{M}_\varphi}.$$

Connection to the defeasibility theory of knowledge. As promised in the Introduction, we show now that in the generalized Belief Revision Theory given by our topological semantics for conditional beliefs, *topological knowledge coincides with the one given by the defeasibility analysis*:

Theorem 4. *Let $\mathcal{M} = (X, \tau, \nu)$ be a topological model. The following are equivalent, for all worlds $x \in X$ and atomic sentences*[10] *p:*

1. $x \in \llbracket Kp\rrbracket^{\mathcal{M}}$;
2. $x \in \llbracket[!\theta]Bp\rrbracket^{\mathcal{M}}$ *for every formula* θ;
3. $x \in \llbracket Bp\rrbracket^{\mathcal{M}_\theta}$ *for every formula* θ *such that* $x \in \llbracket\theta\rrbracket^{\mathcal{M}}$.

[10] The restriction to atomic sentences in the other clauses is necessary because of the so-called Moore sentences: these are epistemic formulas which change their truth value after being learnt.

Proof. The equivalence between (2) and (3) follows immediately from the semantics of dynamic update modalities, so we only prove the equivalence between (1) and (3). For this, let us first put $A := [\![p]\!]^{\mathcal{M}}$, $P := X \setminus \mathrm{Int}(A)$, which gives us $X \setminus P = \mathrm{Int}(A)$. Then by the above equations we have $\mathrm{Int}_{\tau^P}(A) = \mathrm{Int}((X \setminus P) \cup A) \cap P = \mathrm{Int}(\mathrm{Int}(A) \cup A) \cap P = \mathrm{Int}(\mathrm{Int}(A)) \cap P = \mathrm{Int}(A) \cap (X \setminus \mathrm{Int}(A)) = \emptyset$.

To show (1) $\Rightarrow$ (3): assume that $x \in [\![Kp]\!]^{\mathcal{M}}$, so $x \in \mathrm{Int}([\![p]\!]^{\mathcal{M}})$. Let θ be any formula s.t. $x \in [\![\theta]\!]^{\mathcal{M}}$. Note that $\mathrm{Int}([\![p]\!]^{\mathcal{M}}) \subseteq \mathrm{Int}((X \setminus [\![\theta]\!]^{\mathcal{M}}) \cup [\![p]\!]^{\mathcal{M}})$, since $[\![p]\!]^{\mathcal{M}} \subseteq (X \setminus [\![\theta]\!]^{\mathcal{M}}) \cup [\![p]\!]^{\mathcal{M}}$. Then, since $x \in [\![\theta]\!]^{\mathcal{M}}$ and $[\![p]\!]^{\mathcal{M}_\theta} = [\![p]\!]^{\mathcal{M}} \cap [\![\theta]\!]^{\mathcal{M}}$, we have $x \in \mathrm{Int}(X \setminus [\![\theta]\!]^{\mathcal{M}} \cup [\![p]\!]^{\mathcal{M}}) \cap [\![\theta]\!]^{\mathcal{M}} = \mathrm{Int}((X \setminus [\![\theta]\!]^{\mathcal{M}} \cup [\![p]\!]^{\mathcal{M}}) \cap X) \cap [\![\theta]\!]^{\mathcal{M}} = \mathrm{Int}((X \setminus [\![\theta]\!]^{\mathcal{M}} \cup [\![p]\!]^{\mathcal{M}}) \cap (X \setminus [\![\theta]\!]^{\mathcal{M}} \cup [\![\theta]\!]^{\mathcal{M}})) \cap [\![\theta]\!]^{\mathcal{M}} = \mathrm{Int}(X \setminus [\![\theta]\!]^{\mathcal{M}} \cup ([\![p]\!]^{\mathcal{M}} \cap [\![\theta]\!]^{\mathcal{M}})) \cap [\![\theta]\!]^{\mathcal{M}} = \mathrm{Int}_{\tau_{[\![\theta]\!]^{\mathcal{M}}}}([\![p]\!]^{\mathcal{M}_\theta}) \subseteq \mathrm{Cl}_{\tau_{[\![\theta]\!]^{\mathcal{M}}}}(\mathrm{Int}_{\tau_{[\![\theta]\!]^{\mathcal{M}}}}([\![p]\!]^{\mathcal{M}_\theta})) = [\![Bp]\!]^{\mathcal{M}_\theta}$.

For (3) $\Rightarrow$ (1): assume that (3) holds but (1) fails, i.e $x \notin [\![Kp]\!]^{\mathcal{M}}$, and hence $x \in [\![\neg Kp]\!]^{\mathcal{M}}$. By applying (3) to the formula $\theta := \neg Kp$, we obtain that $x \in [\![Bp]\!]^{\mathcal{M}_\theta}$. But since $[\![\theta]\!]^{\mathcal{M}} = [\![\neg Kp]\!]^{\mathcal{M}} = X \setminus \mathrm{Int}(A) = P$, we have $x \in [\![Bp]\!]^{\mathcal{M}_\theta} = \mathrm{Cl}_{\tau_P}(\mathrm{Int}_{\tau_P}(A)) = \mathrm{Cl}_{\tau_P}(\emptyset) = \emptyset$. Contradiction!

5 Comparison with related work

We compare now our topological interpretation of belief with a different (and older) topological interpretation semantics that has been proposed for doxastic logic, using Cantor's derivative operator.

The Co-Derived Set Semantics for Belief. Let (X, τ) be a topological space. We recall that a point x is called a *limit point* (limit points are also called *accumulation points*) of a set $A \subseteq X$ if for each open neighbourhood U of x we have $(U \setminus \{x\}) \cap A \neq \emptyset$. Let $d(A)$ denote the set of all limit points of A. This set is called the *derived set* and d is called the *derived set operator*. For each $A \subseteq X$ we let $t(A) = X \setminus d(X \setminus A)$. We call t the *co-derived set operator*. Also recall that there is a close connection between the derived and co-derived set operators and the closure and interior operators. In particular, for each $A \subseteq X$ we have $\mathrm{Cl}(A) = A \cup d(A)$ and $\mathrm{Int}(A) = A \cap t(A)$. Unlike the closure operator there may exist elements of A that are not its limit points. In other words, in general $A \not\subseteq d(A)$. Also note that for each $x \in X$ we have $x \notin d(x)$, where $d(x)$ is a shorthand for $d(\{x\})$.

Definition 5. *Let $\mathcal{M} = (X, \tau, \nu)$ be a topological model. The* ***co-derived set semantics for*** $\mathcal{L}_{KB}$ *is obtained by extending the standard topological semantics for $\mathcal{L}_K$ (interpreting K as interior) with the following clause:*

$$[\![B\varphi]\!]^{\mathcal{M}} = t([\![\varphi]\!]^{\mathcal{M}})$$

This immediately gives us that $[\![\langle B\rangle\varphi]\!]^{\mathcal{M}} = d([\![\varphi]\!]^{\mathcal{M}})$. We again skip the index $\mathcal{M}$ if it is clear from the context. See [3], [20] and [7] for an overview of the results on the co-derived set semantics. Here we only mention the completeness results for the unimodal language $\mathcal{L}_B$ with the co-derived set semantics: the complete logic of belief over all topological spaces is $wK4 = K + ((p \wedge Bp) \rightarrow BBp)$ [12], while the doxastic logic $KD45$ is complete wrt so-called DSO-spaces. Here, a

DSO-space is a topological space (X, τ) satisfying the following conditions: the T_D-separation axiom[11]; for every $A \subseteq X$ the set $d(A)$ is open; and (X, τ) is dense-in-itself, i.e., $d(X) = X$. See [23] for more details.

Criticism and comparison with our conception. Steinsvold [23] was the first to propose the co-derived set operator as a semantics for belief. However, this interpretation has a major disadvantage: *it entails (not just the possibility, but) the necessity of error*. To explain: all authors agree that one of the main characteristics of belief is the *possibility of error*: it is possible that some of the agent's beliefs are false. In other words, any good semantics for belief should allow for models and worlds at which some beliefs are false. However, we claim that, according to the co-derived semantics *the existence of false beliefs is a necessary fact* (holding for all possible agents at all possible worlds in all possible models!).

Indeed, as we pointed out above, for each $x \in X$ we always have $x \notin d(x)$. So $x \in B(X \setminus \{x\})$. Thus, at a point x the agent believes $X \setminus \{x\}$, which is false (since $x \notin X \setminus \{x\}$). This means that in *any topological model and any world in this model, there is at least one false belief.* Hence, the co-derived set interpretation implies the "necessity of error": *the actual world is always dis-believed.*

We think this consequence is an intuitively undesirable property. It generally prevents any act of learning (updating with) the actual world. Indeed, the main problem of Formal Learning Theory (learning the true world, or the correct possibility, from a given set of possibilities) becomes automatically unattainable. Similarly, the physicist's dream of finding a true "theory of everything" is declared impossible by fiat, *as a matter of logic.* More importantly, even if necessity of error might seem realistic within a Lewisian "large-world interpretation" of possible-world semantics (in which each world must really come with a full description of all the myriad of ontic facts of the world), this property seems completely unrealistic when we adopt the more down-to-earth "small-world" models that are common in Computer Science, Game theory and other applications. In these fields, the "worlds" in any usable model come only with the description of the facts that are *relevant* for the problem at hand: e.g. in a scenario involving the throwing of a fair coin, the relevant fact is the upper face of the coin. A model for this scenario will involve typically only two possible worlds: Head and Tail. Requiring that the agent must always have a false belief means in this context that the agent can never find out which of the coin's faces is the upper one: an obviously absurd conclusion!

There is another objection, maybe even more decisive, against the co-derived set semantics, namely that it can be easily "Gettierized". As mentioned above, we have $\mathrm{Int}(A) = A \cap t(A)$, which means that *in the co-derived set interpretation, knowledge is exactly the same as true (justified) belief.* So this semantics is easily vulnerable to all the well-known Gettier-type counterexamples!

Finally, here is an argument of a more technical nature. As mentioned above, the co-derived set semantics validates the $KD45$ axioms only on DSO-spaces, while our semantics validates them on extremally disconnected spaces. So the

[11] Recall that the T_D separation axiom states that every point is the intersection of a closed and open set. This condition is equivalent to $d(d(A)) \subseteq d(A)$, see e.g., [11].

following result shows that *our topological interpretation "works" on a larger class of models than the co-derived set semantics*:

Proposition 6. *Every DSO-space is extremally disconnected.*

Proof. Let (X, τ) be a *DSO*-space and $U \in \tau$. Recall that for any $A \subseteq X$, $\mathrm{Cl}(A) = d(A) \cup A$. So $\mathrm{Cl}(U) = d(U) \cup U$. Since (X, τ) is a *DSO*-space, $d(U)$ is an open subset of X. Thus, since U is open as well, $d(U) \cup U = \mathrm{Cl}(U)$ is open.

6 Conclusions and future work

In this paper, we proposed a new topological semantics for belief and argued that it is the "correct" one, at least as far as full belief (understood as subjective certainty) is concerned: it is *the "canonical" (most general) semantics for (Stalnaker's axioms for) full belief*. Moreover, our proposal comes with an *independent motivation* and has an *intrinsic philosophical and intuitive value*. Topologically, a point is in the interior of a set P iff it can be sharply distinguished (separated) from all non-P points (by an open set); similarly, a point is in the closure of P iff it is "very close" to P, i.e. it cannot be sharply distinguished from all P points. Thus, an agent knows P if she can sharply distinguish the actual world from all non-P-worlds. Hence, according to our semantics for full belief, an agent (fully) believes P if she cannot sharply distinguish the actual world w from the worlds in which she has knowledge of P. In this sense, one can say that *belief is topologically "very close" to knowledge*: indeed, the agent cannot sharply distinguish it from knowledge. We thus think that our topological semantics perfectly captures the essence of full belief as "subjective certainty".

From a philosophical perspective, the main importance of our paper is that *it connects three different epistemological conceptions* that were proposed as responses to Gettier's challenge: Stalnaker's epistemic definition of full belief (in the spirit of the "knowledge-first" approach), the "knowledge as correctly-justified-belief" approach (underlying the topological semantics of knowledge) and the defeasibility analysis of knowledge. Indeed (as shown in Section 4), in our semantics knowledge is undefeated-true-belief. To show this, we needed the topological analogue of dynamic conditioning (update), as it was already defined in [27].

In on-going work, we also explore the corresponding "static" conditioning, by giving a topological semantics for *conditional belief*; we investigate the properties of the resulting topological belief revision, showing that it satisfies only the AGM Postulates 1-6 (but not postulates 7 and 8). In the same work, we give a complete axiomatization of the logic of conditional beliefs (with the topological semantics), as well as a complete axiomatization of the corresponding dynamic logic (obtained by adding dynamic update operators, as in Section 4). We plan to present these results (as well as the proofs that are missing from the current paper) in a future journal publication.

References

1. J. Barwise. Three views of common knowledge. In *Proceedings of the second Conference on Theoretical Aspects of Reasoning about Knolwedge*, pages 365–379, 1988.
2. P. Battigalli and M. Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 105:356–391, 2002.
3. J. van Benthem and G. Bezhanishvili. Modal logics of space. In *Handbook of spatial logics*, pages 217–298. Springer, Dordrecht, 2007.
4. J. van Benthem, G. Bezhanishvilli, B. ten Cate, and D. Sarenac. Modal logics for products of topologies. *Studia Logica*, 84(3)(369-392), 2005.
5. J. van Benthem and E. Pacuit. Dynamic logics of evidence-based beliefs. *Studia Logica*, 99(1):61–92, 2011.
6. J. van Benthem and D. Sarenac. The geometry of knowledge. *Aspects of universal Logic*, 17, 2005.
7. N. Bezhanishvili and W. van der Hoek. Structures for epistemic logic. In A. Baltag and S. Smets, editors, *Logical and Informational Dynamics. A volume in honour of Johan van Benthem.* Trends in Logic, Springer. To appear.
8. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Cambridge University Press, 2001.
9. A. Chagrov and M. Zakharyaschev. *Modal Logic.* The Clarendon Press, 1997.
10. J.M. Dunn. Generalized ortho negation. In H. Wansing, editor, *Negation: A Notion in Focus (Perspectives in Analytical Philosophy, Bd 7)*, Perspectives in analytical philosophy, pages 3–26. de Gruyter, 1996.
11. R. Engelking. *General topology*, volume 6. Heldermann Verlag, Berlin, second edition, 1989.
12. L. Esakia. Weak transitivity-restitution. *Study in logic*, 8:244–254, 2001. In Russian.
13. B. van Fraassen. Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical logic*, 24:349–377, 1995.
14. E. Gettier. Is justified true belief knowledge? *Analysis*, 23:121–123, 1963.
15. K. Kelly. *The Logic of Reliable Inquiry.* Oxford University Press, 1996.
16. P. Klein. A proposed definition of propositional knowledge. *Journal of Philosophy*, 68:471–482, 1971.
17. K. Lehrer. *Theory of Knowledge.* Westview Press, 2000.
18. K. Lehrer and T. Jr Paxson. Knowledge: Undefeated justified true belief. *Journal of Philosophy*, 66:225–237, 1969.
19. J. C. C. McKinsey and A. Tarski. The algebra of topology. *Ann. of Math. (2)*, 45:141–191, 1944.
20. R. Parikh, L. Moss, and C. Steinsvold. Topology and epistemic logic. In *Handbook of spatial logics*, pages 299–341. Springer, Dordrecht, 2007.
21. R. Sikorski. *Boolean Algebras.* Springer-Verlag, Berlin-Heidelberg-Newyork, 1964.
22. R. Stalnaker. On logics of knowledge and belief. *Philosophical Studies*, 128(1):169–199, 2006.
23. C. Steinsvold. *Topological models of belief logics.* PhD thesis, New York, NY, USA, 2006.
24. S. Vickers. *Topology via logic*, volume 5 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, Cambridge, 1989.
25. F.P.J.M. Voorbraak. *As Far as I Know.* PhD thesis, Utrecht University, 1993.
26. T. Williamson. *Knowledge and its Limits.* Oxford Univ. Press, 2000.
27. J. Zvesper. *Playing with Information.* PhD thesis, University of Amsterdam ILLC PhD Thesis, 2010.