Dynamics of True Belief: Learning by Revision

Nina Gierasimczuk

Department of Applied Mathematics and Computer Science
DTU Compute
Technical University of Denmark

Cape-KR
Cape Town, South Africa, February 13th, 2025

What do we mean by 'learning'?

General **qualitative** model of (exact) learning:

- on the basis of incoming data consistent with an underlying concept
- learner achieves a **desired type of knowledge** of the underlying concept.

This perspective in various ways generalises many popular learning topics:

- one step updates with an incoming piece of information:
  Belief Revision Theory, Dynamic Epistemic Logic
- particular algorithmic probabilistic methods of automatic improvement:
  Machine Learning, Bayesian Learning, Reinforcement Learning

Gierasimczuk, N., Inductive Inference and Epistemic Modal Logic. 31st Annual Conference on Computer Science Logic (CSL 2023)

Outline

Subset Spaces, Learnability, and Solvability
+ historical context: learning in the limit and logic

Topo-characterizations of Learnability and Solvability
+ topological semantics for (epistemic) modal logic

Constructive, Order-driven Learning: Belief Revision
+ truth-tracking under cognitive biases

Outline

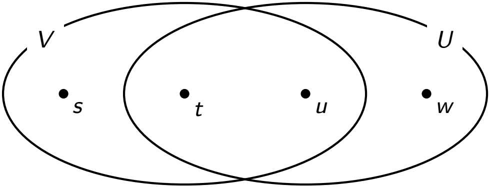
Subset Spaces, Learnability, and Solvability
+ historical context: learning in the limit and logic

Topo-characterizations of Learnability and Solvability
+ topological semantics for (epistemic) modal logic

Constructive, Order-driven Learning: Belief Revision
+ truth-tracking under cognitive biases

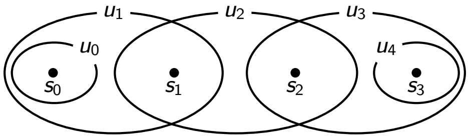
Subset Space

Definition
A **subset space** is $(X, \mathcal{O})$, where $\mathcal{O} \subseteq \mathcal{P}(X)$, X and $\mathcal{O}$ (at most) countable.



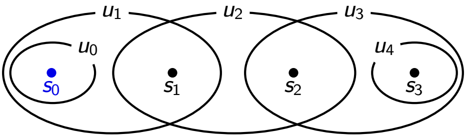
Example: Finite Identifiability

Resulting knowledge: Certainty



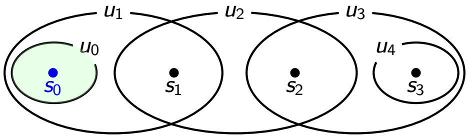
Example: Finite Identifiability

Resulting knowledge: Certainty



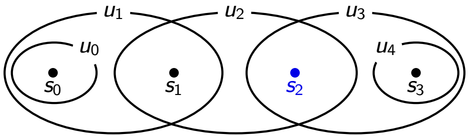
Example: Finite Identifiability

Resulting knowledge: Certainty



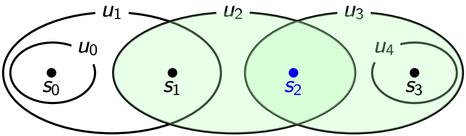
Example: Finite Identifiability

Resulting knowledge: Certainty



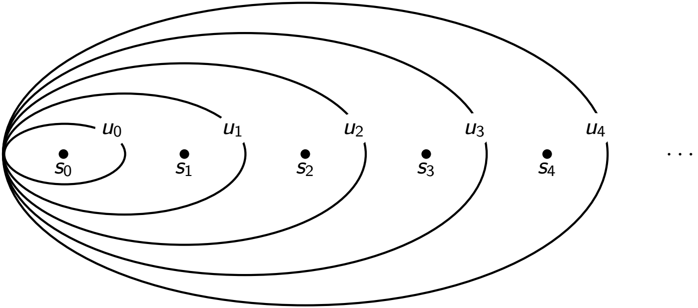
Example: Finite Identifiability

Resulting knowledge: Certainty



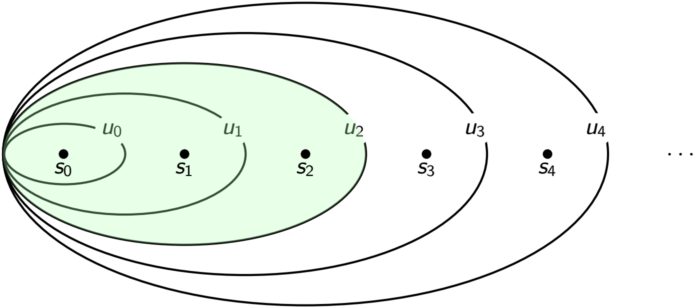
Example: Identifiability in the Limit

Resulting knowledge: undefeated belief



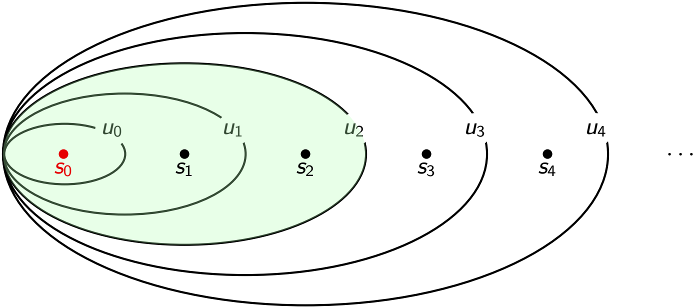
Example: Identifiability in the Limit

Resulting knowledge: undefeated belief



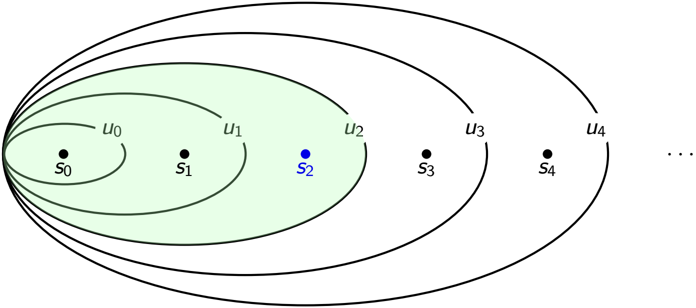
Example: Identifiability in the Limit

Resulting knowledge: undefeated belief

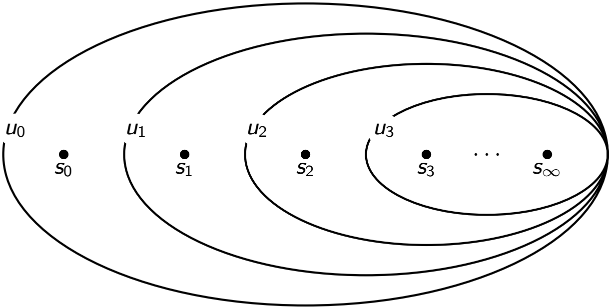


Example: Identifiability in the Limit

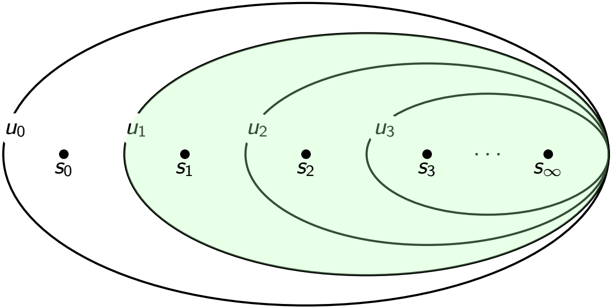
Resulting knowledge: undefeated belief



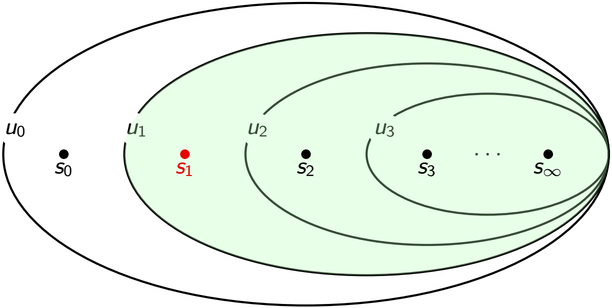
Example: Non-Identifiability in the Limit



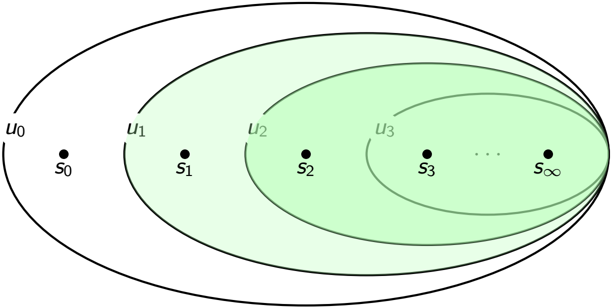
Example: Non-Identifiability in the Limit



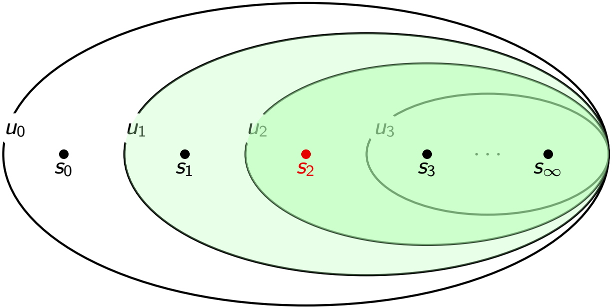
Example: Non-Identifiability in the Limit



Example: Non-Identifiability in the Limit



Example: Non-Identifiability in the Limit



Learning: Streams of Observations

Definition

Let $(X, \mathcal{O})$ be a subset space.

- A **data stream** is an infinite sequence $\vec{O} = (O_0, O_1, \ldots)$ from $\mathcal{O}$.
- A **data sequence** $\vec{O}[n]$ is a finite initial segment of $\vec{O}$ of length $n+1$.

Definition

Take $(X, \mathcal{O})$ and $s \in S$. A data stream $\vec{O}$ is:

- **sound with respect to** s iff every element listed in $\vec{O}$ is true in s.
- **complete with respect to** s iff every observable true in s is listed in $\vec{O}$.

We assume that data streams are sound and complete.

Learning: Learners and Conjectures

Definition

Let $(X, \mathcal{O})$ be a subset space and let σ be a data sequence.

A **learner** L is a function that on σ outputs a conjecture $L(\sigma) \subseteq X$.

Definition

$(X, \mathcal{O})$ is **identified in the limit by** L if for every $x \in X$ and every data stream $\vec{O}$ for x, there is $k \in \mathbb{N}$ s.t.:

$$L(\vec{O}[n]) = \{x\} \text{ for all } n \geq k.$$

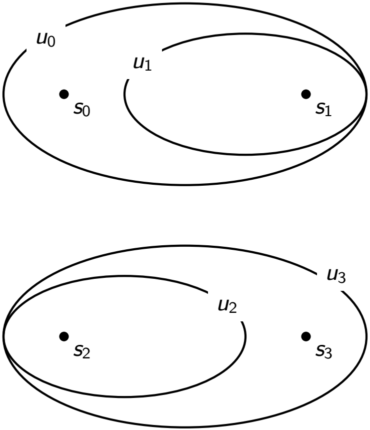
$(X, \mathcal{O})$ is **identifiable in the limit** if it is identified in the limit by a learner L.

Questions, Answers, and Problems

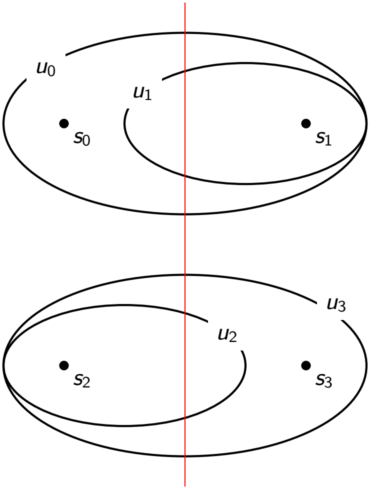
Definition

- A **question** $\mathcal{Q}$ is a partition of X, whose cells A_i are called **answers to** $\mathcal{Q}$.
- Given $x \in A \subseteq X$, $A \in \mathcal{Q}$ is called **the answer to** $\mathcal{Q}$ **at** x, denoted A_x.
- $\mathcal{Q}'$ is a **refinement** of $\mathcal{Q}$ if answers of $\mathcal{Q}$ are disjoint unions of those of $\mathcal{Q}'$.
- A **problem** is a pair $((X, \mathcal{O}), \mathcal{Q})$, where $\mathcal{Q}$ is a question over X.
- $((X, \mathcal{O}), \mathcal{Q}')$ is a **refinement** of $((X, \mathcal{O}), \mathcal{Q})$ if $\mathcal{Q}'$ is a refinement of $\mathcal{Q}$.

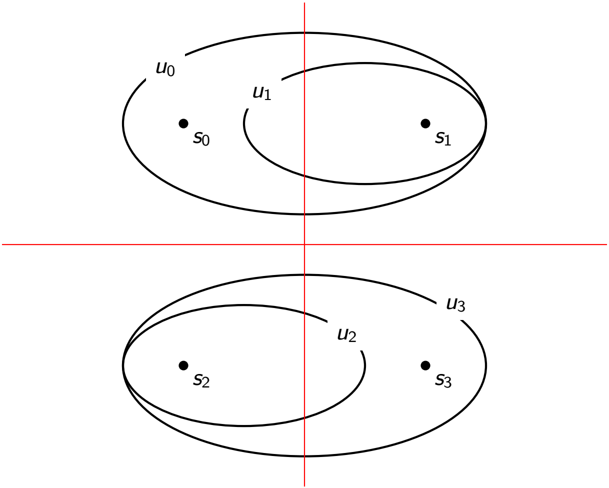
Example: Refinements



Example: Refinements



Example: Refinements



Solving in the Limit

Definition

$((X, \mathcal{O}), \mathcal{Q})$ is **solved in the limit by** L if for every $x \in X$ and every data stream $\vec{O}$ for x, there is $k \in \mathbb{N}$ s.t.:

$$L(\vec{O}[n]) \subseteq A_x \text{ for all } n \geq k.$$

$((X, \mathcal{O}), \mathcal{Q})$ is **solvable in the limit** if solved in the limit by a learner L.

Some historical context

- Hillary Putnam (1965). Trial and error predicates and the solution to...
- E. Mark Gold (1967). Language identification in the limit.
- Ray Solomonoff (1964). A formal theory of inductive inference.

Trial and Error Predicates

A predicate (set) P is decidable if there is a effective procedure φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

Trial and Error Predicates

A predicate (set) P is decidable if there is a recursive function φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

Trial and Error Predicates

A predicate (set) P is decidable if there is a Turing Machine φ such that

$$\begin{array}{lll} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{lll} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

In this context one can think of φ as of a learning function,
Especially if more than two answers are possible.

The quantifier prefix in the definition of trial and error predicates indicates their place in arithmetic hierarchy.

$$
\begin{array}{ccccccccccc}
 & & \mathbf{\Sigma}^0_1 & & & & \mathbf{\Sigma}^0_2 & \cdots & & \mathbf{\Sigma}^0_\alpha & & & \cdots \\
 & \nearrow & & \searrow & & \nearrow & & & \nearrow & & \searrow & & \\
\mathbf{\Delta}^0_1 & & & & \mathbf{\Delta}^0_2 & & & \cdots \quad \mathbf{\Delta}^0_\alpha & & & & \mathbf{\Delta}^0_{\alpha+1} & \cdots \\
 & \searrow & & \nearrow & & \searrow & & & \searrow & & \nearrow & & \\
 & & \mathbf{\Pi}^0_1 & & & & \mathbf{\Pi}^0_2 & \cdots & & \mathbf{\Pi}^0_\alpha & & & \cdots
\end{array}
$$

We will focus on a more general case,
when learner has to pick from more than two options,
in fact, from countably many options.

Outline

Subset Spaces, Learnability, and Solvability

+ historical context: learning in the limit and logic

Topo-characterizations of Learnability and Solvability

+ topological semantics for (epistemic) modal logic

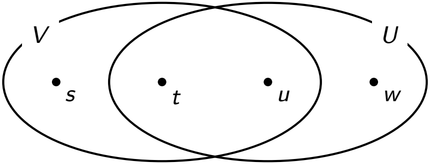
Constructive, Order-driven Learning: Belief Revision

+ truth-tracking under cognitive biases

Definition

A a subset space $(X, \mathcal{O})$ is topological if:

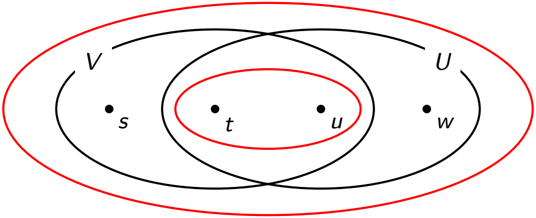
1. $\emptyset \in \mathcal{O}$,
2. $X \in \mathcal{O}$,
3. for any $Y \subseteq \mathcal{O}$, $\bigcup Y \in \mathcal{O}$, and
4. for any finite $Y \subseteq \mathcal{O}$, we have $\bigcap Y \in \mathcal{O}$.



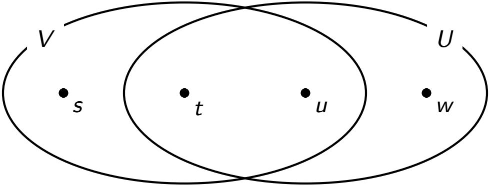
Definition

A a subset space $(X, \mathcal{O})$ is topological if:

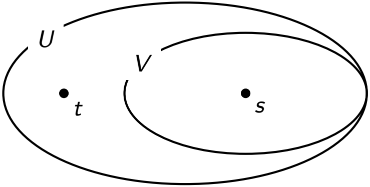
1. $\emptyset \in \mathcal{O}$,
2. $X \in \mathcal{O}$,
3. for any $Y \subseteq \mathcal{O}$, $\bigcup Y \in \mathcal{O}$, and
4. for any finite $Y \subseteq \mathcal{O}$, we have $\bigcap Y \in \mathcal{O}$.



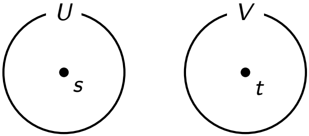
Separability by observations: Illustration



(A) t and u are not separable



(B) weakly separated space, $T0$



(C) strongly separated space, $T1$

DEFINITION

A topological space $(X, \mathcal{O})$ is T_d iff
for every $x \in X$ there is a $U \in \mathcal{O}$ such that $U \setminus \{x\} \in \mathcal{O}$.

T_d is a separation property between $T0$ and $T1$.

DEFINITION

A set A is **locally closed** if $A = U \cap C$, where U is open and C is closed.

Characterization of Solvability in the Limit

Theorem

$((X, \mathcal{O}), \mathcal{Q})$ is solvable in the limit iff $\mathcal{Q}$ has a locally closed refinement.

Corollary

$(X, \mathcal{O})$ is identifiable in the limit iff it is T_d.

A. Baltag, N. Gierasimczuk, S. Smets, On the solvability of inductive problems: a study in epistemic topology, TARK 2015.

Relational semantics for modal logic

Definition (Syntax)

Let P be a countable set of propositional symbols, $p \in P$.

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi$$

Definition (Semantics)

Given a model $M = (W, R, v)$, where $R \subseteq W \times W$, $v : P \to \wp(W)$, $x \in W$:

$$\begin{array}{lll}
M, x \models p & \text{iff} & x \in v(p) \text{ for each } p \in P \\
M, x \models \neg\varphi & \text{iff} & \text{not } M, x \models \varphi \\
M, x \models \varphi \wedge \psi & \text{iff} & M, x \models \varphi \text{ and } M, x \models \psi \\
M, x \models \Box\varphi & \text{iff} & \text{for all } y \in W\text{: if } xRy \text{ then } M, y \models \varphi
\end{array}$$

Some Axioms and Their Epistemic Interpretation

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ (omniscience)

(T) $\Box\varphi \to \varphi$ (truthfullness/reflexivity)

(D) $\Box\varphi \to \neg\Box\neg\varphi$ (consistency/seriality)

(4) $\Box\varphi \to \Box\Box\varphi$ (positive introspection/transitivity)

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$ (negative introspection/Euclidean-ness)

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \rightarrow \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$ (omniscience)

(T) $\Box\varphi \rightarrow \varphi$ (truthfullness/reflexivity)

(D) $\Box\varphi \rightarrow \neg\Box\neg\varphi$ (consistency/seriality)

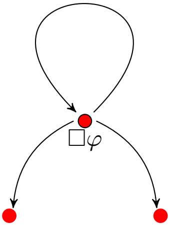
(4) $\Box\varphi \rightarrow \Box\Box\varphi$ (positive introspection/transitivity)

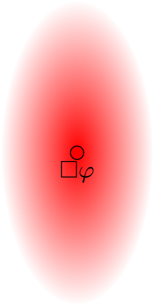
(5) $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$ (negative introspection/Euclidean-ness)

Ax is a logic of a class of models $\mathcal{M}$ iff *Ax* is sound and complete wrt $\mathcal{M}$.

Topological Interpretations

Relational $\Box$ vs Topological $\Box := Int$





Topological Topo-semantics for Modal Logic

Definition (Syntax)

Let P be a countable set of propositional symbols, $p \in P$.

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \Box\varphi$$

Definition

A **topological model** (or a topo-model) $M = (X, \mathcal{O}, v)$ is a topological space $(X, \mathcal{O})$ together with a valuation function $v : P \to \mathcal{P}(X)$.

Definition (Semantics)

Given a topological model $M = (X, \mathcal{O}, v)$ and a state $x \in X$:

$$\begin{array}{lll} M, x \models p & \text{iff} & x \in v(p) \text{ for each } p \in P \\ M, x \models \neg\varphi & \text{iff} & \text{not } M, x \models \varphi \\ M, x \models \varphi \wedge \psi & \text{iff} & M, x \models \varphi \text{ and } M, x \models \psi \\ M, x \models \Box\varphi & \text{iff} & \text{there is } U \in \tau (x \in U \text{ and for all } y \in U\text{: } M, y \models \varphi) \end{array}$$

Sound and Complete Topo-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(T) $\Box\varphi \to \varphi$

(D) $\Box\varphi \to \neg\Box\neg\varphi$

(4) $\Box\varphi \to \Box\Box\varphi$

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$

Sound and Complete Topo-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(T) $\Box\varphi \to \varphi$

(4) $\Box\varphi \to \Box\Box\varphi$

S4 is the topo-logic of all topological spaces (McKinsey & Tarski 1944).

Sound and Complete Topo-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(T) $\Box\varphi \to \varphi$

S4=Topo

(4) $\Box\varphi \to \Box\Box\varphi$

S4 is the topo-logic of all topological spaces (McKinsey & Tarski 1944).

What about T_d-spaces (identifiable in the limit)?

T_d is not topo-definable.

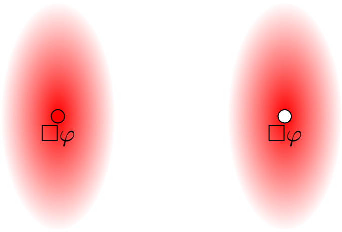
The identifiability-adequate notion of belief is not *topo*-definable.

What about T_d-spaces (identifiable in the limit)?

T_d is not topo-definable.

The identifiability-adequate notion of belief is not *topo*-definable.

But let us, on a whim, change the way we view $\Box$.



DEFINITION (SEMANTICS)

Given a topological model $M = (X, \mathcal{O}, v)$ and a state $x \in X$:

$$
\begin{array}{lll}
M, x \models_d p & \text{iff} & x \in v(p) \\
M, x \models_d \neg\varphi & \text{iff} & \text{not } M, x \models_d \varphi \\
M, x \models_d \varphi \wedge \psi & \text{iff} & M, x \models_d \varphi \text{ and } M, x \models_d \psi \\
M, x \models_d \Box\varphi & \text{iff} & \exists U \in \tau(x \in U \;\&\; \forall y \in U - \{x\}\; M, y \models_d \varphi)
\end{array}
$$

Sound and Complete d-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(T) $\Box\varphi \to \varphi$

(D) $\Box\varphi \to \neg\Box\neg\varphi$

(4) $\Box\varphi \to \Box\Box\varphi$

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$

(w) $(\varphi \wedge \Box\varphi) \to \Box\Box\varphi$

(GL) $\Box(\Box\varphi \to \varphi) \to \Box\varphi$

SOUND AND COMPLETE d-AXIOMATIZATIONS

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(D) $\Box\varphi \to \neg\Box\neg\varphi$

wKD45=dense

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$

(w) $(\varphi \wedge \Box\varphi) \to \Box\Box\varphi$

wKD45 is the d-logic of dense spaces.

Sound and Complete d-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

(D) $\Box\varphi \to \neg\Box\neg\varphi$

(4) $\Box\varphi \to \Box\Box\varphi$

(5) $\neg\Box\varphi \to \Box\neg\Box\varphi$

KD45=DSO

KD45 is the d-logic of DSO-spaces.

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

GL=scattered

(GL) $\Box(\Box\varphi \to \varphi) \to \Box\varphi$

GL is the d-logic of scattered spaces.

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

wK4=Topo

(w) $(\varphi \wedge \Box\varphi) \to \Box\Box\varphi$

wK4 is the d-logic of all topological spaces.

Sound and Complete d-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

K4$=T_d$

(4) $\Box\varphi \to \Box\Box\varphi$

Finally, K4 is the d-logic of all T_d-spaces!

Sound and Complete d-Axiomatizations

Rules

(MP) if $\vdash \varphi$ and $\vdash \varphi \to \psi$, then $\vdash \psi$

(N) if $\vdash \varphi$, then $\vdash \Box\varphi$

Axioms

(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$

And so what...?

(4) $\Box\varphi \to \Box\Box\varphi$

Finally, K4 is the d-logic of all T_d-spaces!

Another way

Get dynamic!

Theorem

Dynamic Logic of Learning Theory is sound and complete with respect to the class of learning models.

Baltag, A., Gierasimczuk, N., Özgün, A., Vargas Sandoval, A.L., and Smets S., A dynamic logic for learning theory. J. Log. Algebr. Meth. Program. 2019.

Outline

Subset Spaces, Learnability, and Solvability
+ historical context: learning in the limit and logic

Topo-characterizations of Learnability and Solvability
+ topological semantics for (epistemic) modal logic

Constructive, Order-driven Learning: Belief Revision
+ truth-tracking under cognitive biases

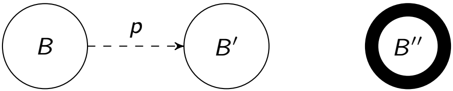
Background

- Learning and belief revision go their separate ways,
- conjecture dynamics is a common theme.
- What are the principles of this dynamics?

$$\begin{tikzcd} B \arrow[r, dashed, "p"] & B' \end{tikzcd}$$

Background

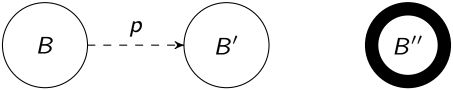
- Learning and belief revision go their separate ways,
- conjecture dynamics is a common theme.
- What are the principles of this dynamics?



Truth-tracking!

Background

- Learning and belief revision go their separate ways,
- conjecture dynamics is a common theme.
- What are the principles of this dynamics?



Truth-tracking!

Order-driven learning: Motivation

- Belief Revision: minimal states give beliefs.
- Computational Learning Theory: co-learning, learning by erasing.
- Philosophy of Science: Ockham's razor.

Plausibility Spaces

A **plausibility space**, $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, consists of an epistemic space $\mathbb{S} = (S, \mathcal{O})$ and a plausibility preorder $\preceq \subseteq S \times S$.

Plausibility Spaces

A **plausibility space**, $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, consists of an epistemic space $\mathbb{S} = (S, \mathcal{O})$ and a plausibility preorder $\preceq \subseteq S \times S$.

Knowledge and Belief

$$\begin{array}{lll} \mathbb{B}_{\mathbb{S}} \models Kp & \text{iff} & S \subseteq p \\ \mathbb{B}_{\mathbb{S}} \models Bp & \text{iff} & min_{\preceq} S \subseteq p. \end{array}$$

For non-well-founded spaces we generalise to:

$$\mathbb{B}_{\mathbb{S}} \models Bp \quad \text{iff} \quad \exists w \, \forall u \preceq w \;\; u \in p.$$

Belief-Revision Methods

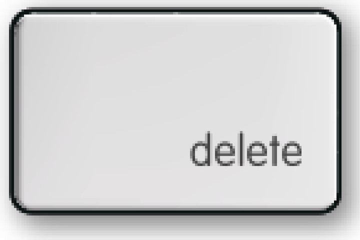
Definition

A **belief-revision method** is a function R that, for any plausibility space $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$ and any observation O outputs a new plausibility space:

$$R(\mathbb{B}_{\mathbb{S}}, O) := (S', \mathcal{O}, \preceq').$$

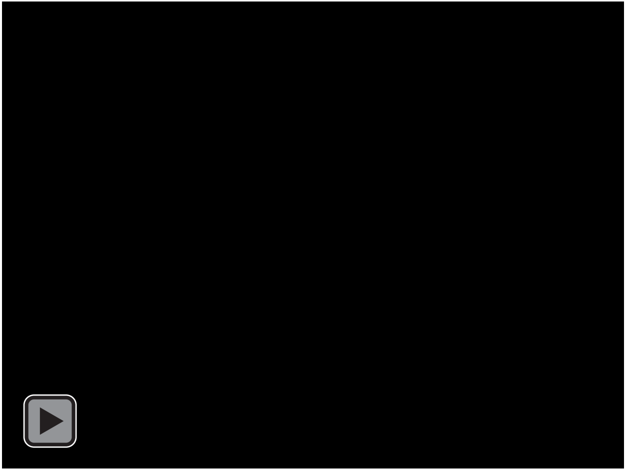
A belief revision R can be iterated in the following way:
$R(\mathbb{B}_{\mathbb{S}}, \sigma * O) := R(R(\mathbb{B}_{\mathbb{S}}, \sigma), O)$, where σ is a finite sequence of observations.

- **Conditioning** eliminates all worlds of S that do not satisfy the observation.



Lexicographic upgrade

- **Lexicographic upgrade** rearranges the preorder by putting all worlds satisfying the observation to be more plausible than others.



Minimal upgrade

- **Minimal upgrade** rearranges the preorder by making only the most plausible states satisfying the observation more plausible than all others, leaving the rest of the preorder the same.

Learning via Belief Revision

Definition

Every belief-revision method R, together with a prior plausibility $\preceq$ generates in a canonical way a learning method $L_R^{\preceq}$ called a **belief-revision-based learning method**, and given by:

$$L_R^{\preceq}((S, \mathcal{O}), \sigma) := \min_{\preceq} R((S, \mathcal{O}, \preceq), \sigma).$$

Definition

An epistemic space $\mathbb{S}$ is **learnable by a belief-revision method** R if there exists a prior plausibility assignment $\preceq$ such that $L_R^{\preceq}$ learns $\mathbb{S}$.

A. Baltag, N. Gierasimczuk, S. Smets. Truth tracking by belief revision. Studia Logica 2018.

Universality Results

Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Universality Results

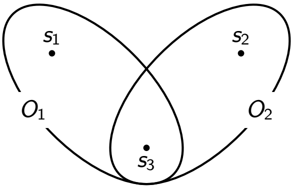
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

Minimal revision is not universal.



Universality Results

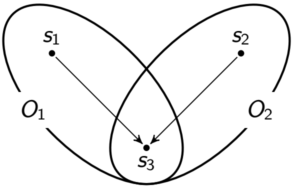
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

Minimal revision is not universal.



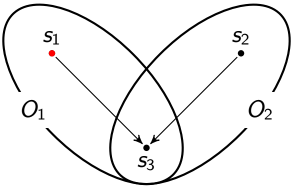
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

Minimal revision is not universal.



Universality Results

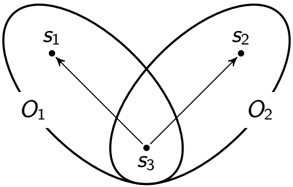
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

Minimal revision is not universal.



Universality Results

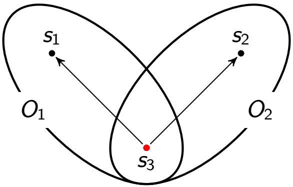
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

Minimal revision is not universal.



Universality Results

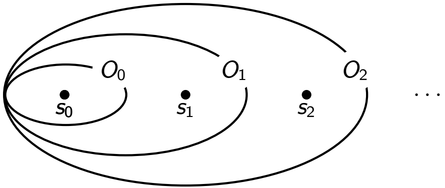
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

Theorem

There is no universal belief-revision method under well-foundedness.



Universality Results

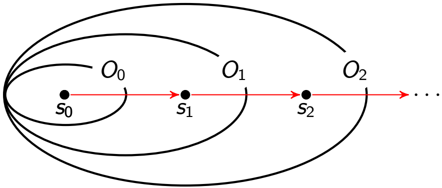
Definition

L is **universal** if it can learn every epistemic space that is learnable.

	Conditioning	Lexicographic	Minimal
Positive Streams	YES	YES	NO

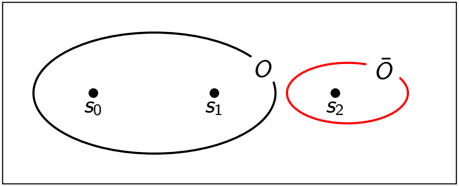
Theorem

There is no universal belief-revision method under well-foundedness.



Is $\neg O$ observable?

An epistemic space $\mathbb{S} = (S, \mathcal{O})$ is **negation-closed** iff if $O \in \mathcal{O}$, then $\bar{O} \in \mathcal{O}$.

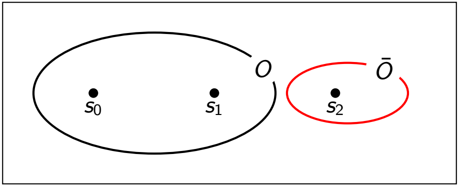


Definition

Let $\mathbb{S} = (S, \mathcal{O})$ be a negation-closed epistemic space. A stream $\vec{O}$ is **fair** with respect to the world s if $\vec{O}$ is complete wrt to s, and contains only finitely many observations O, s.t. $s \notin O$ and every such error is eventually corrected in $\vec{O}$.

Is $\neg O$ observable?

An epistemic space $\mathbb{S} = (S, \mathcal{O})$ is **negation-closed** iff if $O \in \mathcal{O}$, then $\bar{O} \in \mathcal{O}$.



Definition

Let $\mathbb{S} = (S, \mathcal{O})$ be a negation-closed epistemic space. A stream $\vec{O}$ is **fair** with respect to the world s if $\vec{O}$ is complete wrt to s, and contains only finitely many observations O, s.t. $s \notin O$ and every such error is eventually corrected in $\vec{O}$.

	Conditioning	Lexicographic	Minimal
Positive and Negative	YES	YES	NO
Fair Streams	NO	YES	NO

Types of bias

We combine revision methods with three kinds of 'cognitive' limitations:

- issues with accepting the input;
- issues with perceiving the input;
- issues with access to belief state.

We investigate (theoretically and practically) their impact on truth-tracking.

The procedure

After the random generation of an epistemic space, one of the states, s, is randomly designated to be the actual world, and a sound and complete data sequence σ for s is generated. Then a plausibility preorder is randomly generated and each of the (biased) revision methods is called to identify s on σ.

The simulation

Each series of tests consisted of 200 runs, the plausibility spaces consisted of 5 possible worlds and 12 observables, and the incoming data sequence was always longer than the number of observables.

Issues with accepting the input (confirmation bias)

Definition
Given an $(S, \mathcal{O})$, the *stubbornness function* is $D : \mathcal{P}(S) \to \mathbb{N}$.

Definition
R_{CB} is defined in the following way:

$$R_{CB}(\mathbb{B}, \lambda) = \mathbb{B},$$

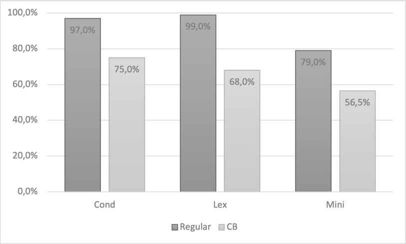
$$R_{CB}(\mathbb{B}, \sigma \cdot p) = \begin{cases} R_1(R_{CB}(\mathbb{B}, \sigma), p) & \text{if} \quad \#p(\sigma) \geq D(\overline{p}), \\ R_{CB}(\mathbb{B}, \sigma) & \text{otherwise.} \end{cases}$$

We obtain $Cond_{CB}$, Lex_{CB}, $Mini_{CB}$.

Results

Proposition
Cond, Lex, Mini are strictly more powerful than $Cond_{CB}$, Lex_{CB}, $Mini_{CB}$.

Corollary
$Cond_{CB}$ and Lex_{CB} are not universal.



Issues with perceiving the input (framing bias)

Definition

Given $(S, \mathcal{O})$, the *framing function* is $FR : \mathcal{O} \to \mathcal{P}(S)$.

Definition (Framing-bias methods)

We define a framing-biased method in the following way:

$$R_{FR}(\mathbb{B}, \lambda) = \mathbb{B},$$

$$R_{FR}(\mathbb{B}, \sigma \cdot p) = R_1(R_{FR}(\mathbb{B}, \sigma), x), \text{ such that } x \in FR(p).$$

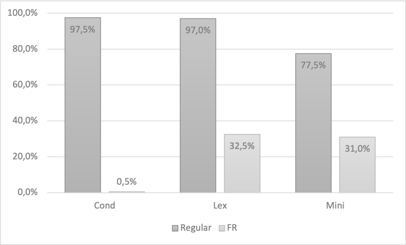
We obtain $Cond_{FR}$, Lex_{FR}, $Mini_{FR}$.

Results

Proposition
$Cond_{FR}$ and Lex_{FR} are not universal.

Proposition
Mini is strictly more powerful than $Mini_{FR}$.

Proposition
Lex_{FR} is universal on fairly framed streams.



Issues with access to belief state (anchoring bias)

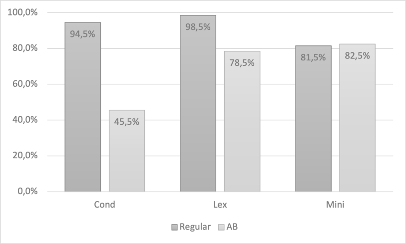
Definition (Anchoring-bias methods)

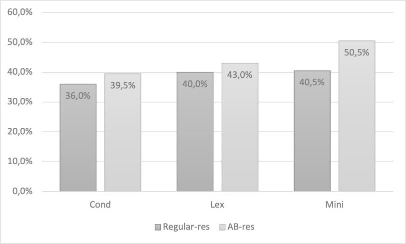
We define a anchoring-biased method in the following way: after receiving p they 'choose' the best world satisfying p. If there are minimal several worlds that are equi-plausible, the method picks one at random.

Additionally, a **resource parameter** (a real number between 0 and 100) halves each time a revision takes place, and the process terminates when the resource is depleted (< 1).

Results

Proposition
$Cond_{AB}$, Lex_{AB} are not universal.





Anchoring ability to select a random world as the candidate for the actual world improves the truth-tracking capability, especially in the case of minimal revision.

Conclusions

Inductive inference for epistemic spaces
can be viewed from a topological perspective
which lends itself to modal dynamic epistemic logic.

Conclusions

Inductive inference for epistemic spaces
can be viewed from a topological perspective
which lends itself to modal dynamic epistemic logic.

Moreover, it can also give a long-term horizon
of truth-tracking to iterated belief revision.

Conclusions

Inductive inference for epistemic spaces
can be viewed from a topological perspective
which lends itself to modal dynamic epistemic logic.

Moreover, it can also give a long-term horizon
of truth-tracking to iterated belief revision.

Recent related work:

- Baltag, Bezhanishivili, and Fernández-Duque, Topology of Surprise. Proc. 19th KR, 2022
- Booth and Singleton, Truth-tracking with Non-expert Information Sources. JAIR 2024.