

Who is Afraid of Minimal Revision?

Edoardo Baccini	Zoé Christoff	Nina Gierasimczuk	Rineke Verbrugge
University of Groningen	University of Groningen	Technical University of Denmark	University of Groningen
e.baccini@rug.nl	z.l.christoff@rug.nl	nigi@dtu.dk	l.c.verbrugge@rug.nl

The principle of minimal change in belief revision theory requires that, when accepting new information, one keeps one’s belief state as close to the initial belief state as possible. This is precisely what the method known as minimal revision does. However, unlike less conservative belief revision methods, minimal revision falls short in learning power: It cannot learn everything that can be learned by other learning methods. We begin by showing that, despite this limitation, minimal revision is still a successful learning method in a wide range of situations. Firstly, it can learn any problem that is finitely identifiable. Secondly, it can learn with positive and negative data, as long as one considers finitely many possibilities. We then characterize the prior plausibility assignments (over finitely many possibilities) that enable one to learn via minimal revision, and do the same for conditioning and lexicographic upgrade. Finally, we show that not all of our results still hold when learning from possibly erroneous information.

1 Introduction

Many human activities aim to find the truth. Consider, for instance, a scientist deciding which of two possible theories, T_1 or T_2 , is the correct one. At the start of their inquiry, the scientist might have some *prior plausibility ordering* that reflects their preference among the two. Over time, new experimental results and observations become available to the scientist, who revises their assessment accordingly using some *revision method*. How should the scientist revise their prior assessment? Some inspiration about the structure of that dynamics can be drawn from belief revision theory. In particular, one natural way to revise one’s beliefs is to accept the newly learned information but change the overall belief state as little as possible. This *minimal revision* method (also known as “conservative upgrade”) was first introduced in [13], where it was shown that the method is consistent with the basic AGM principles of (iterated) belief revision [1], that it naturally accommodates the Ramsey test for conditionals [31], and that, while continually updating the propositional beliefs, it preserves as many as possible of the extended, conditional beliefs.

Despite its natural appeal, minimal revision falls short in its *learning power*, i.e., in its ability to successfully lead an agent to identify the true state of the world among a set of possibilities. In particular, using tools from formal learning theory, [23] and [17, 4, 6] show that, unlike the more radical belief revision methods of *conditioning* (also known as “update”) and *lexicographic revision* (also known as “radical upgrade”), it does not excel as a learning method, because it suffers from memory limitations [17] and has a tendency to fall into cycles [7]. Given these negative results, other revision methods appear *in general* preferable to minimal revision to track the truth. However, the much simpler minimal revision could still be the right method when considering *particular classes* of learning scenarios. Building on the framework introduced in [17, 4, 6], we show that this is indeed the case.

After recalling the framework from [6] in Section 2, we show in Section 3 that: Minimal revision can successfully learn any problem that is finitely identifiable (Section 3.1); It can learn on positive and negative data if the possibilities considered are finite (Section 3.2); Only some prior plausibility

assessments can lead minimal revision to success, and we characterize the class of such plausibility assessments when one considers finitely many possibilities (Section 3.3); Not all our results are preserved when learning by minimal revision with possibly erroneous observations (Section 3.4). In Section 4, we consider conditioning and lexicographic revision and characterize the prior plausibility assessments which allow one to successfully learn via these methods. We conclude in Section 5.

2 Background and Preliminary Definitions

We consider a discrete-time iterative process ordered like the natural numbers: At each time-step, an agent receives an observation from the world, on the basis of which they revise their plausibility order and make a conjecture about which state is the actual one. The aim of the agent is to correctly identify the actual state of the world. Below, we recall the framework developed in [17, 4, 6].

Definition 1 ([6, Def. 1]). An *epistemic space* S is a pair $\langle S, O \rangle$, where S is a set of possible worlds (at most countable), and $O \subseteq \mathcal{P}(S)$ is a set of observables (at most countable). For every $s \in S$, O_s denotes the set of observables $p \in O$ such that $s \in p$. For any two worlds $s, t \in S$, if $O_s = O_t$ then $s = t$ (no two worlds satisfy exactly the same set of observables). An epistemic space is said to be *finite* if $|S| < |\mathbb{N}|$.

The set of states represents the possibilities entertained by an agent, e.g., theories a scientist considers possible. The set of observables represents the observations the agent might come across. Figure 1 provides an example of an epistemic space (on the left). To represent an agent's doxastic state, we enrich the space with a plausibility relation to obtain a plausibility space (Figure 1, right) [21, 13, 8].

Definition 2 ([6, Def. 2]). A *plausibility space* B is the tuple $\langle S, \preceq \rangle$ where $S = \langle S, O \rangle$ is an epistemic space and $\preceq$ is a total preorder, i.e., a total, transitive and reflexive relation over S .

For $s, t \in S$, $s \preceq t$ means that “ s is at least as plausible as t ”. We write $s \prec t$ if $s \preceq t$ and $t \not\preceq s$; $s \simeq t$ if $s \preceq t$ and $t \preceq s$. Since in Section 4 we will also consider preorders that are not necessarily total, we define additional notation: $s \sim t$ if $s \preceq t$ or $t \preceq s$; $s \not\sim t$ if not $s \sim t$.

The assumption that the preorder be total is common in the dynamic epistemic logic literature and is rooted in Grove's semantics [21], inspired in turn by Lewis' sphere system for counterfactuals [26]. Plausibility spaces allow for a simple representation of belief as truth in the most plausible worlds [8]: An agent believes a proposition p iff p is true in all states that the agent considers most plausible.

Definition 3. Let a preorder $\preceq$ (not necessarily total) over a set of elements S be given. An element $s \in S$ is a $\preceq$ -*minimal element* of S iff for all $t \in S$ such that $t \sim s$, $s \preceq t$. We denote by $\min_{\preceq} X$ the set of $\preceq$ -minimal elements of the set $X \subseteq S$. For every $p \in O$, we define the preorders $\preceq_p := \preceq \cap (p \times p)$ and $\preceq_{\bar{p}} := \preceq \cap (S \setminus p \times S \setminus p)$.

The observations that an agent receives (in a step-wise manner) form an infinite stream:

Definition 4 ([6, Def. 4, 5]). A *data sequence* σ is a finite sequence of elements of O ; a *data stream* $\vec{O}$ is an infinite sequence of elements in O . Given a data sequence $\sigma = (O_0, O_1, \dots, O_n)$, we denote by $\text{set}(\sigma)$ the set of elements of O that occur in σ ; given a data stream $\vec{O} = (O_0, O_1, \dots)$, we let $\text{set}(\vec{O})$ be the set of elements of O that occur in $\vec{O}$. Given $n \in \mathbb{N}$, we denote by O_n the n^{th} observation in $\vec{O}$. $\vec{O}[k]$ is the initial segment of $\vec{O}$ of length k . $\sigma * \vec{O}$ is the concatenation of the sequence σ with the data stream $\vec{O}$. The empty sequence is denoted by ε .

We consider streams containing all and only observations that are true in the actual world.

Definition 5 ([6, Defs. 6, 7]). A *data sequence* σ is: *sound with respect to a state* s iff $\text{set}(\sigma) \subseteq O_s$; *complete with respect to a state* s iff $O_s \subseteq \text{set}(\sigma)$. A *data stream* $\vec{O}$ is: *sound with respect to a state* s iff $\text{set}(\vec{O}) \subseteq O_s$; *complete with respect to a state* s iff $O_s \subseteq \text{set}(\vec{O})$.

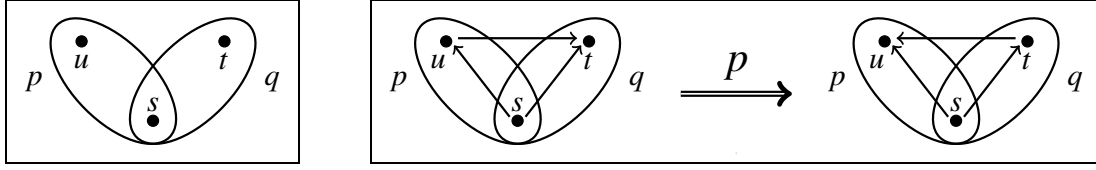


Figure 1: On the left: An epistemic space with $S = \{u, s, t\}$, $O = \{p, q\}$. On the right: a plausibility order on the same space with $t < u < s$ and the revised order $u < t < s$ after observing p (represented by the p -labelled arrow). This is an example of (one-step) *minimal revision* (Def. 12). An arrow from one world to another indicates that the latter is more plausible than the former. We omit reflexive arrows.

At each time-step, a learner uses some belief revision method to revise their plausibility order on the basis of the new observation received. Formally:

Definition 6 ([6, Def. 10]). A **one-step belief revision** is a function R_1 taking a plausibility space $\mathbf{B}$ and an observable $p \in O$ to output the plausibility space $R_1(\mathbf{B}, p)$. A one-step belief revision function R_1 induces an **iterated belief revision function** R in the following way: $R(\mathbf{B}, \varepsilon) = \mathbf{B}$, $R(\mathbf{B}, \sigma * p) = R_1(R(\mathbf{B}, \sigma), p)$.

Any belief-revision method can be used to define a corresponding learning method.

Definition 7 ([6, Def. 8]). A **learning method** L is a function that takes as input an epistemic space $S = \langle S, O \rangle$ and a data sequence σ and outputs a conjecture $L(S, \sigma) \subseteq S$.

Definition 8 ([6, Def. 11]). A **plausibility assignment** $\mathbf{PLAUS}$ is a function that takes as input an epistemic space $S = \langle S, O \rangle$ and assigns to it a plausibility order $\preceq$, converting it into a plausibility space $\mathbf{PLAUS}(S) = \langle S, \preceq \rangle$.

Definition 9 ([6, Def. 12]). Let a belief-revision method R be given. Given any epistemic space S and plausibility assignment $\mathbf{PLAUS}$, a belief-revision method R induces a **canonical learning method** in the following way: $L_R^{\mathbf{PLAUS}}(S, \sigma) := \min R(\mathbf{PLAUS}(S), \sigma)$.

Given an epistemic space S , a plausibility assignment $\mathbf{PLAUS}$, a one-step belief-revision method R_1 , and $\vec{O}[n]$, we denote with $\preceq_{\vec{O}[n]}^{R_1}$ the plausibility order in the updated epistemic space $R(\mathbf{PLAUS}(S), \vec{O}[n])$.

As the central criterion of successful learning we take *identifiability in the limit*, which originally comes from the studies of inductive grammar inference [20], and is often combined with computational restrictions on learners and the classes they learn. In this paper, we abstract away from such constraints and focus on purely structural considerations.

Definition 10 ([6, Def. 18], originally in [20]). Let an epistemic space $S = \langle S, O \rangle$ be given. A state $s \in S$ is **identified in the limit by a learning method** L iff for any sound and complete data stream $\vec{O}$ for s , there is an $n \in \mathbb{N}$ such that for all $k \geq n$, $L(S, \vec{O}[k]) = \{s\}$. The **epistemic space** S is **identified in the limit by learning method** L iff all $s \in S$ are identified in the limit by learning method L . The **epistemic space** S is **identifiable in the limit (learnable)** iff it is identified in the limit by some learning method.

Instead of such general, unrestricted learning functions, we will focus on more constructive ones, namely, those guided by some underlying belief revision methods.

Definition 11 ([6, Def. 19]). An **epistemic space** S is **identifiable in the limit by the belief revision method** R if there exists a prior plausibility assignment $\mathbf{PLAUS}$ such that $L_R^{\mathbf{PLAUS}}$ identifies S in the limit. A **belief revision method** R is **universal on a class** C of epistemic spaces if every space in C is identifiable by the belief revision method. A **belief revision method** R is **universal (tout court)** if it is universal in the class of all learnable spaces.

3 Learning with Minimal Revision

What is the least disruptive way of revising plausibility spaces? Such a minimal policy was first discussed in [30] and then generalised to total preorders in [13]. Upon receiving p , minimal revision moves all the most plausible p -worlds in front of all other worlds, leaving the rest of the plausibility order unchanged. Minimal revision is *conservative* in the sense that it leaves as much as possible of the old plausibility ordering unchanged [13, 17].¹ In our setting, it is defined in the following way.

Definition 12 (Minimal Revision, [8]). *Let $\mathbf{B} = \langle S, O, \preceq \rangle$ be a plausibility space and $p \in O$. The one-step revision function $\mathbf{mini}(\mathbf{B}, p)$ generates the plausibility space $\mathbf{B}' = \langle S, O, \preceq' \rangle$, where $\preceq'$ is such that if $s \in \min_{\preceq} p$ and $t \notin \min_{\preceq} p$, then $s \preceq' t$ and $t \not\preceq' s$; $s \preceq' t$ iff $s \preceq t$, otherwise.*

Minimal revision suffers from a number of shortcomings. For instance, it can fail to stabilize when revising with true higher-order information [7], and it fails to learn the least complex problems in the learning setting of [23]. With respect to the framework adopted in this paper, it has been shown that minimal revision is not universal, in the sense that not all spaces that are learnable can be learned via minimal revision [17, 4, 6]. Consider the space in Figure 1 from [6]. First, note that the space is learnable by the cruder *conditioning* method (Definition 19): At each time-step, eliminate all worlds at which the new information is false. In this case, any initial plausibility order would do, as long as $u \prec s$ and $t \prec s$. Yet, there is no plausibility assignment allowing identification of this space in the limit via minimal revision: In order to identify either u or t , both of them must be strictly more plausible than s ; but then s is not identified, since no observation will ever suffice to make s strictly more plausible than u and t .

We are now ready to state our main research question: Even if minimal revision cannot identify all learnable epistemic spaces, are there interesting classes of epistemic spaces that it can identify? In the sections below, we give a positive answer to this question and consider some notable classes of spaces in which minimal revision is universal. All omitted proofs are included in the Appendix.

3.1 Minimal Revision is Universal on the Class of Finitely Identifiable Spaces

The success criterion of identifiability in the limit in Definition 10 is often contrasted with a stronger criterion of conclusive learning (see [19, 14]), adapted from finite identifiability as proposed in [16, 24, 22]. Finite identifiability was characterized in [27], studied, e.g., in [25, 32], and applied to problems in AI, e.g., in [11, 12, 29]. It requires not only that the learner *converges* to a correct conjecture, but also that they know with certainty that the conjecture is true. The latter condition of certainty can be rendered as the requirement that for any stream the learner is a once-defined function, i.e., throughout the learning process it only has one-shot at the correct guess [19]. In this section, we will show that **mini** can identify in the limit the class of epistemic spaces that can be finitely identified, i.e., those spaces for which there exists a learner that can identify the actual world with a single guess.

Definition 13 ([27]). *Let an epistemic space $S = \langle S, O \rangle$ be given. A state $s \in S$ is **finitely identified** by a learning method L iff for any sound and complete data stream $\vec{O}$ for s , when inductively given $\vec{O}$, L outputs at some point a single conjecture s . The **epistemic space S is finitely identified** by learning method L iff all $s \in S$ are finitely identified by learning method L . The **epistemic space S is finitely identifiable (conclusively learnable)** just in case it is finitely identified by some learning method.*

In [27], in a computable setting, it has been shown that this class of spaces is characterized by the existence of the following map:

¹For a systematic analysis of the learning-theoretic properties of the canonical learning methods induced by minimal revision and other belief revision methods, see [17].

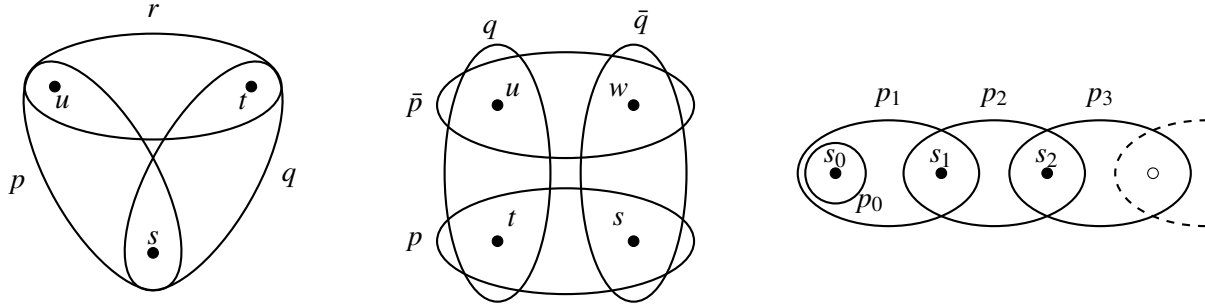


Figure 2: Example of three epistemic spaces that are finitely identifiable. On the left: $S = \{s, u, t\}$, $O = \{p, q, r\}$. At the center: $S = \{s, u, t, w\}$, $O = \{p, \bar{p}, q, \bar{q}\}$. On the right: $S = \{s_i | i \in \mathbb{N}\}$, $O = \{p_i : i \in \mathbb{N}\}$, and for all s_i , $O_i = \{p_j \in O : j = i \text{ or } j = i + 1\}$.

Definition 14 ([18, Def. 10.8], originally in [27]). Let $S = \langle S, O \rangle$ be an epistemic space, and let $s \in S$ be given. A **definite finite tell-tale map** is a total map assigning to each $s \in S$ a **definite finite tell-tale set** $D_s \subseteq O$ such that: (i) D_s is finite; (ii) $D_s \subseteq O_s$; (iii) for any $t \in S$, if $D_s \subseteq O_t$, then $s = t$.

Informally, the existence of a definite finite tell-tale map on a space means that each world is uniquely identified by some finite conjunction of observables. Fig. 2 shows examples of spaces with definite finite tell-tale maps. The characterisation below adapts to our setting the original theorem in [27].

Proposition 1. An epistemic space S is finitely identifiable iff it has a definite finite tell-tale map.

We show now that **mini** can learn the class of finitely identifiable spaces, i.e., it can identify this class in the limit. We do this by showing that **mini** can learn any space that has a definite finite tell-tale map. We start with the following Lemma:

Lemma 1. Let an epistemic space $S = \langle S, O \rangle$ be given. Let **PLAUS** be the plausibility assignment that assigns to S the preorder $\preceq$ such that for all $s, t \in S$, $s \preceq t$. Then, for any state s and for any sound data-stream $\vec{O}$ for s , $s \in L_{\text{mini}}^{\text{PLAUS}}(S, \vec{O}[n])$ for all $n \in \mathbb{N}$.

Proof. We prove by induction on n that $s \in \min_{\preceq_{\vec{O}[n]}}^{\text{mini}}$. [Base Case: $n = 0$] The sequence $\vec{O}[0]$ is empty, and as such the plausibility model we consider is the initial plausibility model. Since for all $t \in S$, $s \preceq t$, we have that s is a minimal element of $\preceq$. [Inductive Step] Assume as inductive hypothesis that for all $m \leq n$, it is the case that $s \in \min_{\preceq_{\vec{O}[m]}}^{\text{mini}}$. Consider now $n + 1$, and consider the observable p occurring in the position $n + 1$ in the stream $\vec{O}$, i.e., the observable after the sequence $\vec{O}[n]$. By assumption, $\vec{O}$ is sound with respect to s , and thus $s \in p$. In addition, by inductive hypothesis we have that $s \in \min_{\preceq_{\vec{O}[n]}}^{\text{mini}} p$. Hence, $s \in \min_{\preceq_{\vec{O}[n+1]}}^{\text{mini}}$ by Def. 12. Therefore, in any sound and complete data-stream $\vec{O}$ for all $n \in \mathbb{N}$, $s \in \min_{\preceq_{\vec{O}[n]}}^{\text{mini}}$, and thus $s \in L_{\text{mini}}^{\text{PLAUS}}(S, \vec{O}[n])$. $\square$

Theorem 1. Minimal revision is universal on the class of spaces that are finitely identifiable.

Proof. Let S be an epistemic space that is finitely identifiable. Then, by Prop. 1, S has a definite finite tell-tale map. Let the plausibility order $\preceq$ be given such that for all $t, s \in S$, $s \preceq t$. Assume towards a contradiction that the learner induced by **mini** and $\preceq$ does not identify the space S , i.e., there is a state $s \in S$ and a sound and complete data-stream $\vec{O}$ for s such that there is no $n \in \mathbb{N}$ such that for all $k \geq n$, $L_{\text{mini}}^{\text{PLAUS}}(S, \vec{O}[k]) = \{s\}$. Consider such a data-stream. Since $\vec{O}$ is sound and complete with respect to s ,

and since there exists a definite finite tell-tale set D_s for s , there is some $n \in \mathbb{N}$ such that $D_s \subseteq \text{set}(\vec{O}[n])$. Consider such a sequence $\vec{O}[n]$. Since s is not identifiable in the limit, there exists a $k \geq n$, such that $L_{\text{mini}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) \neq \{s\}$. By Lemma 1, it must be the case that $s \in L_{\text{mini}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. Therefore, there must exist a $t \neq s$ such that $t \in L_{\text{mini}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. It follows that $t \in \bigcap \text{set}(\vec{O}[k])$ (otherwise, it would have become strictly less plausible than s for some $k' \leq k$). Since $D_s \subseteq \text{set}(\vec{O}[k])$, we have $t \in \bigcap D_s$. By Def. 14, we have $t = s$. Contradiction. $\square$

The theorem above states that any space that is finitely identifiable can be learned via **mini** by using an initially ‘unbiased’ plausibility order, which considers all possibilities equally plausible. This plausibility assignment together with **mini** induces a method that patiently waits for enough data to dismiss all but one possibility. Note, however, that starting from a prior that ranks all states as equally plausible is not necessary for learning finitely identifiable spaces via **mini**, i.e., there are finitely learnable spaces that can be learned via **mini** and a plausibility assignment where for at least two states s, t , $s \prec t$. For instance, the space on the right in Figure 2 can be identified by ranking the states in a way such that $s_i \preceq s_j$ iff $i \leq j$.

Let us also note that **mini** can identify spaces that are not finitely identifiable and thus the class of finitely identifiable epistemic spaces is properly included in the class of spaces on which **mini** is universal. For instance, the epistemic space with $S = \{s, t\}$, $O = \{p, q\}$, $O_s = \{p, q\}$ and $O_t = \{q\}$ is a space identifiable via **mini** using the plausibility order $t \prec s$. This space is not finitely identifiable since $O_t \subseteq O_s$ and therefore $t \in S$ has no definite finite tell-tale set.

3.2 Minimal Revision is Universal on Positive and Negative Data on Finite Spaces

In the inductive inference literature, it is common to distinguish between learning from only positive information and from positive and negative information. This distinction is given a lot of attention, because the latter setting is much more informative, and so it leads to different learning powers of the same learning methods (see, e.g., [3]). In our case, learning from positive and negative information means that for every observable its complement can also be observed. In logic parlance, this is expressed as closure under negation (i.e., if p can be observed, so can $\neg p$). We will therefore speak of epistemic spaces that are closed under negation [6]. Negation-closed spaces are a subclass of the class of *strongly separated* spaces [5].

Definition 15. An epistemic space $\mathbf{S} = \langle S, O \rangle$ is: **strongly separated** when for every $s, t \in S$, $O_s \not\subseteq O_t$; **negation-closed** iff for every $p \in O$, there is $\bar{p} \in O$ such that $p = S \setminus \bar{p}$.

Figure 2 shows an example of a strongly separated space that is not negation-closed (left) and an example of a negation-closed space (center). From [17], we know that **mini** is not universal on positive and negative data on the class of all negation-closed spaces, and thus the same holds for the larger class of strongly separated spaces. Nevertheless, we can show that **mini** is in fact universal on the class of *finite* strongly separated spaces and thus on the class of *finite* negation-closed spaces.

Lemma 2. Let $\mathbf{S}$ be a finite strongly separated epistemic space. Then $\mathbf{S}$ has a definite finite tell-tale map.

Proof. Consider a strongly separated epistemic space $\mathbf{S}$. Take an arbitrary $s \in S$. Consider the set $D^s := \bigcup_{s' \in S \setminus \{s\}} (O_s \setminus O_{s'})$. Since the space is strongly separated, for all $s' \in S$ such that $s' \neq s$, $O_s \setminus O_{s'} \neq \emptyset$. Thus for all $s' \neq s$, $s' \notin \bigcap D^s$. Since the space is finite, D^s is finite, and $D^s \subseteq O_s$ by construction. By Def. 14, D^s is a definite finite tell-tale set for s . Since s was arbitrary, $\mathbf{S}$ has a definite finite tell-tale map. $\square$

Theorem 2. Minimal revision is universal on the class of finite strongly separated epistemic spaces.

Proof. By Lemma 2 and Prop. 1, the class of finite strongly separated epistemic spaces is a subset of the class of finitely identifiable spaces. By Th. 1, **mini** is universal on the class of finitely identifiable spaces, and thus on the class of finite strongly separated spaces. $\square$

Corollary 1. *Minimal revision is universal on positive and negative data on the class of finite negation-closed epistemic spaces.*

Therefore, one can safely learn from positive and negative data using **mini** when only finitely many states are possible. Let us note here that this result has been foreshadowed by Boutilier in his seminal work [13], where minimal revision is applied to finitely-grounded knowledge sets over propositional logic (which is closed on negation).

3.3 What Priors for Minimal Revision?

We have considered above some classes of spaces where **mini** is universal. However, in those and all other classes on which **mini** is universal, only some plausibility orders will allow learning via **mini**. We call these orders *appropriate*:

Definition 16. *Let S be an epistemic space. A plausibility order $\preceq$ is appropriate for learning S via the belief revision method R if the canonical learning method induced by R and the plausibility assignment assigning $\preceq$ to S identifies S in the limit.*

Figure 3 shows a plausibility order that is not appropriate for learning a space via **mini**: The state s (in red) is not identified with the (sound and complete) data-stream in which q occurs once first, followed by an infinite stream of ps . With this stream, the state u will forever remain equi-plausible to s , and thus s will never be the uniquely minimal element of the space.

Below, we give a characterisation of the class of appropriate plausibility orders for learning via **mini** on *finite* epistemic spaces. Our focus on finite spaces is justified by the fact that they are often sufficient for a variety of applications in belief revision theory, knowledge representation, and dynamic epistemic logic (where possible worlds are identified with valuations over a finite set of propositions, see, e.g., [15]). Furthermore, the full characterisation of the class of spaces for which **mini** is universal is still unknown, so we cannot resort to knowledge about the underlying structure of the epistemic space in our analysis.

Let us remark that our focus here is different from that in [6], where the focus is on whether a belief revision method is universal in the sense that for any learnable space one can construct *some* plausibility order that guarantees successful learning via that method. The interest in [6] therefore lies in the existence of a suitable plausibility order on every learnable space. In contrast, we aim at characterising for **mini** (and later in Section 4 also for conditioning and lexicographic upgrade) the class of *all* prior plausibility orders that are suited for learning every space that is learnable via the method under consideration. Of course, an appropriate order for a method can only exist on spaces that can be learned via that method, and thus [6] provides the limits of applicability of our analysis. Let us start with some preliminary lemmas.

Lemma 3. *If a plausibility order $\preceq$ is appropriate to learn an epistemic space S via **mini**, then for all $s \in S$, there exists a $p \in O_s$ such that $s \in \min_{\preceq} p$.*

Proof. Assume that $\preceq$ is appropriate to learn S via **mini**. Assume towards a contradiction that there exists an s such that for all $p \in O_s$, $s \notin \min_{\preceq} p$. Consider now a sound and complete data stream such that it enumerates all the elements in O_s . We show by induction on the length n of any finite segment of $\vec{O}[n]$

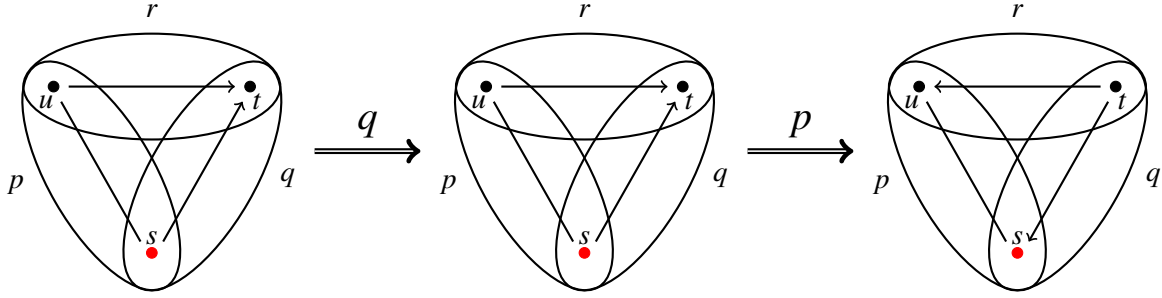


Figure 3: Example of a plausibility order that is not appropriate to learn a space via **mini**. The epistemic space is given by $S = \{s, u, t\}$ and $O = \{p, q, r\}$; the prior plausibility order (leftmost space) is $t \prec s$, $t \prec u$, $u \simeq s$. The labelled arrows represent two revision steps, first when observing q , and then p .

that for all $p \in O_s$ there exists a t such that $t \prec_{\vec{O}[n]}^{\text{mini}} s$ and $t \in p$. [Base case $n = 0$] Follows directly from the assumption. [Inductive step: $n = m + 1$] Assume that for $O[m]$, for all $p \in O_s$ there is a t such that $t \in p$ and $t \prec_{O[m]}^{\text{mini}} s$. Consider now $p \in O_s$ that occurs in $\vec{O}$ after $O[m]$. By inductive hypothesis, it follows that there is a $t \in S$ such that $t \prec_{O[m]}^{\text{mini}} s$ and $t \in p$. Hence $t \prec_{O[m+1]}^{\text{mini}} s$. Hence $s \notin \min_{\prec_{O[m+1]}^{\text{mini}}}$. $\square$

The above lemma expresses that an appropriate plausibility order guarantees that in every state it is possible to make an observation that ‘favors’ that state.

Lemma 4. Suppose that a plausibility order $\preceq$ is appropriate to learn an epistemic space S via **mini**. For all $s \in S$, for all $t \in S$ such that $s \simeq t$ and $t \neq s$, if $|\{p \in O_s : t \notin p\}| < |\mathbb{N}|$, then there exists a $p \in O_s$ such that $s \in \min_{\preceq} p$ and $t \notin p$.

Proof. Towards a contradiction, assume that there are $s, t \in S$ such that: $s \simeq t$, $t \neq s$; $|\{p \in O_s : t \notin p\}| < |\mathbb{N}|$; for all $p \in O_s$ if $s \in \min_{\preceq} p$ then $t \in p$. Let $\vec{O}$ be a stream and $n \in \mathbb{N}$ be such that $\text{set}(\vec{O}[n]) = \{p \in O_s : t \notin p\}$, which exists since $|\{p \in O_s : t \notin p\}| < |\mathbb{N}|$. Denote with $\preceq_{\vec{O}[n]}^{\text{mini}}$ the order in **mini**(PLAUS(S), $\vec{O}[n]$). Since by assumption for all $p \in O_s$ if $t \notin p$ then $s \notin \min_{\preceq} p$, and since by construction for all $p \in \text{set}(\vec{O}[n])$, $t \notin p$, we have that $s \notin \min_{\preceq} p$ for all $p \in \text{set}(\vec{O}[n])$. It follows that $s \simeq_{\vec{O}[n]}^{\text{mini}} t$ (otherwise one of them would be a minimal element of some $p \in \vec{O}[n]$). Consider now any extension of the sequence $\vec{O}[n]$ that enumerates all remaining elements of O_s , i.e. all elements in $O_s \setminus \{p \in O_s : t \notin p\}$. By construction all $p \in \text{set}(\vec{O})$ occurring after the initial sequence $\vec{O}[n]$ are such that $t \in p$. Since $t \simeq_{\vec{O}[n]}^{\text{mini}} s$ and $\vec{O}$ is sound, it follows that for all $m \geq n$, $t \simeq_{\vec{O}[m]}^{\text{mini}} s$ by Def. 12. Thus, there is no m such that $\min_{\preceq_{\vec{O}[m]}^{\text{mini}}} = \{s\}$. Hence, $\preceq$ is not appropriate for learning S via **mini**. $\square$

The lemma 4 ensures that an appropriate plausibility order cannot be deceived by sound and complete data streams such as the one illustrated in Figure 3. The problem there is that a distinct state from s , namely u , is equi-plausible to s and minimal in all observables in which s is minimal. Lemma 4 excludes such configurations of points in finite epistemic spaces.

Lemma 5. Let $S = \langle S, O \rangle$ be an epistemic space, and let $\preceq$ be a plausibility order over S . Consider a state s and assume that there exists a $p \in O_s$ such that $s \in \min_{\preceq} p$. Then for any finite sequence $\vec{O}[n]$ of a sound data stream $\vec{O}$ for s , it holds that $s \in \min_{\preceq_{\vec{O}[n]}^{\text{mini}}} p$.

Proof. We prove this by induction on n . [Base case: $n = 0$] Follows directly from the assumption. [Inductive step] Assume that for all $m < n$, $s \in \min_{\vec{O}[m]}^{\text{mini}} p$. Consider now the p' occurring in $\vec{O}$ after the initial segment $\vec{O}[m]$. Assume towards a contradiction that $s \notin \min_{\vec{O}[m+1]}^{\text{mini}} p$. Hence there is $t \in p$ such that $t \prec_{\vec{O}[m+1]}^{\text{mini}} s$. Since $s \in \min_{\vec{O}[m]}^{\text{mini}} p$ by inductive hypothesis, then $t \not\prec_{\vec{O}[m]}^{\text{mini}} s$, hence $s \preceq_{\vec{O}[m]}^{\text{mini}} t$, since $\preceq$ is total. Since $s \notin \min_{\vec{O}[m]}^{\text{mini}} p'$ (otherwise it would be a minimal element in p after updating on p'), there is $t' \prec_{\vec{O}[m]}^{\text{mini}} s$, and $t \in p'$. By transitivity, $t' \prec_{\vec{O}[m]}^{\text{mini}} t$. Hence, $t \notin \min_{\vec{O}[m]}^{\text{mini}} p'$. Thus by Def. 12, $s \prec_{\vec{O}[m+1]}^{\text{mini}} t$, which contradicts the assumption that $t \prec_{\vec{O}[m+1]}^{\text{mini}} s$. $\square$

To characterise the plausibility orders appropriate for minimal revision, we make use of the concept of a *finite tell-tale map*. We have already seen a concept of this kind in Definition 14. The concept was first introduced by Dana Angluin in her seminal work [2]. Broadly speaking, a finite tell-tale map assigns to each possible state a special finite set of observables (a finite tell-tale), which guides learning and guarantees that the state can be identified. We consider a specific variant of tell-tale maps that satisfies a number of conditions regarding not only the available observables but also the plausibility order of the states.

Definition 17. Let $S = \langle S, O \rangle$ be a finite epistemic space and $\preceq$ be a plausibility order over S be given. A *finite space mini tell-tale map* for $\preceq$ and S is a total map assigning to each state $s \in S$ a *finite space mini tell-tale set* $F_s \subseteq O$ such that: (i) $F_s \subseteq O_s$; (ii) there is a $p \in F_s$ such that $s \in \min_{\preceq} p$ and for all $t \neq s$, $t \simeq s$ such that $t \in \min_{\preceq} p$, there exists $q \in F_s$ such that $s \in \min_{\preceq} q$ and $t \notin \min_{\preceq} q$.

Theorem 3. Let $\preceq$ be a plausibility order on a finite epistemic space $S = \langle S, O \rangle$. The plausibility order $\preceq$ is appropriate to learn S via *mini* iff there exists a finite space mini tell-tale map for S and $\preceq$.

Proof. [$\Rightarrow$] Let $S = \langle S, O \rangle$ be finite, and let $\preceq$ be appropriate for learning S via *mini*. Since $O_s \subseteq O$, it trivially satisfies (i) from Def. 17. Assume towards a contradiction that for all $p \in O_s$, either $p \notin \min_{\preceq} p$ or there exists a $t \neq s$ such that $t \in \min_{\preceq} p$ and there is no $q \in O_s$ such that $s \in \min_{\preceq} q$ and $t \notin \min_{\preceq} q$. By Lemma 3 we know that there is a $p \in O_s$ such that $s \in \min_{\preceq} p$. By assumption, it follows that there exists a $t \in \min_{\preceq} p$ such that there is no $q \neq p \in O_s$ such that $s \in \min_{\preceq} q$ and $t \notin \min_{\preceq} q$. By Lemma 4, it follows that the set $|\{p \in O_s : t \notin p\}| = |\mathbb{N}|$. Contradiction.

[$\Leftarrow$] Let $S = \langle S, O \rangle$ be a finite epistemic space, and let $\preceq$ be given. Assume there exists a finite space mini tell-tale map for $\preceq$ and S . Let $s \in S$ be arbitrary and F_s be its corresponding finite space mini tell-tale set. Since S is finite, also O_s is a finite space mini tell-tale set. Consider an arbitrary sound and complete stream $\vec{O}$ for s . Since O_s is finite and $\vec{O}$ is complete, there exists an $n \in \mathbb{N}$ such that $O_s \subseteq \text{set}(\vec{O}[n])$. Since O_s is a finite space mini tell-tale set for s , by Def. 17 there exists a $p \in O_s$ such that $s \in \min_{\preceq} p$. Let $\vec{O}[m]$ be the sequence where the last element is the first occurrence of a $p \in O_s$ such that $s \in \min_{\preceq} p$ in $\vec{O}$. By Lemma 5, $s \in \min_{\vec{O}[m-1]}^{\text{mini}} p$. By Def. 12, $s \in \min_{\vec{O}[m]}^{\text{mini}}$, and thus $s \in L_{\text{mini}}^{\text{PLAUS}}(S, O[m])$. If there is no $t \neq s$ such that $t \in L_{\text{mini}}^{\text{PLAUS}}(S, O[m])$, then for all $m' \geq m$, $L_{\text{mini}}^{\text{PLAUS}}(S, O[m]) = \{s\}$ (since $\vec{O}$ is sound with respect to s will forever be the uniquely minimal element for all $p \in O_s$). If there is a $t \neq s$ such that $t \in L_{\text{mini}}^{\text{PLAUS}}(S, O[m])$, then $t \simeq_{\vec{O}[m]}^{\text{mini}} s$. This implies that $s \simeq t$, and $t \in \min_{\preceq} p$. Since by assumption S has a finite space mini tell-tale, there is q such that $s \in \min_{\preceq} q$ and $t \notin \min_{\preceq} q$, thus $t \notin q$. Since p was the first occurrence of a $p \in O_s$ such that $s \in \min_{\preceq} p$, since $q \in O_s$ and $s \in \min_{\preceq} q$, there is $m' > m$ such that q is the last observable occurring in the initial segment $\vec{O}[m']$ of $\vec{O}$, since $\vec{O}$ is complete. Take the smallest such s' . Since $s \in \min_{\vec{O}[m]}^{\text{mini}}$, and $\vec{O}$ is sound with respect to s , $s \in \min_{\vec{O}[m'-1]}^{\text{mini}}$ and also $s \in \min_{\vec{O}[m'-1]}^{\text{mini}} q$, and since $t \notin q$, $s \prec_{\vec{O}[m']}^{\text{mini}} t$ by Def. 12. Since t was arbitrary, for all $t \in S$ such that $t \neq s$ and $t \simeq_{\vec{O}[m]}^{\text{mini}}$, there

is $k > m$ such that $s \prec_{\vec{O}[k]}^{\text{mini}} t$. Since the space is finite, there can only be finitely many $t \neq s$ such that $t \in \min_{\vec{O}[m]}^{\text{mini}}$, and thus there is k' such that for all such $t \in S$, $s \prec_{\vec{O}[k']}^{\text{mini}} t$, and $\min_{\vec{O}[k']}^{\text{mini}} = \{s\}$. By Def. 12 and the soundness of $\vec{O}$, for all $k'' \geq k$, $L_{\text{mini}}^{\text{PLAUS}}(S, O[k'']) = \{s\}$. $\square$

3.4 No Guarantees for Minimal Revision when Learning with Errors

Theorem 3 characterises the appropriate priors when learning via **mini** under the assumption that no erroneous observation can occur (i.e., the data streams are sound). Does the characterisation still hold when considering learning scenarios in which errors can occur? Here, we consider fair streams [6], i.e., data streams that can contain only finitely many errors, all of which will eventually be corrected.

Definition 18 ([6, Def. 27]). *Let $S = \langle S, O \rangle$ be a negation-closed epistemic space. A data stream $\vec{O}$ from O is **fair with respect to the state s** precisely in case: (i) $\vec{O}$ is complete with respect to s ; (ii) there is an $n \in \mathbb{N}$ such that for all $k \geq n$, $s \in O_k$; (iii) for every $i \in \mathbb{N}$ such that $s \notin O_i$, there is a $k > i$ such that $O_k = \bar{O}_i$.²*

When considering fair data streams, not only Theorem 3 is no longer an adequate characterisation of the appropriate plausibility orders to learn via **mini**, but also **mini** is not anymore a universal method on the class of negation-closed finite epistemic spaces.³

Proposition 2. *The following two hold: (i) There exist plausibility orders appropriate for learning some negation-closed space via **mini** on sound and complete data streams that are not appropriate for learning the same space via **mini** on fair data streams; (ii) Minimal revision is not universal on fair data-streams on the class of all finite negation-closed epistemic spaces.*

Proof. [Point (i)] Consider the epistemic space at the center of Fig. 2. The space is negation-closed and the plausibility order in which for all $s, t \in S$, $t \simeq s$ is appropriate to identify the space via **mini** on sound and complete data-streams by the proof of Th. 1. Consider the state s and consider the data stream starting with the sequence $q, \bar{p}, p$ followed by an infinite stream of $\bar{q}$. This stream is fair and prevents **mini** from ever identifying s , since w will always be a minimal element with s after the initial segment $q, \bar{p}, p$. [Point (ii)] Consider again the epistemic space at the center of Fig. 2. The only appropriate plausibility order to learn that space via **mini** on sound and complete data streams is such that $s, t \in S, t \simeq s$. This order is not appropriate for learning the space via **mini** on fair data streams as shown above for (i). $\square$

We leave to future work a characterisation of the plausibility assignments that are appropriate for minimal revision on fair data streams (and of the class of spaces on which they exist).

4 What Priors are Appropriate for Stronger Methods?

We turn to considering how one ought to assign priors when employing more radical belief revision methods than minimal revision. In particular, we consider *conditioning* and *lexicographic upgrade*.

Definition 19 (Conditioning [8]). *Let $B = \langle S, O, \preceq \rangle$ be a plausibility space, and let $p \in O$. The one-step revision function $\text{cond}(B, p)$ generates the plausibility space $B' = \langle S', O', \preceq' \rangle$, where $S' = S \cap p$, $O' = \{q \in \mathcal{P}(S') : \exists q' \in O \text{ such that } q = q' \setminus \bar{p}\}$, and $\preceq' = \preceq_p$.*

²All definitions in Section 2 and Definition 16 can be straightforwardly adapted for the case of fair data-streams.

³**mini**'s sensitivity to errors was highlighted in [28] (Proposition 3.3), where it was applied to cases of cognitive bias, in which some observations are (systematically) misinterpreted by the revising agent.

Definition 20 (Lexicographic Upgrade [8]). Let $\mathbf{B} = \langle S, O, \preceq \rangle$ be a plausibility space, and $p \in O$. The one-step revision function $\mathbf{lex}(\mathbf{B}, p)$ generates the plausibility space $\mathbf{B}' = \langle S, O, \preceq' \rangle$, where $s \preceq' t$ if and only if: $s \preceq_p t$, or $s \preceq_{\bar{p}} t$, or $s \in p$ and $t \notin p$.

When revising with **cond** upon receiving some observation p , an agent simply eliminates all the worlds that do not satisfy p , and restricts the plausibility order to the worlds that do satisfy p . On the other hand, lexicographic upgrade does not eliminate worlds, but rather moves all the p -worlds in front of the $\neg p$ -worlds ([30, 9]).

For the case of conditioning and lexicographic revision, we give a fully general characterisation of appropriate plausibility orders. In particular, we consider spaces that are not necessarily finite, as well as preorders that are not necessarily total.⁴ As done for **mini**, we start by defining a tell-tale map variant.

Definition 21. Let $S = \langle S, O \rangle$ be an epistemic space and $\preceq$ be a preorder over S be given. A **generalised conditioning tell-tale map** for $\preceq$ and S is total map assigning to each state $s \in S$ a set $F_s \subseteq O$ such that: (i) F_s is finite; (ii) $F_s \subseteq O_s$; (iii) for all $t \in S$, if $t \sim s$, $F_s \subseteq O_t$ and $s \neq t$, then $s \prec t$; (iv) for all finite $F'_s \supseteq F_s \subseteq O_s$, for all $t \in S$, if $t \not\sim s$ and $F'_s \subseteq O_t$, then there exists a $v \in S$ such that $v \prec t$ and $F'_s \subseteq O_v$.

One can then prove that the existence of a generalised conditioning tell-tale characterises appropriate preorders for conditioning as well as lexicographic upgrade:

Theorem 4. Let $\preceq$ be a preorder over an epistemic space S . The following are equivalent: (i) There exists a conditioning tell-tale map for S and $\preceq$; (ii) $\preceq$ is appropriate to learn S via **cond**; (iii) $\preceq$ is appropriate to learn S via **lex**.

It is easy to see that the equivalence between (i) and (ii) in Theorem 4 fails when considering fair streams (Definition 18): Since **cond** irreparably eliminates the actual state as soon as an error occurs, no preorder is appropriate to learn via **cond** on fair streams. Finally, since any lexicographic upgrade on a finite initial segment of a fair data stream can be shown to be equivalent to the lexicographic upgrade obtained by removing the corrected errors from the sequence [6], the equivalence between (i) and (iii) in Theorem 4 still holds, and thus the existence of a generalised conditioning tell-tale map also characterises the appropriate preorders for learning via **lex** on fair streams in negation-closed spaces.

5 Conclusion

We showed that minimal revision is universal when used to learn finitely identifiable spaces (Theorem 1), and that it is consequently universal on finite strongly separated spaces (Theorem 2). Thus, it can always be used to learn from positive and negative data on finite spaces (Corollary 1) unless erroneous observations possibly occur (Proposition 2). In addition, we characterised the prior plausibility orders that are appropriate for learning finite spaces via minimal revision (Theorem 3) as well as the preorders appropriate for conditioning and lexicographic upgrade (Theorem 4).

Beyond the scope of this paper, several questions are left for future research. First, concerning minimal revision: What is the class of epistemic spaces on which the method is universal? What priors are appropriate for learning via minimal revision on arbitrary spaces, and possibly with wrong information? Second, beyond minimal revision, what priors are appropriate for other revision methods? Finally, we plan to consider appropriate priors from the viewpoint of doxastic attitudes, in the spirit of [14]: What type of doxastic attitudes correspond to holding an appropriate prior?

⁴Definition 16 can be straightforwardly adapted to the case of arbitrary preorders.

Acknowledgements

Rineke Verbrugge acknowledges support from the project “Hybrid Intelligence: Augmenting Human Intellect”, a 10-year Gravitation programme funded by the Dutch Ministry of Education, Culture and Science through the Netherlands Organisation for Scientific Research (grant number 024.004.022). Zoé Christoff acknowledges support from the project “Democracy on Social Networks” (VENI project number VI.Veni.201F.032) financed by the Netherlands Organisation for Scientific Research (NWO). Edoardo Baccini acknowledges support from the Evert Willem Beth Foundation of the Royal Netherlands Academy of Arts and Sciences (KNAW) (grant number KNAW WF/743 – 13). Edoardo Baccini would also like to thank the “Reasoning, Rationality and Science” research group at the Ruhr-University Bochum for insightful discussion on an earlier version of this manuscript. All authors thank three anonymous referees for helpful comments.

References

- [1] Carlos E. Alchourrón, Peter Gärdenfors & David Makinson (1985): *On the Logic of Theory Change: Partial Meet Contraction and Revision Functions*. *Journal of Symbolic Logic* 50(2), pp. 510–530, doi:10.2307/2274239.
- [2] Dana Angluin (1980): *Inductive inference of formal languages from positive data*. *Information and Control* 45(2), pp. 117–135, doi:10.1016/S0019-9958(80)90285-5.
- [3] Dana Angluin & Carl H. Smith (1983): *Inductive inference: Theory and methods*. *ACM Comput. Surv.* 15(3), pp. 237–269, doi:10.1145/356914.356918.
- [4] Alexandru Baltag, Nina Gierasimczuk & Sonja Smets (2011): *Belief revision as a truth-tracking process*. In: *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, TARK XIII*, ACM, New York, NY, USA, pp. 187–190, doi:10.1145/2000378.2000400.
- [5] Alexandru Baltag, Nina Gierasimczuk & Sonja Smets (2016): *On the solvability of inductive problems: A study in epistemic topology*. In Ram Ramanujam, editor: *Electronic Proceedings in Theoretical Computer Science*, 215, Open Publishing Association, p. 81–98, doi:10.4204/eptcs.215.7.
- [6] Alexandru Baltag, Nina Gierasimczuk & Sonja Smets (2019): *Truth-tracking by belief revision*. *Studia Logica* 107, pp. 917–947, doi:10.1007/s11225-018-9812-x.
- [7] Alexandru Baltag & Sonja Smets (2011): *Keep changing your beliefs, aiming for the truth*. *Erkenntnis* 75, pp. 255–270, doi:10.1007/s10670-011-9294-y.
- [8] Alexandru Baltag & Sonja Smets (2016): *A qualitative theory of dynamic interactive belief revision*. In Horacio Arló-Costa, Vincent F. Hendricks & Johan van Benthem, editors: *Readings in Formal Epistemology: Sourcebook*, Springer International Publishing, Cham, pp. 813–858, doi:10.1007/978-3-319-20451-2_39.
- [9] Johan van Benthem (2007): *Dynamic logic for belief revision*. *Journal of Applied Non-Classical Logics* 17(2), pp. 129–155, doi:10.3166/jancl.17.129-155.
- [10] Lenore Blum & Manuel Blum (1975): *Toward a mathematical theory of inductive inference*. *Information and Control* 28(2), pp. 125–155, doi:10.1016/S0019-9958(75)90261-2.
- [11] Thomas Bolander & Nina Gierasimczuk (2015): *Learning actions models: Qualitative approach*. In Wiebe van der Hoek, Wesley H. Holliday & Wen-Fang Wang, editors: *Logic, Rationality, and Interaction - 5th International Workshop, LORI 2015 Taipei, Taiwan, October 28-31, 2015, Proceedings, Lecture Notes in Computer Science* 9394, Springer, pp. 40–52, doi:10.1007/978-3-662-48561-3_4.
- [12] Thomas Bolander & Nina Gierasimczuk (2018): *Learning to act: Qualitative learning of deterministic action models*. *Journal of Logic and Computation* 28(2), pp. 337–365, doi:10.1093/logcom/exx036.

- [13] Craig Boutilier (1993): *Revision sequences and nested conditionals*. In: *Proceedings of the 13th International Joint Conference on Artificial Intelligence - Volume 1, IJCAI'93*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, p. 519–525, doi:10.5555/1624025.1624098.
- [14] Cédric Dégremont & Nina Gierasimczuk (2011): *Finite identification from the viewpoint of epistemic update*. *Information and Computation* 209(3), pp. 383–396, doi:10.1016/j.ic.2010.08.002.
- [15] Hans van Ditmarsch, Wiebe van der Hoek & Barteld Kooi (2007): *Dynamic Epistemic Logic*. Synthese Library Series, Springer, Cham, doi:10.1007/978-1-4020-5839-4.
- [16] R. V. Freivald & Rolf Wiehagen (1979): *Inductive inference with additional information*. *J. Inf. Process. Cybern.* 15, pp. 179–185. Available at <https://api.semanticscholar.org/CorpusID:598084>.
- [17] Nina Gierasimczuk (2010): *Knowing one's limits: Logical Analysis of Inductive Inference*. Ph.D. thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam, Amsterdam, The Netherlands. Available at <https://eprints.illc.uva.nl/id/eprint/2094/>. ILLC Dissertation Series DS-2010-11.
- [18] Nina Gierasimczuk, Vincent F Hendricks & Dick de Jongh (2014): *Logic and learning*. In Alexandru Baltag & Sonja Smets, editors: *Johan van Benthem on Logic and Information Dynamics*, Springer, pp. 267–288, doi:10.1007/978-3-319-06025-5_10.
- [19] Nina Gierasimczuk & Dick de Jongh (2012): *On the complexity of conclusive update*. *The Computer Journal* 56(3), pp. 365–377, doi:10.1093/comjnl/bxs059.
- [20] E Mark Gold (1967): *Language identification in the limit*. *Information and Control* 10(5), pp. 447–474, doi:10.1016/S0019-9958(67)91165-5.
- [21] Adam Grove (1988): *Two modellings for theory change*. *Journal of Philosophical Logic* 17(2), pp. 157–170, doi:10.1007/bf00247909.
- [22] Klaus P. Jantke & Hans-Rainer Beick (1981): *Combining postulates of naturalness in inductive inference*. *J. Inf. Process. Cybern.* 17(8/9), pp. 465–484.
- [23] Kevin Kelly (1999): *Iterated belief revision, reliability, and inductive amnesia*. *Erkenntnis* 50(1), pp. 7–53, doi:10.1023/A:1005444112348.
- [24] R Klette & R Wiehagen (1980): *Research in the theory of inductive inference by GDR mathematicians-A survey*. *Information Sciences* 22, pp. 149–169, doi:10.1016/S0020-0255(80)80006-5.
- [25] Steffen Lange & Thomas Zeugmann (1992): *Types of monotonic language learning and their characterization*. In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory, COLT '92*, Association for Computing Machinery, New York, NY, USA, p. 377–390, doi:10.1145/130385.130427.
- [26] David Lewis (1973): *Counterfactuals*. Library of Philosophy and Logic, Blackwell, Oxford, doi:10.2307/2273738.
- [27] Yasuhito Mukouchi (1992): *Characterization of finite identification*. In: *International Workshop on Analogical and Inductive Inference*, Springer, pp. 260–267, doi:10.1007/3-540-56004-1_18.
- [28] Panagiotis Papadamos & Nina Gierasimczuk (2023): *Cognitive bias and belief revision*. In Rineke Verbrugge, editor: *Proceedings Nineteenth conference on Theoretical Aspects of Rationality and Knowledge, TARK 2023, Oxford, United Kingdom, 28-30th June 2023, EPTCS* 379, pp. 441–454, doi:10.4204/EPTCS.379.34.
- [29] Joseph Singleton & Richard Booth (2024): *Truth-tracking with non-expert information sources*. *J. Artif. Intell. Res.* 81, pp. 619–641, doi:10.1613/JAIR.1.15273.
- [30] Wolfgang Spohn (1988): *Ordinal conditional functions: A dynamic theory of epistemic states*. In William L. Harper & Brian Skyrms, editors: *Causation in Decision, Belief Change, and Statistics: Proceedings of the Irvine Conference on Probability and Causation*, Springer Netherlands, Dordrecht, pp. 105–134, doi:10.1007/978-94-009-2865-7_6.
- [31] Robert Stalnaker (1968): *A theory of conditionals*. In Nicholas Rescher, editor: *Studies in Logical Theory*, Blackwell, pp. 98–112, doi:10.1007/978-94-009-9117-0_2.

- [32] Ana Lucía Vargas-Sandoval (2020): *On the Path to the Truth*. Ph.D. thesis, Institute for Logic, Language and Computation, Universiteit van Amsterdam, Amsterdam, The Netherlands. Available at <https://eprints.illc.uva.nl/id/eprint/2176/>. ILLC Dissertation Series DS-2020-07.

Appendix

Proof of Proposition 1

Proposition 1. *An epistemic space S is finitely identifiable if and only if there exists a definite finite tell-tale map for S .*

Proof. $[\Rightarrow]$ Consider an epistemic space $S = \langle S, O \rangle$ and assume that it is finitely identifiable. Therefore there exists a learner L that finitely identifies every state $s \in S$ on any sound and complete data stream for s . Consider an arbitrary state $s \in S$ and consider an arbitrary sound and complete data stream $\vec{O}$ for s . By Def. 13, it follows that there exists an $n \in \mathbb{N}$ such that $L(S, \vec{O}[n]) = \{s\}$. Define $D_s := \text{set}(\vec{O}[n])$. Assume towards a contradiction that D_s is not a definite finite tell-tale set for s . It follows that there exists a $t \in S$ such that $t \neq s$ and $D_s \subseteq O_t$. Consider now a data stream for $\vec{O}_t$ for t such that $\vec{O}_t[n] = \vec{O}[n]$. It follows that $L(S, \vec{O}_t[n]) = \{s\}$. Thus, L does not finitely identify t contrary to our initial assumption.

$[\Leftarrow]$ Consider an epistemic space $S = \langle S, O \rangle$ and assume there exists a definite finite tell-tale map for S . Let an enumeration $s_0, s_1, \dots, s_i, \dots$ of the elements of S be given. Define L in the following way: $L(S, \vec{O}[k]) = \{s_i\}$, where $i \leq k$ is the smallest i such that $D_{s_i} \subseteq \text{set}(O[k])$, if such s_i exists and no guess has been made in the past; else, L makes no guess. Consider an arbitrary state s_j and a sound and complete data stream $\vec{O}$ for s_j . Since $\vec{O}$ is complete and D_{s_j} is finite, there exists an $n \in \mathbb{N}$ such that $D_{s_j} \subseteq \text{set}(\vec{O}[n])$. Consider the smallest $n' \leq n$ for which $D_{s_j} \subseteq \text{set}(\vec{O}[n'])$ and $j \leq n'$. Such an n' exists. Consider indeed the least n for which $D_{s_j} \subseteq \text{set}(\vec{O}[n])$: If $n \geq j$, then $n' := n$; If $n < j$, set $n' := j$ since $D_{s_j} \subseteq \text{set}(\vec{O}[n]) \subseteq \text{set}(\vec{O}[j])$. We now need to show that L has not made guesses for any $k < n'$. Assume towards a contradiction that there is $k < n'$, such that $L(S, \vec{O}[k]) = \{s_{j'}\}$. Since n' is the smallest $n \in \mathbb{N}$ for which $D_{s_j} \subseteq \vec{O}[n]$ and $j \leq n'$, it follows that $D_{s_j} \not\subseteq \text{set}(\vec{O}[k])$. Thus $s_{j'} \neq s_j$. By definition of L , it follows that $D_{s_{j'}} \subseteq \text{set}(\vec{O}[k]) \subseteq O_{s_j}$. But then $D_{s_{j'}}$ is not a definite finite tell-tale for $s_{j'}$, which contradicts our assumption. Since no previous guess has been made by L before observing the sequence $\vec{O}[n']$, $L(S, \vec{O}[n']) = \{s_j\}$ (if $L(S, \vec{O}[n']) \neq \{s_j\}$, then there would be $s_i \neq s_j$ with $i < j$ and such that $D_{s_i} \subseteq \text{set}(\vec{O}[n']) \subseteq O_{s_j}$, and thus D_{s_i} would not be a definite finite tell-tale set for s_i contrary to our assumptions). By definition of L , no later guess will be output by L . Thus, L finitely identifies s_j . Since s_j was arbitrary, L finitely identifies S . $\square$

Proof of Theorem 4

Theorem 4. *Let S be an epistemic space and let $\preceq$ be a preorder over S . The following are equivalent: (i) There exists a conditioning tell-tale map for S and $\preceq$; (ii) $\preceq$ is appropriate to learn S via **cond**; (iii) $\preceq$ is appropriate to learn S via **lex**.*

We prove Theorem 4 by proving the following two propositions:

Proposition 3. *Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$. The preorder $\preceq$ is appropriate to learn S via **cond** if and only if there exists a generalised conditioning tell-tale map for S and $\preceq$.*

Proposition 4. *Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$. The preorder $\preceq$ is appropriate to learn S via **lex** if and only if there exists a generalised conditioning tell-tale map for S and $\preceq$.*

Proposition 3 proves that (i) and (ii) in Theorem 4 are equivalent. Proposition 4 proves that (i) and (iii) in Theorem 4 are equivalent.

To prove both propositions, we rely on the following definition and lemma.

Definition 22 (Locking sequence [6, Def. 22], originally in [10]). Let an epistemic space $\mathbf{S} = \langle S, O \rangle$, a possible world $s \in S$, a learning method L , and a data sequence σ be given. Sequence σ is called a locking sequence for s if σ is sound with respect to s and for each data sequence τ with $s \in \bigcap \text{set}(\tau)$, $L(\mathbf{S}, \sigma * \tau) = L(\mathbf{S}, \sigma) = \{s\}$.

Lemma 6 ([6, Lemma 1]). If learning method L identifies possible world s in the limit, then there exists a locking sequence for s and L .

Proof of Proposition 3

Proposition 3. Let $\preceq$ be a preorder on an epistemic space $\mathbf{S} = \langle S, O \rangle$. The preorder $\preceq$ is appropriate to learn $\mathbf{S}$ via **cond** if and only if there exists a generalised conditioning tell-tale map for $\mathbf{S}$ and $\preceq$.

Proof. $[\Rightarrow]$ Let a preorder $\preceq$ and an epistemic space $\mathbf{S}$ be given. Assume that the preorder $\preceq$ is appropriate to learn $\mathbf{S}$ via $\preceq$ and **cond**. Let **PLAUS** be the plausibility assignment assigning the preorder $\preceq$ to $\mathbf{S}$. By Def. 16, the learner $L_{\text{cond}}^{\text{PLAUS}}$ identifies the space S in the limit. By Lemma 6, for each state $s \in S$, there exists a locking sequence σ_s for $L_{\text{cond}}^{\text{PLAUS}}$ and s .

Consider an arbitrary s and define $F_s := \text{set}(\sigma_s)$. F_s is finite, since σ_s is finite, and thus satisfies point (i) in Def. 21. Furthermore, $F_s \subseteq O_s$ since $\text{set}(\sigma_s) \subseteq O_s$, and thus satisfies point (ii) in Def. 21.

We now prove that F_s satisfies (iii) of Def. 21. Assume for contradiction that there is a $t \in S$ such that $t \sim s$, $F_s \subseteq O_t$ and $t \neq s$, and $s \not\prec t$. Then, $t \simeq s$, or $t \prec s$. [Case: $t \prec s$] Consider now a sound and complete data stream $\vec{O}$ such that there exists n such that $\vec{O}[n] = \sigma_s$. By Def. 19, and the fact that $F_s \subseteq O_t$, t is in the updated space after $\vec{O}[n]$, just like s . Consider the preorder $\preceq_{\vec{O}[n]}^{\text{cond}}$, i.e., the preorder in the updated space **cond**(**PLAUS**($\mathbf{S}$), $\vec{O}[n]$). Since $t \prec s$, we have that $t \prec_{\vec{O}[n]}^{\text{cond}} s$. Thus $s \notin L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n])$, since $s \notin \min_{\preceq_{\vec{O}[n]}^{\text{cond}}}$.

But by assumption, σ_s is a locking sequence for s and $L_{\text{cond}}^{\text{PLAUS}}$, and thus $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n]) = \{s\}$. [Case: $t \simeq s$] Consider now a sound and complete data stream $\vec{O}$ such that there exists n such that $\vec{O}[n] = \sigma_s$. By Def. 19, and the fact that $F_s \subseteq O_t$, t is in the updated space after $\vec{O}[n]$, just like s . Since $t \simeq s$, we have that $t \simeq_{\vec{O}[n]}^{\text{cond}} s$. By assumption, σ_s is a locking sequence for s and $L_{\text{cond}}^{\text{PLAUS}}$, and thus $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n]) = \{s\}$, and thus $s \in \min_{\preceq_{\vec{O}[n]}^{\text{cond}}}$. But since $t \simeq_{\vec{O}[n]}^{\text{cond}} s$, then $t \in \min_{\preceq_{\vec{O}[n]}^{\text{cond}}}$. Hence $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n]) \neq \{s\}$.

We now prove that F_s satisfies (iv) of Def. 21. Assume by contradiction that there exists a set $F'_s \supseteq F_s \subseteq O_s$ such that there exists a $t \in S$ such that $t \not\sim s$, $F'_s \subseteq O_t$, and there is no v such that $v \prec r$ and $F'_s \subseteq O_v$. Consider now a sound and complete data-stream $\vec{O}$ such that for some $n \in \mathbb{N}$, $\vec{O}[n] = \sigma_s$, and for some $k \geq n$, $\text{set}(\vec{O}[k]) = F'_s$. Since $\vec{O}[n]$ is a locking sequence for s , $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n]) = \{s\}$ and for all sequences τ for s , $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n] * \tau) = \{s\}$. Hence also $L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) = \{s\}$. By assumption, we know that there is a $t \in S$ such that $t \not\sim s$ and $\text{set}(\vec{O}[k]) \subseteq O_t$. Therefore, such a t is in the updated model **cond**(**PLAUS**($\mathbf{S}$), $\vec{O}[k]$), and moreover $t \not\sim_{\vec{O}[k]}^{\text{cond}} s$. By assumption, for such a t , there is no v such that $v \prec r$ and $F'_s \subseteq O_v$. Hence, for all $v \in S$, either $v \not\sim t$ or $F'_s \not\subseteq O_v$. Consider such a v . If $F'_s \not\subseteq O_v$, then v has been eliminated. Therefore, for all $v \sim t$, either v has been eliminated or t is at least as good as v in the updated model, i.e., $t \preceq_{\vec{O}[k]}^{\text{cond}} v$. Therefore, if $v \sim_{\vec{O}[k]}^{\text{cond}} t$, then $t \preceq_{\vec{O}[k]}^{\text{cond}} v$. Thus, in the model updated on the sequence $\vec{O}[k]$, we have $t \in \min_{\preceq_{\vec{O}[k]}^{\text{cond}}}$, and thus $t \in L_{\text{cond}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. Contradiction.

$[\Leftarrow]$ Let a preorder $\preceq$ and an epistemic space $\mathbf{S}$ be given such that for each s there exists a set F_s as in Def. 21. Let the plausibility assignment **PLAUS** be given that assigns the preorder $\preceq$ to the space $\mathbf{S}$. Consider now an arbitrary world s , and a sound and complete data stream $\vec{O}$ for s . Since F_s is finite, there exists an $n \in \mathbb{N}$ such that $F_s \subseteq \text{set}(\vec{O}[n])$. We now show that $L_{\text{cond}}^{\text{PLAUS}}$ identifies s in the limit, by showing

that s is the unique minimal element in $\mathbf{cond}(\mathbf{PLAUS}(\mathbf{S}), \vec{O}[n])$, and that it is the unique minimal element for any $k > n$. Consider $t \in S$ distinct from s in the model obtained after updating on $\vec{O}[n]$. By definition of $\mathbf{cond}$, it must be the case that $\text{set}(\vec{O}[n]) \subseteq O_t$ and thus $F_s \subseteq O_t$. Consider the preorder $\preceq_{\vec{O}[n]}^{\mathbf{cond}}$ in the updated space $\mathbf{cond}(\mathbf{PLAUS}(\mathbf{S}), \vec{O}[n])$. We consider two cases.

[Case 1: $t \sim_{\vec{O}[n]}^{\mathbf{cond}} s$] It follows that $t \sim s$. Since $F_s \subseteq \text{set}(O[n])$, and since $\text{set}(O[n]) \subseteq O_t$ (since otherwise t would have been eliminated), $F_s \subseteq O_t$. By Def. 21 it follows that $s \prec t$, and thus $s \prec_{\vec{O}[n]}^{\mathbf{cond}} t$. Since t is arbitrary, it follows that for all $t \sim_{\vec{O}[n]}^{\mathbf{cond}} s$, $s \prec_{\vec{O}[n]}^{\mathbf{cond}} t$, and thus $s \in \min_{\vec{O}[n]}^{\mathbf{cond}}$.

[Case 2: $t \not\sim_{\vec{O}[n]}^{\mathbf{cond}} s$] Consider a $t \not\sim_{\vec{O}[n]}^{\mathbf{cond}} s$. Since $t \not\sim_{\vec{O}[n]}^{\mathbf{cond}} s$, and since $\text{set}(\vec{O}[n]) \subseteq O_t$, we have that $F_s \subseteq \text{set}(\vec{O}[n])$. By Def. 21 there exists a $v \in S$ such that $v \prec t$ and $\text{set}(\vec{O}[n]) \subseteq O_v$. It follows that $v \prec_{\vec{O}[n]}^{\mathbf{cond}} t$, and thus $t \notin \min_{\vec{O}[n]}^{\mathbf{cond}}$. Since t was arbitrary, there is no $t \in S$ such that $t \not\sim_{\vec{O}[n]}^{\mathbf{cond}} s$ and $t \in \min_{\vec{O}[n]}^{\mathbf{cond}}$.

It follows that, $\min_{\vec{O}[n]}^{\mathbf{cond}} \mathbf{cond}(\mathbf{PLAUS}(\mathbf{S}), \vec{O}[n]) = \{s\}$. Thus $L_{\mathbf{cond}}^{\mathbf{PLAUS}}(\mathbf{S}, \vec{O}[n]) = \{s\}$.

We now show that no new minimal element is created. Consider an arbitrary $k > n$, and assume by contradiction that some $t \neq s$ is a minimal element after updating on the sequence $\vec{O}[k]$. Note that $s \in \min_{\vec{O}[k]}^{\mathbf{cond}}$, since $\vec{O}$ is sound with respect to s , and $\mathbf{cond}$ does not change the relative order of the worlds. Also, $F_s \subseteq \text{set}(\vec{O}[n]) \subseteq \text{set}(\vec{O}[k])$. Since $\text{set}(\vec{O}[k]) \subseteq O_t$, we have $F_s \subseteq O_t$. Either $t \sim_{\vec{O}[k]}^{\mathbf{cond}} s$ or $t \not\sim_{\vec{O}[k]}^{\mathbf{cond}} s$. In the first case, $t \sim s$. Thus, by assumption, $s \prec t$, and by definition of conditioning $s \prec_{\vec{O}[k]}^{\mathbf{cond}} t$, which contradicts the assumption that t is minimal. Therefore, it must be the case that $t \not\sim_{\vec{O}[k]}^{\mathbf{cond}} s$. This implies that $t \not\sim s$. But since $F_s \subseteq \text{set}(\vec{O}[k])$, there exists a $v \in S$ s.t. $v \prec t$ and $\text{set}(\vec{O}[k]) \subseteq O_v$. It follows that, $v \prec_{\vec{O}[k]}^{\mathbf{cond}} t$ and therefore t is not a minimal element of $\mathbf{cond}(\mathbf{PLAUS}(\mathbf{S}), \vec{O}[k])$. Contradiction. $\square$

Proof of Proposition 4

Proposition 4. *Let $\preceq$ be a preorder on an epistemic space $\mathbf{S} = \langle S, O \rangle$. The preorder $\preceq$ is appropriate to learn $\mathbf{S}$ via **lex** if and only if there exists a general conditioning tell-tale map for $\mathbf{S}$ and $\preceq$.*

To prove the proposition above we need a number of facts about **lex**.

Fact 1. *Let $\preceq$ be a preorder on an epistemic space $\mathbf{S} = \langle S, O \rangle$. If $s \sim t$, then $s \sim_p^{\mathbf{lex}} t$, for any $p \in O$.*

Proof. Assume that $s \sim t$. Then $s \preceq t$ or $t \preceq s$. Without loss of generality, assume that $s \preceq t$. There are four cases. [Case 1: $s \in p$ and $t \in p$] Since $s \in p$, $t \in p$ and $s \preceq t$ by assumption, we have that $s \preceq_p^{\mathbf{lex}} t$ by Def. 20. Hence $s \sim_p^{\mathbf{lex}} t$. [Case 2: $s \notin p$ and $t \notin p$] Since $s \notin p$, $t \notin p$ and $s \preceq t$ by assumption, we have that $s \preceq_p^{\mathbf{lex}} t$ by Def. 20. Hence $s \sim_p^{\mathbf{lex}} t$. [Case 3: $s \in p$ and $t \notin p$] If $s \in p$ and $t \notin p$, then $s \preceq_p^{\mathbf{lex}} t$ by Def. 20. Hence, $s \sim_p^{\mathbf{lex}} t$. [Case 4: $s \notin p$ and $t \in p$] If $s \notin p$ and $t \in p$, then $s \preceq_p^{\mathbf{lex}} t$ by Def. 20. Hence, $s \sim_p^{\mathbf{lex}} t$. $\square$

Fact 2. *Let $\preceq$ be a preorder on an epistemic space $\mathbf{S} = \langle S, O \rangle$. Consider $s, t \in S$ such that $s \not\sim t$, and consider a sound data stream $\vec{O}$ for s . Then, for any $n \in \mathbb{N}$, if $\text{set}(\vec{O}[n]) \subseteq O_t$, then $s \not\sim_{\vec{O}[n]}^{\mathbf{lex}} t$.*

Proof. By induction on n . [Base Case: $n = 0$] Given that $\vec{O}[0]$ is empty, the updated space is the same as the initial space, thus $t \not\sim s$. [Inductive Step] Assume that for all $n < m$, $t \not\sim_{\vec{O}[n]}^{\mathbf{lex}} s$ in the corresponding update $\mathbf{lex}(\mathbf{PLAUS}(\mathbf{S}), \vec{O}[n])$. Therefore, $t \not\sim_{\vec{O}[m-1]}^{\mathbf{lex}} s$. Hence, $t \not\sim_{\vec{O}[m-1]}^{\mathbf{lex}} s$ and $s \not\sim_{\vec{O}[m-1]}^{\mathbf{lex}} t$. Assume $\text{set}(\vec{O}[m]) \subseteq O_t$, then $t, s \in p$, where $p \in O_s$ is the proposition occurring in $\vec{O}$ after the sequence $\vec{O}[m-1]$.

By Def. 20, since $s, t \in p$, we have that $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ if and only if $s \preceq_{\vec{O}[m-1]}^{\text{lex}} t$, and $t \preceq_{\vec{O}[m]}^{\text{lex}} s$ if and only if $t \preceq_{\vec{O}[m-1]}^{\text{lex}} s$. Using the inductive hypothesis, we then conclude that $s \not\preceq_{\vec{O}[m]}^{\text{lex}} t$ and $t \not\preceq_{\vec{O}[m]}^{\text{lex}} s$. Hence, $s \not\sim_{\vec{O}[m]}^{\text{lex}} t$. $\square$

Fact 3. Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$. Consider $s, t \in S$ such that $s \prec t$, and consider a sound data stream $\vec{O}$ for s . We have that for any $n \in \mathbb{N}$, $s \prec_{\vec{O}[n]}^{\text{lex}} t$.

Proof. Proof by induction on n . [Base Case: $n = 0$] Given that $\vec{O}[0]$ is empty, the updated space is the same as the initial space, thus $s \prec t$. [Inductive Step] Assume that for all $n < m$, $s \prec_{\vec{O}[n]}^{\text{lex}} t$ in the corresponding update $\text{lex}(\text{PLAUS}(S), \vec{O}[n])$. Therefore, $s \prec_{\vec{O}[m-1]}^{\text{lex}} t$. Since the stream is sound with respect to s , we have $s \in p$, where $p \in O$ occurs in the stream $\vec{O}$ after the sequence $\vec{O}[m-1]$. Either $t \in p$, or $t \notin p$. If $t \in p$, then $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ and $t \not\preceq_{\vec{O}[m]}^{\text{lex}} s$ by Def. 20; hence, $s \prec_{\vec{O}[m]}^{\text{lex}} t$. If $t \notin p$, then again $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ and $t \not\preceq_{\vec{O}[m]}^{\text{lex}} s$ by Def. 20, and thus $s \prec_{\vec{O}[m]}^{\text{lex}} t$. $\square$

Fact 4. Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$. If there is an $s \in S$ such that $s \in \min_{\preceq}$, then for any sound data stream $\vec{O}$ for s , for any $n \in \mathbb{N}$, $s \in \min_{\preceq_{\vec{O}[n]}^{\text{lex}}} \text{lex}(\text{PLAUS}(S), \vec{O}[n])$.

Proof. By induction on n . [Base case] Follows from the fact that $s \in \min_{\preceq}$. [Inductive step] Assume that for all $n < m$, $s \in \min_{\preceq_{\vec{O}[n]}^{\text{lex}}}$, where $\preceq_{\vec{O}[n]}^{\text{lex}}$ is the preorder in the space $\text{lex}(\text{PLAUS}(S), \vec{O}[n])$. This is also true for $n = m - 1$. Therefore, $s \in \min_{\preceq_{\vec{O}[m-1]}^{\text{lex}}}$; hence, for all $t \sim_{\vec{O}[m-1]}^{\text{lex}} s$, $s \preceq_{\vec{O}[m-1]}^{\text{lex}} t$. Now consider $t \sim_{\vec{O}[m]}^{\text{lex}} s$. There are two possibilities: either $t \sim_{\vec{O}[m-1]}^{\text{lex}} s$ or not, i.e., $t \not\sim_{\vec{O}[m-1]}^{\text{lex}} s$. [Case 1: $t \sim_{\vec{O}[m-1]}^{\text{lex}} s$] Since $\vec{O}$ is sound with respect to s , we have $s \in p$, where $p \in O$ is the observable occurring in $\vec{O}$ after the initial segment $\vec{O}[m-1]$. Either $t \notin p$ or $t \in p$. If $t \notin p$, then $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ by Def. 20; if $t \in p$, then since $s \preceq_{\vec{O}[m-1]}^{\text{lex}} t$ by inductive hypothesis, we have $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ by Def. 20. [Case 2: $t \not\sim_{\vec{O}[m-1]}^{\text{lex}} s$] Since $\vec{O}$ is sound with respect to s , we have $s \in p$, where $p \in O$ is the observable occurring in $\vec{O}$ after the initial segment $\vec{O}[m-1]$. Either $t \notin p$ or $t \in p$. If $t \notin p$, then $s \preceq_{\vec{O}[m]}^{\text{lex}} t$ by Def. 20; if $t \in p$, then $s \not\preceq_{\vec{O}[m]}^{\text{lex}} t$ by Def. 20. Hence, for all $t \in S$, if $t \sim_{\vec{O}[m]}^{\text{lex}} s$, then $s \preceq_{\vec{O}[m]}^{\text{lex}} t$. Hence $s \in \min_{\preceq_{\vec{O}[m]}^{\text{lex}}}$. $\square$

We can now show that the existence of a general conditioning tell-tale map (Definition 21) is necessary (Lemma 7, right to left direction of Proposition 4) and sufficient (Lemma 8, left to right direction of Proposition 4) for a preorder to be appropriate to learn via **lex**.

Lemma 7. Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$. If $\preceq$ is appropriate to learn S via **lex**, then there exists a conditioning tell-tale map for S and $\preceq$.

Proof. Let $\preceq$ be a preorder on an epistemic space $S = \langle S, O \rangle$ that is identifiable via lexicographic upgrade. Let the plausibility assignment **PLAUS** be given that assigns $\preceq$ to S . Assume that $L_{\text{lex}}^{\text{PLAUS}}$ identifies S in the limit. By Lemma 6, for every state $s \in S$, there exists a locking sequence σ_s . Consider an arbitrary state s . Let $F_s := \text{set}(\sigma_s)$. We now show that F_s is a conditioning tell-tale for s . F_s is finite, since σ_s is finite, and thus satisfies point (i) of Def. 21. $F_s \subseteq O_s$, since σ_s is sound with respect to s . Thus, F_s satisfies point (ii) of Def. 21. We now show that F_s satisfies (iii) in Def. 21. Assume by contradiction that there is some $t \sim s$ such that $F_s \subseteq O_t$, $s \neq t$ and $s \not\prec t$. Since $s \not\prec t$, either $t \prec s$ or $t \simeq s$. [Case 1: $t \prec s$] Consider a sound and complete data stream $\vec{O}$ for which there is an $n \in \mathbb{N}$ such that $\vec{O}[n] = \sigma_s$.

Since $t \prec s$, $F_s \subseteq O_t$ and $\vec{O}[n]$ is sound with respect to s , we have that $t \prec_{\vec{O}[n]}^{\text{lex}} s$ by Fact 3. Hence, $s \notin \min_{\prec_{\vec{O}[n]}^{\text{lex}}} \text{lex}(\text{PLAUS}(\mathbf{S}), \vec{O}[n])$, and thus $s \notin L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n])$, which contradicts the fact that σ_s is a locking sequence for s . [Case 2: $t \simeq s$] Similar as the case above: s is a minimal element if and only if t is a minimal element. And hence we contradict the assumption that σ_s is a locking sequence.

We now show that F_s satisfies (iv) of Def. 21. Assume in order to derive a contradiction that there exists a finite set $F'_s \supseteq F_s \subseteq O_s$ such that some $t \in S$ with $t \not\sim s$, and $F'_s \subseteq O_t$ has no v such that $v \prec t$ and $F'_s \subseteq O_v$. Consider a data stream $\vec{O}$ such that there is an $n \in \mathbb{N}$ such that $\vec{O}[n] = \sigma_s$, and for which there is a $k \geq n$ such that $\text{set}(\vec{O}[k]) = F'_s$. Since $\vec{O}[n]$ is a locking sequence for s , $L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) = \{s\}$. Consider now $t \in S$ as above: for all $v \in S$, either $v \not\sim t$ or $F'_s \not\subseteq O_v$. We now show that t is a minimal element of $\prec_{\vec{O}[k]}^{\text{lex}}$ in the updated plausibility space $\text{lex}(\text{PLAUS}(\mathbf{S}), \vec{O}[k])$. Consider $v \sim_{\vec{O}[k]}^{\text{lex}} t$: either $v \preceq_{\vec{O}[k]}^{\text{lex}} t$ or $t \preceq_{\vec{O}[k]}^{\text{lex}} v$. If $v \preceq_{\vec{O}[k]}^{\text{lex}} t$, it must be the case that $F'_s \subseteq O_v$, otherwise v would have become strictly less plausible than t at some point and would have stayed so, since $F'_s \subseteq O_t$. But by assumption, it follows that $v \not\sim t$, and hence, either $v \not\sim t$ or, $t \preceq v$. If $v \not\sim t$, then $v \not\sim_{\vec{O}[k]}^{\text{lex}} t$, since both $F'_s \subseteq O_t$ and $F'_s \subseteq O_v$. But $v \sim_{\vec{O}[k]}^{\text{lex}} t$, and thus contradiction. Therefore, it must be the case that $t \preceq v$. Since both t, v satisfy all the propositions occurring in $\vec{O}[k]$, we have that $t \preceq_{\vec{O}[k]}^{\text{lex}} v$. In all cases, for all $v \sim_{\vec{O}[k]}^{\text{lex}} t$, we have $t \preceq_{\vec{O}[k]}^{\text{lex}} v$, which implies that $t \in \min_{\prec_{\vec{O}[k]}^{\text{lex}}}$, and thus $t \in L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. But since σ_s is a locking sequence and $s \in \cap \text{set}(\vec{O}[k])$, we may conclude $L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) = \{s\}$. Contradiction. $\square$

Lemma 8. *Let $\preceq$ be a preorder on an epistemic space $\mathbf{S} = \langle S, O \rangle$. If there exists a conditioning tell-tale map for S and $\preceq$, then $\preceq$ is appropriate to learn $\mathbf{S}$ via lex .*

Proof. Let a space $\mathbf{S}$ be given and let PLAUS be a plausibility assignment assigning a preorder $\preceq$ to $\mathbf{S}$ such that there exists a conditioning tell-tale map for S and $\preceq$.

Consider a state $s \in S$ and a sound and complete data stream $\vec{O}$ for s . Consider a set F_s for s as in Def. 21. Since the data stream is complete and F_s is finite, there exists an $n \in \mathbb{N}$ such that $F_s \subseteq \text{set}(\vec{O}[n])$.

We now show that $L_{\text{lex}}^{\text{PLAUS}}$ outputs s after updating on $\vec{O}[n]$, and forever after that. Consider a $t \neq s$ in $\text{lex}(\text{PLAUS}(\mathbf{S}), \vec{O}[n])$. Either $t \sim_{\vec{O}[n]}^{\text{lex}} s$ or $t \not\sim_{\vec{O}[n]}^{\text{lex}} s$.

[Case 1: $t \sim_{\vec{O}[n]}^{\text{lex}} s$] Either $t \not\sim s$ or $t \sim s$. If $t \not\sim s$, then by Fact 2, we have $\text{set}(\vec{O}[n]) \not\subseteq O_t$, and thus there exists an $m \leq n$ such that $s \preceq_{\vec{O}[m]}^{\text{lex}} t$, and $t \not\preceq_{\vec{O}[m]}^{\text{lex}} s$ by Def. 20. Since $\vec{O}$ is sound with respect to s , we have that $s \prec_{\vec{O}[n]}^{\text{lex}} t$ by Fact 3. If $t \sim s$, then there are two possibilities: $F_s \subseteq O_t$ or $F_s \not\subseteq O_t$. If $F_s \subseteq O_t$, then $s \prec t$ by assumption, and by Fact 3, $s \prec_{\vec{O}[n]}^{\text{lex}} t$. If $F_s \not\subseteq O_t$, then there exists a $p \in \text{set}(\vec{O}[n])$ such that $t \not\in \text{set}(\vec{O}[n])$. Therefore, $s \prec_{\vec{O}[n]}^{\text{lex}} t$ for some $m \leq n$. By Fact 3, $s \prec_{\vec{O}[n]}^{\text{lex}} t$.

[Case 2: $t \not\sim_{\vec{O}[n]}^{\text{lex}} s$] Since $t \not\sim_{\vec{O}[n]}^{\text{lex}} s$, we have that $t \not\sim s$ by Fact 1. It must be the case that $\text{set}(\vec{O}[n]) \subseteq O_t$, otherwise, $s \prec_{\vec{O}[n]}^{\text{lex}} t$. By assumption, we therefore know that there exists a $v \prec t$ such that $\text{set}(\vec{O}[n]) \subseteq O_v$. Since both $\text{set}(\vec{O}[n]) \subseteq O_v$ and $\text{set}(\vec{O}[n]) \subseteq O_t$, we have that $v \prec_{\vec{O}[n]}^{\text{lex}} t$. Therefore, $t \notin \min_{\prec_{\vec{O}[n]}^{\text{lex}}}$.

We therefore know that for all $t \sim s$, $s \preceq_{\vec{O}[n]}^{\text{lex}} t$, and there is no $t \not\sim_{\vec{O}[n]}^{\text{lex}} s$ such that for all $v \sim_{\vec{O}[n]}^{\text{lex}} t$, $t \preceq_{\vec{O}[n]}^{\text{lex}} v$. Therefore $\min_{\prec_{\vec{O}[n]}^{\text{lex}}} = \{s\}$, and thus $L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[n]) = \{s\}$.

We now need to show that for all $k \geq n$, $L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) = \{s\}$. Assume that there is some k for which $L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k]) \neq \{s\}$. By Fact 4, we know that $s \in L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. Therefore, there must exist

a t such that $t \neq s$ and $t \in L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$.

[Case 1: $t \sim_{\vec{O}[n]}^{\text{lex}} s$] Since s is the uniquely minimal element after the sequence $\vec{O}[n]$, it must be the case that $s \prec_{\vec{O}[n]}^{\text{lex}} t$. Since $\vec{O}$ is sound, by Fact 3, $s \prec_{\vec{O}[k]}^{\text{lex}} t$. Contradiction. [Case 2: $t \not\sim_{\vec{O}[n]}^{\text{lex}} s$] Then $t \not\prec s$, and $F_s \subseteq \text{set}(\vec{O}[n]) \subseteq \text{set}(\vec{O}[k]) \subseteq O_t$, otherwise, t would have become less plausible than s . But then, by assumption, there exists a v such that $\text{set}(\vec{O}[k]) \subseteq O_v$ and $v \prec t$. Since $\text{set}(\vec{O}[k]) \subseteq O_v$, and $v \prec t$, we have $v \prec_{\vec{O}[k]}^{\text{lex}} t$ (since they always satisfy the same propositions, their relative plausibility never changes). Thus $t \notin \min_{\prec_{\vec{O}[k]}^{\text{lex}}}$, hence $t \notin L_{\text{lex}}^{\text{PLAUS}}(\mathbf{S}, \vec{O}[k])$. Contradiction. $\square$