The Topology of Learning

Ethan Lu

Last Updated: December 10, 2021

Contents

0	Abo	out		2
1	Set-Theoretic Preliminaries			
	1.1	Propos	sitional Structures	2
	1.2	Topolo	gical Operators	5
	1.3	Open and Closed Sets		
	1.4	Topolo	ogies	
		1.4.1	Topological Bases	1
		1.4.2	Separation Properties	.1
2 Context of Inquiry			Inquiry 1	2
-	2.1	Metaphysical Basis		
		2.1.1	Metric Similarity	.3
	2.2	The En	npirical Basis	.6
		2.2.1	Information Gathering	.6
		2.2.2	Measurement and its Consequences	.7
		2.2.3	Between Topology and Empiricism	8
		2.2.4	The Problems of Induction and Metaphysics	9
		2.2.5	Distribution Laws and Their Failures	21
		2.2.6	Simplicity	22
	2.3	The Er	otetic Basis	23
3	Methodology 24			
0	3.1	Epister	mic Virtues of Deduction	24
	3.2 Topological Characterizations of Inductive Solvability		gical Characterizations of Inductive Solvability	28
	3.3	Ockha	m's Razor and its Epistemic Iustification	29
	3.4	Neglig	ibility	50
		0.0		

0 About

These notes are from the class 80-524/824: The Topology of Learning as taught by Professor Kevin Kelly. Consequently, I claim no credit for any of the ideas presented below, but I hope that you find them interesting nonetheless!

I should also say that (as of the time of writing) these notes are still somewhat of a work in progress, particularly towards the later sections. If you find any typoes and such, I'd definitely appreciate a ping about it! ⁽²⁾

1 Set-Theoretic Preliminaries

1.1 Propositional Structures

We'll begin this class with same basic set theoretic definitions that'll guide our discussions throughout. As our basic setup, let W be the set of all worlds (presumably there should be some discussion on why such a set should even exist, but we'll take it as given for now). Given any logical proposition P (e.g. "the sky is blue"), we can naturally identify it with the set of all worlds in which it holds true, and hence write $P \subseteq W$.

Given this identification, we can then naturally lift the logical operations defined on propositions to operations defined on our sets.

Definition 1.1 (Logical Operations on Propositions). Given any $A, B \subseteq W$, we define the following:

- Disjunction: $A \lor B := A \cup B$
- Conjunction: $A \land B := A \cap B$
- Negation: $\neg A := W \setminus A$.
- Difference: $A B := A \setminus B = A \cap \neg B$.

Exercise 1.2. Check that these definitions work, in the sense that $A \land B$ is the set of all worlds where either A or B is true, $A \lor B$ is the set of all worlds where both A and B hold, and so on.

While the above gives us the tools for working with arbitrary propositions, more often than not we'll want us to restrict to propositions of a particular class (e.g. decidable propositions, verifiable propositions), etc. This motivates the following definition:

Definition 1.3. A collection \mathcal{A} on a set W is just a set $\mathcal{A} \subseteq \mathcal{P}(W)$ of propositions on W.

Definition 1.4 (Operations on Collections). Given any collection A of propositions, we define the following:

• Disjunction: $\forall \mathcal{A} := \cup \mathcal{A}$.

- Conjunction: $\wedge \mathcal{A} := \bigcap \mathcal{A}; \bigcap \emptyset = W.$
- Dualization: $\neg A := \{\neg A : A \in A\}.$
- Localization: $\mathcal{A}_w := \{A \in \mathcal{A} : w \in A\}.$
- Restriction $\mathcal{A} \mid B := \{A \cap B : A \in \mathcal{A}\}.$
- Subsets: $\mathcal{A}_{\subseteq B} := \{A \in \mathcal{A} : A \subseteq B\} = \mathcal{P}(B) \cap \mathcal{A}.$
- Supersets: $\mathcal{A}_{\supseteq B} := \{A \in \mathcal{A} : A \supseteq B\}.$
- Quotients: $A/\sim := \{[A] : A \in A; A \text{ respects } \sim (\text{i.e. that } x \in A \implies [x] \subseteq A)\}.$

Proposition 1.5. $A \mid B = A_{\subseteq B}$ *if and only if* A *is closed under intersection with* B.

Of course, every collection is inherently tied to the set it's associated with, but referring to both can get cumbersome. To be lazy, we making the following definition.

Definition 1.6. A propositional structure (W, A) is just a pairing of a set W and a collection $A \subseteq \mathcal{P}(W)$.

Of course, working with arbitrary propositional structures is a task far too complicated for anyone's imagination, so typically we classify these structures based on the types of *closure laws* they satisfy: some examples follow.

EXAMPLE 1.7 Closure laws of propositional structures

- Boolean algebras: closed under negation and finite conjunction/disjunction.
- Sigma algebra: closed under negation and countable conjunctions/disjunctions.
- Sigma ideal: closed under subsets and countable unions (relative to a σ -algebra).
- Topologies: closed under finite conjunctions and arbitrary disjunction.
- Topological bases: closed (kind of) under finite conjunctions and satisfies the covering axiom.

While these closure laws may initially seem somewhat arbitrary, the contexts in which we'll be using them will naturally come with motivation for these properties.

Now suppose we have two propositional structures (W, A), (V, B). Given a function $\phi : W \to V$, we can lift it to collections and structures in the obvious way:

$$\phi[B] = \{\phi(b) \mid b \in B\}$$

$$\phi[\mathcal{A}] = \{\phi[B] \mid B \in \mathcal{A}\}$$

$$\phi[(W, \mathcal{A})] = (\phi(W), \phi[\mathcal{A}])$$

$$\phi^{-1}[C] = \{b \mid \phi(b) \in C\}$$

$$\phi^{-1}[\mathcal{B}] = \{\phi^{-1}[C] \mid B \in \mathcal{B}\}$$

$$\phi^{-1}[(V, \mathcal{B})] = (\phi^{-1}(V), \phi^{-1}[\mathcal{B}])$$

Using these definitions, we're now ready to define the notion of continuity between propositional structures.

Definition 1.8 (Maps between structures). A map $\phi : (W, \mathcal{A}) \to (V, \mathcal{B})$ is **continuous** if $\phi^{-1}[\mathcal{B}] \subseteq \mathcal{A}$ (equivalently, $\phi[\mathcal{A}] \supseteq \mathcal{B}$).

It's called a **homeomorphism** if it's bijective, continuous, and has a continuous inverse. We call any two structures for which a homeomorphism exists homeomorphic.

Proposition 1.9. *The relation* ~ *on the proper class of propositional structures given by homeomorphism is an equivalence relation.*

Definition 1.10 (Operations on structures). We lift once more to define the following operations on structures:

- Dual: $\neg(W, \mathcal{A}) := (W, \neg \mathcal{A}).$
- Localization: $(W, \mathcal{A})_w := (W, \mathcal{A}_w)$.
- Restriction: $(W, \mathcal{A}) \mid B := (B, \mathcal{A} \mid B)$.
- Weak restriction. $(W, \mathcal{A})[B] := (B, \mathcal{A}_{\subseteq B}).$
- Quotient: $(W, \mathcal{A})/\sim := (W/\sim, \mathcal{A}_w/\sim).$

Finally, we can define the categoric theoretical product and co-product of propositional structures:

Definition 1.11 (Product).

$$\prod_{i \in I} (W_i, \mathcal{A}_i) := \left(\prod_{i \in I} W_i, \left\{ \prod_i A_i \mid A_i \in \mathcal{A}_i, \text{ and } A_i = W_i \text{ for all but finitely many } i \right\} \right)$$

Definition 1.12 (Coproduct).

$$\sum_{i\in I} (W_i, \mathcal{A}_i) := \left(\bigsqcup_{i\in I} W_i, \left\{\bigcup_{i\in I} A_i \mid A_i \in \mathcal{A}_i\right\}\right).$$

1.2 Topological Operators

Now given an arbitrary propositional structure (W, A), we'll start by defining some topological operators on the underlying space, motivated by the notions of *verifiability* and *refutability*. An important thing to note is that in this section, although these operators are most naturally associated to cases where A is a topology, this initial section will make no assumptions on the structure of A.

Definition 1.13. Given a structure (W, A), we define the interior operator int $: \mathcal{P}(W) \to \mathcal{P}(W)$, via

$$\operatorname{int} B := \bigcup \mathcal{A}_{\subseteq B} = \bigcup \{ A \in \mathcal{A} \mid A \subseteq B \}$$

This can be interpreted as the set of all worlds in *B* such that *B* is **verified** by an element of A, or alternatively, the portion of *B* covered by A.

Remark 1.14. When \mathcal{A} is a topology, we have $int : \mathcal{P}(W) \to \mathcal{A}$.

Proposition 1.15. The following hold:

- 1. int $W = \bigcup A \subseteq W$, with equality if and only if A covers W.
- 2. *Extensivity:* int $A \subseteq A$.
- *3. Isotony:* $A \subseteq B$ *implies* int $A \subseteq int B$
- 4. *Idempotency:* int int A = int A.
- 5. $\bigcup_{i \in \alpha} \operatorname{int} A_i \subseteq \operatorname{int} \bigcup_i A_i$ (arbitrary union).
- 6. int $\bigcap_i A_i \subseteq \bigcap_i$ int A_i (arbitrary intersection).

All the properties above are an immediate consequence of the definition and left as an exercise to the reader. Dual to the interior operator is the *exterior operator*, which we define now.

Definition 1.16. We define the exterior operator $ext : \mathcal{P}(W) \to \mathcal{P}(W)$, via

 $\operatorname{ext} A := \operatorname{int} \neg A.$

Using the interpretation above, this set corresponds to the set of all worlds where A is **refuted** by an element of A.

Using both of these definitions, we can now define the boundary operator

Definition 1.17. The boundary operator ∂ : $\mathcal{P}(W) \rightarrow \mathcal{P}(W)$, is defined via

 $\partial A := \neg(\operatorname{int} A \cup \operatorname{ext} A)$

which is the set of all worlds where *A* is not **decidable**.

Proposition 1.18. ∂ *is* self-dual: *that is*, $\partial A = \partial \neg A$ *for all* $A \subseteq W$.

Proof. Just using the definition, we have that

 $\partial A = \neg(\operatorname{int} A \cup \operatorname{int} \neg A) = \neg(\operatorname{int} \neg A \cup \operatorname{int} \neg \neg A) = \partial \neg A.$

An alternative proof is the fact that if *A* is not decidable, neither is $\neg A$.

Proposition 1.19. *For all* $A \subseteq W$ *,* $W = int A \sqcup \partial A \sqcup ext A$ *.*

Proof. By extensivity, int $A \cap \text{ext} A = \emptyset$. Then by definition ∂A will mop up the rest.

Next is the closure operator, a familiar object from e.g. the theory of metric spaces. **Definition 1.20.** We define the closure operator $cl : \mathcal{P}(W) \to \mathcal{P}(W)$ via

$$\operatorname{cl} A := \neg \operatorname{ext} A.$$

This is the set of all worlds where *A* is not refuted by a true element of A.

As expected, we have the analogue of proposition 1.15: **Proposition 1.21.** *The following hold:*

- 1. $\operatorname{cl} \emptyset = \neg \cup \mathcal{A} \supseteq \emptyset$, with equality if and only if \mathcal{A} covers W.
- 2. *Extensivity:* $cl A \supseteq A$.
- *3. Isotony:* $A \subseteq B$ *implies* $\operatorname{cl} A \subseteq \operatorname{cl} B$
- 4. *Idempotency:* $\operatorname{cl} \operatorname{cl} A = \operatorname{cl} A$.
- 5. $\bigcup_{i \in \alpha} \operatorname{cl} A_i \subseteq \operatorname{cl} \bigcup_i A_i$ (arbitrary union).
- 6. cl $\bigcap_i A_i \subseteq \bigcap_i cl A_i$ (arbitrary intersection).
- 7. cl $A = int A \sqcup \partial A = A \cup \partial A$.

As before, these follow quickly using duality of the interior and closure operators.

Now we turn to properties of the **frontier operator**

Definition 1.22. We define the frontier operator frnt : $\mathcal{P}(W) \rightarrow \mathcal{P}(W)$ via

frnt $A := \operatorname{cl} A \setminus A$.

This is the set of all worlds where *A* is false but not refuted by a true element of A. From this observation, it should be clear that frnt $A = \partial A \cap \neg A$, which can also be shown from the fact that $\partial A = \neg \operatorname{int} A \cap \operatorname{cl} A$.

Proposition 1.23. We have the following fundamental partitions:

- $W = \operatorname{int} B \sqcup \operatorname{frnt} B \sqcup \operatorname{frnt} \neg B \sqcup \operatorname{int} \neg B$.
- $\operatorname{cl} B = \operatorname{int} B \sqcup \operatorname{frnt} B \sqcup \operatorname{frnt} \neg B$.
- $B = \operatorname{int} B \sqcup \operatorname{frnt} \neg B$.

1.3 Open and Closed Sets

Prior to our discussion of topologies, which will serve as a specialization of the things discussed in this section, we'll take a moment to talk about the general notion of open and closed sets for arbitrary structures.

Definition 1.24. Given a structure (W, A), we call a set B **open** iff $B \subseteq \text{int } B$ iff B = int B. We denote the collection of all such open sets O.

The interpretation here is that "No matter how *B* is true, it is verified by some true element of A".

Proposition 1.25. *O is closed under arbitrary unions.*

Proof. Let $w \in \bigcup_{i \in X} U_i$, where each $U_i \in O$. It suffices to show that there exists $A \in A$ such that $w \in A \subseteq \bigcup U_i$. Picking *i* with $w \in U_i$, and then corresponding $A \in A$ with $w \in A \subseteq U_i$ will do the job.

Proposition 1.26. int *B* is the largest open subset of *B*.

Now we can define closed sets:

Definition 1.27. Given a structure (W, A), we call a set *B* **closed** iff $\neg B$ is open iff $\neg B \subseteq \operatorname{int} \neg B$ iff $B = \operatorname{cl} B$. The interpretation here is that "No matter how *B* is false, it is refuted by some true element of A".

Proposition 1.28. *The collection of closed sets is closed under arbitrary union and* cl B *is the smallest closed set containing B.*

Proposition 1.29. *For any* $B \subseteq W$ *,* ∂B *is closed.*

Definition 1.30. A set *B* is **clopen** iff it's both closed and open iff $\partial B = \emptyset$. The interpretation is that *B* is **decided** by members of *A*.

Definition 1.31. A set *B* is **locally closed** if it's frontier is closed.

Proposition 1.32. *The following are equivalent:*

- 1. *B* is locally closed.
- 2. $B = \operatorname{cl} B \setminus \operatorname{cl} \operatorname{frnt} B$.
- 3. $B = \text{ext frnt } B \setminus \text{ext } B$.
- 4. *B* is the difference of two open sets.
- 5. *B* is the difference of two closed sets.

Proof. We'll do only $(5) \implies (1)$: the rest are straightforward. Towards doing so, we calculate: $\operatorname{cl}\operatorname{frnt}(A \cap B) =$ $= \operatorname{cl}(\operatorname{cl}(A \cap B) \cap \neg(A \cap B))$ $\subseteq \operatorname{cl}(\operatorname{cl}(A \cap B) \cap \operatorname{cl}(\neg A \cup \neg B))$ $\subseteq \operatorname{cl}(A \cap B) \cap (\operatorname{cl} \neg A \cup \operatorname{cl} \neg B)$ $\subseteq \operatorname{cl}(A \cap B) \cap (\neg A \cup \operatorname{cl} \neg B)$ $\subseteq \operatorname{cl}(A \cap B) \cap (\neg A \cup \operatorname{cl} \neg B)$ $\subseteq \operatorname{rtt}(A \cap B)$

1.4 Topologies

Now with these concepts under our belt, we're ready to delve directly into the realm of topology. Our fundamental definition is, of course, the definition of a topology:

Definition 1.33. A topology τ on a set W is a propositional structure on W that is closed under finite intersections and arbitrary unions. In this context, any set $B \in \tau$ is said to be *open*.

Remark 1.34. Typically one also requires \emptyset , $W \in \tau$, but this is actually implied by the properties above if we use the convention of the empty union being empty and the empty intersection to be W.

Remark 1.35. Topological spaces are preserved under restriction, quotient, and co-product, but not necessarily under product. Since the intersection of topologies is a topology, given any propositional structure A, we can define the **topological completion** A^* of A to be the smallest topology containing A. We then define the **topological product** to be the topological completion of the product.

Now we can strengthen the distributivity properties we had earlier:

Proposition 1.36. *The following hold if* (W, τ) *is a topological space,:*

- 1. int $\bigcap_{i \le n} A_i = \bigcap_{i \le n} \text{ int } A_i$ (finite intersection).
- 2. cl $\bigcup_{i < n} A_i = \bigcup_{i < n} cl A_i$ (finite union).
- 3. int W = W.
- 4. $\operatorname{cl} \emptyset = \emptyset$.

Perhaps unsurprisingly, it turns out that given just an interior-type operator, one can naturally recover the topology associated to it, and vice versa. We formalize this as follows.

THEOREM 1.37 Kuratowski

Given an operator int : $\mathcal{P}(W) \rightarrow \mathcal{P}(W)$, the set int[W] is a topology iff the following conditions are satisfied.

- 1. int W = W.
- 2. int $A \subseteq A$.
- 3. $\operatorname{int} \operatorname{int} A = \operatorname{int} A$.
- 4. int $\bigcap_{i < n} A_i = \bigcap_{i < n} \text{ int } A_i$ (finite intersection).

Similarly, given $cl : \mathcal{P}(W) \to \mathcal{P}(W)$, the set $\neg cl[W]$ is a topology iff

- 1. $\operatorname{cl} W = W$.
- 2. $\operatorname{cl} A \supseteq A$.
- 3. $\operatorname{cl}\operatorname{cl} A = \operatorname{cl} A$.
- 4. int $\bigcup_{i < n} A_i = \bigcup_{i < n} \text{ int } A_i$ (finite intersection).

Furthermore, if τ is a topology, the interior and closure operators as defined before satisfy all of the above.

Now we'll take a brief digression to discuss algebras, which will be critical later in our discussions of complexity.

Definition 1.38. A propositional structure A on a set W is an **algebra** if it's closed under negation and finite intersections. It's a σ -**algebra** if these hold for countable intersections. As before, the intersection of algebras remains an algebra, so we can define the algebra generated by any fixed structure A to be the smallest algebra containing A.

Proposition 1.39 (Disjunctive Normal Form DNF). *Every element of the field generated by* (W, A) *can be expressed as a finite union of finite intersections of elements of* $A \cup \neg A$.

Given any topology τ , we call the algebra generated by τ the **collection of constructible propositions**, which is the set of all finite union of locally closed propositions.

Definition 1.40. Given any constructible *B*, we define it's **complexity** to be the number of propositions in its minimal-length topological DNF. The diagram looks something like

this:

This leads naturally to the following complexity classes:

- Π_n^C : the set of propositions with a TDNF of length at most *n* and a closed disjunct.
- Σ_n^C : the set of propositions with a TDNF of length *n* and no closed disjunct.
- $\Delta_n^C := \Pi_n^C \cap \Sigma_n^C$.

These nest in the following way, exactly analogously to the **Borel hierarchy**.



The σ -algebra generated by a topology is called the collection of **Borel propositions**.

Remark 1.41. The enumeration above also serves as a way to measure complexity through **counting retractions**. That is, if τ is the topology of verifiability, elements of Σ_n^C , Π_n^C are precisely the propositions where you can "converge to a correct judgement" while only retracting at most *n* times. For example, if $A \in \Sigma_1^C$ (i.e. if *A* is open), such a strategy would be to declare *A* false until *A* is verified, which will work always. A similar construction works for e.g. locally closed sets, as well as further up the hierarchy.

1.4.1 Topological Bases

Oftentimes, the full structure of a topology is too rich to admit any sort of reasonable description. In order to circumvent this issue, it's frequently useful to work with topologies indirectly through their bases, which frequently admit easier descriptions.

Definition 1.42. We call a propositional structure A a **topological base** if $\cup A = W$ and for any $A, B \in A$, there exists $C \in A$ with $C \subseteq A \cap B$. In this case, we have

$$\mathcal{A}^* = \{ \cup B \mid B \subseteq \mathcal{A} \}$$

(recall that * denotes the topological completion as defined in Remark 1.35). We call a propositional structure A a **topological base for a topology** τ if every open set in τ contains some element of A. If a topology admits a countable base, we call it **second-countable**.

Proposition 1.43. *Every topology is a base for itself.*

Proposition 1.44. *Bases are preserved under product, restriction, quotient, and co-product.*

1.4.2 Separation Properties

Since topological structures can be arbitrarily poorly behaved, it's often convenient to have stronger classifications of topologies that we can restrict to.

Definition 1.45. We make the following definitions, which actually differ slightly from the standard terminology. Throughout, we fix a given topological structure (W, τ) .

- w, v are distinguishable iff $(\exists A \in \tau)(w \in A \text{ iff } v \notin A)$.
- w, v are mutually distinguishable if there exists $A, B \in \tau$ with $w \in A \setminus B$ and $v \in B \setminus A$.
- *A*, *B* are separable if there exists disjoint $A', B' \in \tau$ such that $A \subseteq A'$ and $B \subseteq B'$
- w, v are separable if $\{w\}, \{v\}$ are separable.

When we're dealing with arbitrary topological spaces, it'll be useful to have the following classification of separation properties:

Definition 1.46.

- *T*⁰ (Kolmogorov). Distinct points are distinguishable.
- *R*⁰ (symmetric). Distinguishable points are mutually distinguishable.
- *T*₁ (Frechet). Distinct points are mutually distinguishable.
- *R*₁ (pre-regular). Distinguishable points are separable.

- *T*₂ (Hausdorff). Distinct points are separable.
- Regular: Closed propositions are separable from disjoint singletons.
- T_3 (regular Hausdorff). Regular and T_0 .
- Normal: Disjoint closed propositions are separable.
- T_4 (normal Hausdorff). Normal and T_1 .

Proposition 1.47. *The following hold.*

- $T_1 = R_0 + T_0$
- $T_2 = R_1 + T_0$.
- $R_1 \implies R_0, T_2 \implies T_1 \implies T_0.$
- *T*₁ *is equivalent to every singleton being closed.*
- *Regular implies pre-regular* (*R*₁).
- *Regular Hausdorff implies Hausdorff (T₃ implies T₂).*
- Normal Hausdorff implies Hausdorff (*T*₄ implies *T*₂).
- Normal Hausdorff implies regular Hausdorff (T₄ implies T₃).

Proposition 1.48. Counterexamples showing that all these inclusions are strict:

- $\neg T_0$: $W = \{0, 1\}; \mathcal{B} = \{W\}$ (the indiscrete topology).
- $T_0 \setminus T_1 : W = \{0, 1\}; \mathcal{B} = \{W, \{1\}\}$ (the Sierpinski space).
- $T_1 \setminus T_2 : W = \mathbb{N}; \quad \mathcal{B} = \{A \subseteq \mathbb{N} : \neg A \text{ is finite }\}.$ (the co-finite topology).
- $T_2 \setminus T_3 : W = \mathbb{N}; \quad \mathcal{B} = \{an + b : a, b \in \mathbb{R}; n \in \mathbb{Z}\} \cap \mathbb{N}$
- $T_3 \setminus T_4$: Product of two Sorgenfrey lines (\mathbb{R} topologized by half-open intervals).

2 Context of Inquiry

Now with these topological prerequisites under our belt, we're finally ready to see some applications! In particular, given the title of this class, we'll be interested in how we can use topological ideas to characterize learning and scientific inquiry. The basic premise we'll be working under then is to place any attempt at scientific inquiry into a **context** that frames the questions we're asking and the answers we might hope to receive. For the purposes of this class, we can separate this context into three distinct topological bases that we'll want to consider:

- The **Metaphysical** base \mathcal{M} that measures similarity among worlds.
- The **Empirical** base \mathcal{E} that captures the information or data that can **possibly** be received.
- The **Erotetic** base *Q* that contains the statements of all the questions/answers that we're hoping to investigate.

As per usual, it can be shown that each of these propositional structures is a topological base, and hence generates a corresponding topology $\mathcal{M}^*, \mathcal{E}^*, \mathcal{Q}^*$. We then define a **context of inquiry** as follows.

Definition 2.1. A context of inquiry is a tuple $(W, \mathcal{M}, \mathcal{E}, \mathcal{Q})$ where

- $\mathcal{M}, \mathcal{E}, \mathcal{Q}$ are topological bases on W,
- *E*, *Q* are countable (since both information states and logical propositions are),
- $\mathcal{E} \subseteq \mathcal{M}^*$ (i.e. that any observable difference is real in the metaphysical sense),
- $\mathcal{E} \subseteq \mathcal{Q}^*$ (data can be recorded, which is more of a convenience thing),
- and $Q \subseteq M^*$ (the questions being asked are structural and non-trivial).

Definition 2.2. Empiricism is the belief that real differences are empirical, i.e. that

 $\mathcal{M}\subseteq \mathcal{E}^{*}.$

With all of these elements defined, we'll now go on to explore the basic properties we expect out of each of these topological spaces.

2.1 Metaphysical Basis

2.1.1 Metric Similarity

We first turn our attention to the *metaphysical basis*, which is a way of measuring similarity between worlds. Typically, we'll assume that the corresponding topology is **metrizable**, which will hence motivate the discussion below.

Recall first that we call $\rho : W \times W \rightarrow \mathbb{R}$ a **metric** if it is nonnegative (and satisfies $\rho(x, y) = 0 \iff x = y$), symmetric, and satisfies the triangle inequality. In the context we'll be working with, we'll be interpreting such a metric ρ as a measure of dissimilarity, which naturally leads to the properties of symmetry and nonnegativity above. We also note that the triangle inequality can be interpreted as a bound on the "failure of transitivity" of dissimilarity, so in some sense the metric axioms can be interpreted exactly as just

axioms for similarity. Given any metric, we define the **open ball** of radius r around x as

$$B_{\rho}(x,r) := \{ y \mid \rho(x,y) < r \}$$

for any $x \in W$, r > 0. (When the underlying metric is clear from context the subscript will often be dropped.)

Using standard arguments from analysis then yields the following:

Proposition 2.3. For any metric ρ , the set of all balls $\mathcal{B}_{\rho} := \{B_{\rho}(w, r) \mid w \in W, r > 0\}$ is a basis for a topology (referred to as the **metric topology** for obvious reasons). We denote this topology by $(\mathcal{B}_{\rho})^*$.

Under our assumption of ρ being a similarity metric, we can naturally extend the interpretations we assigned to topological operators to interpretations within the context of the metric topology.

Lemma 2.4. The following interpretations hold.

- 1. The closure of A is the set of worlds that are arbitrarily similar to A's worlds. Alternatively, this is the set of all fine-tuning of A's worlds or just the fact that A is arbitrarily close to being true.
- 2. The frontier of A is the set of worlds where A is false, but is arbitrarily close to being true in the same vein as above.
- 3. The exterior of A is the set of worlds where A is **not** arbitrarily close to being true (i.e. there is a positive degree of similarity that bounds these worlds uniformly away from A).
- 4. Dual to this, the interior of A is the set of worlds where A is not arbitrarily close to being false, i.e. that A is discretely or easily true.

Lemma 2.5. The following also hold.

- *A is closed iff "A is arbitrarily close to being true" implies A, which is equivalent to A being closed under fine-tuning.*
- *A is open iff A entails "A is not arbitrarily close to being false".*
- *A is clopen iff A vs.* $\neg A$ *is a structural (discrete) distinction.*

While having a concrete metric to work with can be nice, having to come up with such a function can often be difficult or present ideological issues (e.g. what matters more, the color of the sky or the mass of the earth?). One of the nice things about the topological approach we've taken is that we can somewhat skirt this issue by just working with the metric topologies, discarding the concrete values that are provided by the metric. To this end, we make the following distinction.

Definition 2.6. We call a topological space (W, τ) **metrizable** if there exists a metric ρ such that $\tau = (\mathcal{B}_{\rho})^*$.

Given two metrics, it's also often useful to have criterion to check whether or not they generate the same underlying metric topology.

Proposition 2.7. Two metrics ρ , ρ' are equivalent iff they generate the same topology iff every ρ ball contains a ρ' ball and vice versa.

Proposition 2.8. *If* ρ *is a metric, then so is* $\rho \wedge 1$ *and the two are equivalent.*

Proof. Clearly every $\rho \wedge 1$ ball contains a ρ ball of the same radius. On the other hand any ρ ball of radius r contains a $\rho \wedge 1$ ball of radius $r \wedge 0.5$.

It turns out also that metrizable topologies are nicer to work with because they admit stronger separation properties.

Proposition 2.9. Metrizable implies normal Hausdorff.

Proof. First let $x, y \in W$ be distinct, and $d := \rho(x, y) > 0$. The two open sets B(x, d/4), B(y, d/4) then witness that τ is Frechet.

Now suppose $A, B \subseteq W$ are closed. For every $x \in A \subseteq B^c$, since B is closed, there exists $r_x > 0$ with $B(x, r_x) \subseteq B^c$, with a similarly argument working for all $y \in B$. The two open neighborhoods

$$\bigcup_{x \in A} B(x, r_x/4), \bigcup_{y \in A} B(y, r_y/4)$$

thus show normality.

Proposition 2.10. Assuming second countability, metrizability is equivalent to being T_3 .

The following slightly stronger theorem also holds.

Proposition 2.11. *Metrizability is equivalent to being* T_3 *and having a* σ *-locally finite base.*

Proposition 2.12. *The indiscrete topology on any set with more than two elements is not metrizable, nor is the Sierpinski topology.*

Example 2.13

- The discrete topology is generated by the metric $\rho(x, y) = \mathbf{1}_{x \neq y}$. This corresponds to a universe where there are no arbitrary similarities between worlds.
- \mathbb{R} equipped with the Euclidean metric.
- \mathbb{R}^n equipped with the Euclidean metric.

- Suppose {(W_i, ρ_i)}_{i∈ℕ} are metric spaces, with sup_i ρ_i < ∞. Then we can equip the product space ∏_i W_i with the metric ρ(**x**, **y**) := ∑ 2⁻ⁱρ(x_i, y_i), which generates the product metric topology. (This construction can be used to "merge topologies" if e.g. similarity with respect to multiple parameters are being considered).
- A specific instance of above: the Cantor space is the countable product of {0, 1} with the discrete metric.
- Using proposition 2.8 and the above, we also see that metrizable spaces are closed under countable topological product.
- Any vector space equipped with a norm: e.g. function spaces equipped with the *L^p* norms (this example will be relevant for the contexts of curve fitting/the description of scientific laws).

2.2 The Empirical Basis

2.2.1 Information Gathering

Now we turn to the empirical basis, which gives us a way to formalize the process of **information gathering** and the tasks achievable with it. The basic setup is as follows:

- Every world *w* ∈ *W* comes attached with a set *S_w* of possible situations or states that the world can be in.
- Every such situation comes attached with a signal that rules out worlds in which the signal is not possible.
- We denote the set of worlds that are not ruled by such a situation *s* by $E_{w,s}$.

Axiom 2.14 (Information is True). $w \in E_{w,s}$ for all signals *s* possible in *w*.

Axiom 2.15 (Truth implies possiblity). For all $v \in E_{w,s}$, there is some situation $r \in S_v$ with $E_{v,r} = E_{w,s}$.

Argument. By definition, the signal produced by w in situation s does not rule out v, and hence v must produce the same signal in some situation r, which is all we wanted.

We now set $\mathcal{E} := \{E_{w,s} \mid w \in W, s \in S_w\}$ as the definition of our empirical basis.

Proposition 2.16 (Truth is possiblity). $\mathcal{E}_w = \{E_{w,s} \mid s \in S_w\}.$

Proof. It suffices to show \subseteq since \supseteq is clear. Fixing any such $E \in \mathcal{E}_w$, there must exist $v \in W$, $r \in S_v$ with $E = E_{v,r}$. The fact that truth implies possibility concludes.

Axiom 2.17. $W = \bigcup \mathcal{E}$.

Argument. The signal of "no information" rules out nothing, hence must correspond to all of *W*.

After establishing these axioms, which should be relatively uncontroversial, it's time to come to one that might be more hotly contested: that is, the axiom of **diligence**, which practically just says that all information that is possible to receive will eventually be received with due diligence. Towards this end, we make the following definitions.

Definition 2.18. $E \ni w$ is *lucky* if it is *possible* to obtain information as strong as E in w. $E \ni w$ is *achievable* if information as least as strong as E will always be achieved in w.

Axiom 2.19 (Diligence). *All information is achievable.*

Axiom 2.20 (Cumulativity). *Assuming diligence,* \mathcal{E}_w *is a topological basis.*

Proof. If $A, B \in \mathcal{E}_w$, then in w, you will eventually achieve information at least as strong as $A \cap B$. This information C is precisely the desired subset of $A \cap B$.

2.2.2 Measurement and its Consequences

EXAMPLE 2.21 Measurement

Suppose the question you're interested in answering is the value of some real-valued quantity *X*. Then a natural information basis for *X* would be the set of open intervals (containing the true value of *X*), which corresponds to the standard topology on \mathbb{R} . The same analysis naturally extends to the measurement of multiple parameters *X*, *Y*, *Z*, ··· .

EXAMPLE 2.22 Curve Fitting

Suppose the question you're interested in answering is the value of some function $f : \mathbb{R} \to \mathbb{R}$. Then a natural information basis would be finitely many observations of the function f and its values, i.e. open intervals/rectangles around (X, f(X)). In this setup, information states are then just the set of all functions incident to each observation, which necessarily generates a very weak topology.

Proposition 2.23. *The topology above is not* T_0 *over arbitrary functions but is* T_1 *over continuous functions.*

Proof. Any two functions that differ at only a single point are not distinguishable, showing failure of T_0 . Now let f, g be distinct continuous functions. Since they're not equal, they must disagree at some point; up to rescaling and translation (which preserve the topology), we might as well assume that f(0) = 0, g(0) = 1. By continuity, we can find $\varepsilon > 0$ such that f < 1/4, g > 3/4 on $[-\varepsilon, \varepsilon]$. Setting $A = (-\varepsilon, \varepsilon) \times [-1/4, 1/4], B = (-\varepsilon, \varepsilon) \times [3/4, 5/4]$ thus witnesses that $f \in A \setminus B, g \in B \setminus A$ as desired.

Proposition 2.24. *Lipschitz functions are dense in this topology; that is, given any finite set of rectangles, there exists a lipschitz function passing through all of them.*

Proof. Given any finite list of rectangles, let (x_i, y_i) enumerate a sequence of points with increasing *x* coordinate such that each rectangle contains one such point. The function that connects this sequence of points to each other via straight lines is Lipschitz and in the corresponding open set. Since collections of such finite lists form a basis, this shows the claim.

Proposition 2.25. No sets are separable in this topology, and hence this topology is not Hausdorff.

Proof. The above shows that any two open sets intersect nontrivially.

Proposition 2.26. *No* L^p *ball is open in the restriction of this topology to the space of continuous functions.*

Remark 2.27. Note that the product topology on $\mathbb{R}^{\mathbb{R}}$ is a refinement of this one, since e.g. a base for the former would be $\{f \mid f(x) \in (a, b)\}$ which contains a base for the latter in it's generated topology. Furthermore, this containment is easily seen to be strict; letting *x*, *a*, and *b* as above and $y \notin (a, b)$. Then the same construction as in Proposition 2.24 yields a function *f* in an arbitrary open set but not in $\{f \mid f(x) \in (a, b)\}$.

2.2.3 Between Topology and Empiricism

We begin this section with some terminology, which will help us transition between topology and information gathering.

Definition 2.28. Let *A* be a proposition and *E* an information state.

- *E* verifies *A* if $E \subseteq A$.
- *E* refutes *A* if $E \subseteq A^c$.
- *E* decides *A* if either $E \subseteq A$ or $E \subseteq A^c$.

As a consequence, we see that *A* will be verified in world *w* precisely when

$$w \in \bigcup \{ E \in \mathcal{E} \mid E \subseteq A \} = \operatorname{int} A.$$

So we have an interpretation of the interior operator as a modal operator corresponding to "will be verified." Indeed, we see that this interpretation makes sense in light of the Kuratowski axioms:

- 1. Extensivity: *A* cannot be verified unless true.
- 2. Idempotency: "*A* will be verified" is verified as soon as *A* is.
- 3. If a conjunction is verified, so is each conjunct. Conversely, if each conjunct will be verified, then eventually all of them will be.

Remark 2.29. It's also easily checked directly that verifiability satisfies the topological axioms.

As a consequence, we see that empirical verifiability is precisely openness in the empirical topology.

2.2.4 The Problems of Induction and Metaphysics

We begin this section with a motivating example.

Example 2.30

Consider any of the "standard setups for induction," that is, you have some hypothesis you think to be true (e.g. the sun will rise tomorrow, bread will nourish tomorrow, all crows are black, etc.), and you're attempting to find some way to verify or justify this belief. A reasonable way of framing this within the framework that we've been working in is to say that reality is some infinite sequence $(a_0, a_1, \dots) \in 2^{\omega}$ corresponding to observations of the hypothesis you're dealing with. Information states will then be finitely many observations of such a sequence, corresponding to "cones" of prefixes within the standard Cantor space and it's associated topology. Throughout science, many different theories and hypotheses have been verified in a setup similar to this one, i.e. seeing long sequences of 1s and taking that to indicate that the hypothesis is true in general. However, doing anything in this way is inherently fallible, since any finite amount of information cannot capture the entire truth, an issue brought up by Sextus, Hume, and Popper.

Definition 2.31. Given a proposition *A*, you are said to face the **problem of induction** for *A* if *A* is true but you will never be able to verify it.

Remark 2.32. This set is exactly frnt $\neg A$ by definition 1.22.

Dual to this issue is the problem of metaphysics, also addressed by the likes of Hume, Kant, and Popper.

Definition 2.33. Given a proposition *A*, you are said to face the **problem of metaphysics** for *A* if *A* is false but you will never be able to refute it. Similarly to above, this means that the problem of metaphysics for *A* is faced in exactly frnt *A*.

Proposition 2.34. *The problem of induction for* A *is exactly the problem of metaphysics for* $\neg A$ *.*

Proposition 2.35. *A* is closed iff frnt $A = \emptyset$ iff you never face the problem of metaphysics for A. (In other words, A is refutable).

With these definitions in hand, we can now revisit the motivating example above. If we quotient out by the answers we're looking for, letting *A* be the worlds in which the hypothesis is true (i.e. that the true world is all ones) and *B* be the set of worlds for which there is some zero entry in the sequence, we actually recover the Sierpinski space, with the corresponding partitions of the space as below: (sorry, I got too lazy to make my own diagram)



Performing the same quotient business, we can also see that this situation actually exactly captures a "point null hypothesis" in the measurement problem over \mathbb{R} or an "exact law" problem in curve fitting, where we can let $A = \{X = 0\}, B = \neg A$, with the corresponding problems of induction/metaphysics matching exactly to the one seen above.

Remark 2.36. Above, we saw that reputability is sufficient to guarantee avoiding the problem of metaphysics, but as we'll show now, that's actually not strictly necessary.

Definition 2.37. *A* is **verifutable** if *A* entails that it will be verified that *A* is refutable.

So *A* is verifutable iff frnt *A* (the problem of metaphysics) is refutable (closed), which is equivalent to frnt frnt $A = \emptyset$, which in turn is equivalent to *A* being locally closed.

Definition 2.38. An information state *E* is a **trigger** for *A* if *A* is refutable given *E*; i.e., that $E \subseteq \neg \operatorname{frnt} A$. We thus have that the **canonical trigger** $\operatorname{trg} A := \neg \operatorname{frnt} A$.

Definition 2.39. A verifutable proposition *A* is **live** given *E* is *A* is triggered but not refuted by *E*.

Example 2.40

Consider the same Cantor space example as before, except where we are now interested in counting precisely how many times 0 shows up. If we let $A = \{1\}, B = \{f : |f^{-1}(0)| = 1\}, C = (A \cup B)^c$, we get the following:



The canonical triggers for these sets are then $\operatorname{trg} A = W$, $\operatorname{trg} B = \neg A$, $\operatorname{trg} C = C$.

Remark 2.41. A similar game can be played for any collection of sets $\{A_i\}_{i \in \mathbb{N}}$ with $A_i \subseteq A_{i+1}$ and each $\bigcup_{i>n} A_i$ open. For example, this same example works for counting polynomial degree, number of free parameters, etc. In this case, each A_i is necessarily locally closed and hence verifutable, and each trg $A_{i+1} = A_i$.

2.2.5 Distribution Laws and Their Failures

As remarked previously, the asymmetry in topology between the intersection and union operations actually makes a lot of sense in the context of verifiability, and has deep connections to the issues discussed in the previous section. In particular, usually the problem of induction is understood in terms of infinite conjunctions. For example, you can verify for each day that the sun will rise, but you will never verify that the sun always rises. (Formally, int $\bigcap_{i \in I} A_i \not\supseteq \bigcap_{i \in I} \operatorname{int} A_i$, for infinite *I*..

Definition 2.42. You face the problem of induction for the collection $\mathcal{A} = \{A_i : i \in I\}$ in w due entirely to infinite conjunction if $w \in \bigcap_{i \in I} \operatorname{int} A_i \setminus \operatorname{int} \bigcap_{i \in I} A_i$.

So the failure of distributivity of interior over conjunction is entirely captured by the problem of induction due to infinite conjunction. We can also define a similar operator for the failure of distributivity of refutability over conjunction.

Definition 2.43. You face the Duhem-Quine problem for the collection $\mathcal{A} = \{A_i : i \in I\}$ in w if $\cap \mathcal{A}$ will be refuted in w but no single element of \mathcal{A} will be.

Parsing through the logic, we see that this holds if

$$w \in \operatorname{ext}\left(\bigcap_{i \in I} A_i\right) \setminus \bigcup_{i \in I} \operatorname{ext} A_i = \neg \bigcup_{i \in I} \operatorname{ext} A_i \setminus \neg \operatorname{ext}\left(\bigcap_{i \in I} A_i\right)$$
$$= \bigcap_{i \in I} \operatorname{cl} A_i \setminus \operatorname{cl} \cap_{i \in I} A_i$$

and hence this problem is precisely faced $\bigcap_{i \in I} \operatorname{cl} A_i \setminus \operatorname{cl} \bigcap_{i \in I} A_i$, and closure distributes over conjunction "up to" the Duhem-Quine problem. Dualizing again, we have the following

Definition 2.44. You face Co-Duhem's problem of credit assignment for the collection $\mathcal{A} = \{A_i : i \in I\}$ in w if $\cup \mathcal{A}$ will be verified in w but no element of \mathcal{A} will be, which is equivalent to

$$w \in \operatorname{int}\left(\bigcup_{i\in I} A_i\right) \setminus \bigcup_{i\in I} \operatorname{int} A_i.$$

2.2.6 Simplicity

According to Popper, "The epistemological questions which arise in connection with the concept of simplicity can all be answered if we equate this concept with degree of falsifiability." Towards this end, we make the following definitions.

Definition 2.45. A potential falsifier of *A* is an information state incompatible with *A* i.e., an information state that entails $\neg A$. We say that *A* is **as simple** as *B* if every potential falsifier of *B* is a potential falsifier of *A*.

Remark 2.46. This is just the standard topological specialization pre-order on propositions. Indeed,

Each potential falsifier of *B* is a potential falsifier of *A*

$$\iff$$
 each information state that entails $\neg B$ entails $\neg A$
 \iff int $\neg B \subseteq \neg A$
 $\iff A \subseteq cl B.$

Some virtues of this definition are that it is a pre-order on propositions and makes sense in many standard cases right (e.g., polynomial degree, number of free parameters, number of particles, etc.).

However, a definition like this also comes with it's own downsides. In particular, with this definition:

- Logical entailment implies simplicity: that is, *A* ∩ *B* is always as simple as *A*, which also means that the tautology is maximally complex!
- Also, the inclusion relation is quite crude; if $A \subseteq \operatorname{cl} B, C \not\subseteq \operatorname{cl} B$, then $A \cup C$ has no simplicity relationship to B.

In order to circumvent this, we can consider alternative definitions of simplicity.

Definition 2.47. Given propositions *A*, *B*, we say that you face the **problem of induction** from *A* to *B* if you are under pressure to conclude *A* instead of *B* but whenever you do so, you risk error in a *B* world. In other words, this is the proposition

$$A \triangleleft B := A \backslash B \cap \operatorname{cl}(B \backslash A).$$

We can use this new concept to define a better simplicity order on propositions.

Definition 2.48. We say that *A* is simpler than *B* given *E* if *E* does not rule out some problem of induction from *A* to *B*. In symbols, $(A \triangleleft B) \cap E \neq \emptyset$. In this case, we write $A \ll_E B$.

We say $A \ll B$ if $A \ll_W B$.

Definition 2.49. We say that *B* is **Ockham** given *E* if no answer *A* is simpler than *B* given *E*. That is,

$$\begin{aligned} \operatorname{Ock}(B \mid E) & \Longleftrightarrow \ (\forall A \in \mathcal{Q}) \neg (A \ll_E B) \\ & \longleftrightarrow \ (\forall A \in \mathcal{Q}) E \cap (A \triangleleft B) = \emptyset \\ & \Leftrightarrow \ E \cap \bigcup_{A \in \mathcal{Q}} (A \triangleleft B) = \emptyset \\ & \Leftrightarrow \ E \setminus \cap_{A \in \mathcal{Q}} (A \triangleleft B) = \emptyset \\ & \Leftrightarrow \ E \subseteq \bigcap_{A \in \mathcal{Q}} (A \triangleleft B). \end{aligned}$$

Remark 2.50. Any closed proposition is Ockham.

2.3 The Erotetic Basis

We finally turn to discussion of the erotetic basis Q, the last element of our context of inquiry. As mentioned briefly above, elements of such a basis are meant to be **target**

conclusions for inquiry, meaning that, in an ideal world, they should be concluded eventually if true. Of course, doing so is not always possible, as we'll investigate more in the next section. As per usual, we also have a convenient interpretation that allows us to pass between topological and methodological concepts.

Lemma 2.51. The following interpretations hold.

- int *A* is the set where "You will have to conclude that *A* holds."
- ext *A* is the set where "You will have to conclude that $\neg A$ holds."
- *clA* is the set where "You will never have to conclude that $\neg A$ "
- ∂A is the set where "A will never need to decide A."
- frnt *A* is the set where "*A* is false, but you never need to conclude so."

3 Methodology

3.1 Epistemic Virtues of Deduction

Now that we've gotten the context of inquiry under our belts, we're ready to formalize the actual process of inquiry, which in some sense can just be thought of as "a function from information states to conclusions." Formally, we say the following.

Definition 3.1. Given a context $(W, \mathcal{M}, \mathcal{E}, \mathcal{Q})$, we call any function $L : \mathcal{E} \to \sigma(\mathcal{Q})$ a **propositional method** and any function $L : \mathcal{Q} \times \mathcal{E} \to \sigma(\mathcal{Q})$ a **reasoned method**. We interpret L(A | E) to mean "the reason for concluding *A* given *E*."

Given a propositional method *L*, we can naturally lift it to a reasoned method by just ignoring the question at hand; that is, by setting $\check{L}(A | E) = LE$. Similarly, given a reasoned method *L*, we can restrict it to a propositional method by simply intersecting the reason given for all answers, i.e. by setting $\hat{L}E = \bigcap_{A \in Q} L(A | E)$.

Example 3.2

- (Conclude Nothing): The constant function $A \mid E \mapsto W$ is a method.
- (The Complete Deductive Method): The identity *A* | *E* → *E* is a method, which is basically just "conclude the exact amount of information you have."

Despite seeming somewhat trivial, the complete deductive method actually comes with a number of **epistemic virtues** that we might seek to replicate in other methods. We start by defining these below:

Definition 3.3. We say that a reasoned method *L* **avoids false reasons** if it never says a false reason; that is,

$$\forall w \in W, A \in \mathcal{Q}_w E \in \mathcal{E}_w, w \in L(A \mid E)$$

Definition 3.4. We say that a method *L* **answers the question** if, given **any** true answer in a specific world, the method eventually receives enough information to conclude that answer (or something possibly stronger) from then onward **for the same reason**. In other words, we have that for all $w \in W$, $A \in Q_w$ that there exists $E \in \mathcal{E}_w$ such that

$$F \in \mathcal{E}_w, F \subseteq E \implies L(A \mid F) \subseteq L(A \mid E) \subseteq A.$$

"For every true answer, you eventually believe it."

Definition 3.5. *L* **deductively solves** a problem (W, \mathcal{M} , \mathcal{E} , \mathcal{Q}) if *L* avoids false reasons and answers the question.

Definition 3.6. We say that a method *L* is **consistent** if

$$\hat{L}E \cap E \neq \emptyset \quad \forall E \in \mathcal{E}_w$$

Definition 3.7. We say that a method *L* **remembers the data** if it always concludes something **at least** as strong as the information it receives: that is,

$$L(A \mid E) \subseteq E \quad \forall A \in \mathcal{Q}, E \in \mathcal{E}$$

Definition 3.8. We say that a method *L* is **monotonic** if, given any further information it always concludes something **at least** as strong as the conclusion it's made before **for the same reason**; that is,

$$A \in \mathcal{Q}E, F \in \mathcal{E}_w, F \subseteq E \implies L(A \mid F) \subseteq L(A \mid E).$$

Remark 3.9. Learning theory is non-monotonic: more information may cause you to drop previous conclusions. Deduction is monotonic, however: new information never makes you retract your previous conclusions. This is controversial in statistics. Monotonicity is a major upsell for deduction.

Proposition 3.10. The complete deductive method

- 1. Avoids false reasons,
- 2. Answers the question (if it's deductive!),
- 3. Is consistent,
- 4. Remembers the data,
- 5. and is monotonic.

Given the proposition above, the only thing that might seem lacking is the conditional on 2: that is, the fact that this method does not always answer the question. It turns out that in **inductive** problems, it will actually be impossible to answer the question while also maintaining all the other virtues, which we investigate now.

Theorem 3.11

The following are equivalent:

- 1. Each answer in Q is open (verifiable) in \mathcal{E}^* .
- 2. The complete deductive method has all of the deductive virtues.
- 3. Some method answers the question and avoids error.

In either (and hence all) of these cases, we call the problem **deductive**.

Proof. (1) \implies (2): Fix $w \in W, A \in Q_w$. Since *A* is open, we know that there exists $E \in \mathcal{E}$ such that $w \in E \subseteq A$. For any $F \subseteq E$ with $F \in \mathcal{E}_w$, we clearly have that $L(A | F) = F \subseteq E \subseteq A$ as desired. (2) \implies (3): this is obvious. (3) \implies (1): Fix any such *L* and $w \in A \in Q$. Since *L* answers the question, we can find E_w with $LE_w \subseteq A$. Now for any $v \in E_w$ we know that $E_w \in \mathcal{E}_v$ by definition of the latter, so since *L* avoids error, we must have $v \in LE_w \implies E_w \subseteq A$. We thus see that every element of *A* admits a neighborhood of it contained in *A*, which shows that *A* is open.

With the result above then, we obtain a complete classification of problems that can be solved with all of our desired deductive virtues. The issue, of course, is that not all the problems that we're interested are solving fall into this category; indeed, given our previous discussion on the problems of induction and metaphysics, even the most basic of questions have no hope of ever being deductive.

The practical conclusion we have to make, then, is that we use induction as a method because we have to: if the problem is not deductive, we have to change our notion of success.

In particular, we'll see that giving up avoiding false reasons and being strictly monotonic in exchange for actually answering the question is a reasonable compromise that we can usually come to. Throughout the remainder of this section, then we'll see that the notion of success that we come to normally will naturally enforces bounds on the topological complexity of the question and vice versa. **Definition 3.12.** We say that a method *L* **eliminates error** if, whenever it draws an incorrect conclusion, there is further information that will cause it to "see it's error". Formally, this is encoded by the assertion that, for all $w \in W$, $A \in Q_w$, $E \in \mathcal{E}_w$

$$w \notin L(A \mid E) \implies (\exists F \in \mathcal{E}_w; F \subseteq E) (\forall G \in \mathcal{E}_w; G \subseteq F) L(A \mid G) \nsubseteq L(A \mid E).$$

EXAMPLE 3.13 Finite or infinite

Consider the following context, which corresponds to asking the question "is the world finite" with possible answers "yes" or "I don't know":

- $W = \omega + 1$
- $\mathcal{E} = \{n^+ \mid n \in \omega\}$
- $\mathcal{Q} = \{\omega, W\}.$

where $n^+ := \{ \alpha \in W \mid \alpha > n \}$. Then the reasoned method

$$\omega \mid n^+ \mapsto \{n\}, W \mid n^+ = n^+$$

answers the question and eliminates false reasons.

Remark 3.14. It may be possible for a reasoned method to answer the question when no propositional method can do so.

How bad is convergence? The restrictions we've placed above still allow for arbitrarily many returns to a false answer, which seems bad. One such measure of convergence could be the following, which will actually turn out to be too strong:

Definition 3.15. We say a method **eliminates false answers** if, for all $w \in W$, $A \in Q$,

$$w \notin A \implies (\exists E \in \mathcal{E}_w) \, (\forall F \in \mathcal{E}_w; F \subseteq E) \, L(A \mid F) \notin A,$$

that is, that the method eventually receives enough information to never conclude A again.

EXAMPLE 3.16 Finite or infinite redux

Consider the following context:

- $W = \omega + 1$
- $\mathcal{E} = \{n^+ \mid n \in \omega\}$
- $A = \{n \in \omega \mid n \text{ is even}\}$

• $\mathcal{Q} = \{A, W\}.$

In this context, it is impossible to answer the question and to eliminate false answers, but it is possible to answer the question, eliminate false conclusions, and eliminate false reasons.

Proof. To see the first point, suppose that arbitrary *L* answers the question. Then for each even $n \in \omega$, $Ln^+ \subseteq A$. So *L* does not eliminate false *A* in world ω . To see the second part, simply observe that the same method from last time works.

Definition 3.17. We say that a reasoned method *L* **preserves true reasons** if, for all $w \in W$, $A \in Q$, $E, F \in \mathcal{E}_w$, $F \subseteq E$ that

$$w \in L(A \mid E) \implies L(A \mid F) \subseteq L(A \mid E).$$

Another measure: No reversal without refutation.

Definition 3.18. We say that a reasoned method *L* **preserves non-refuted reasons** if, for all $A \in Q, E, F \in \mathcal{E}_w, F \subseteq E$ that

 $F \cap L(A \mid F) \neq \emptyset \implies L(A \mid F) \cap F \subseteq L(A \mid E).$

Proposition 3.19. *The two definitions above are equivalent.*

Proof. ⇒: fix *E*, *F* ∈ *E*; *F* ⊆ *E* with $L(A | E) \cap F \neq \emptyset$. Then there exists $w \in L(A | E) \cap F$. Since *L* preserves true reasons, $L(A | F) \subseteq L(A | E)$. \Leftarrow : fix *E*, *F* as before and $w \in L(A | E)$. Then $L(A | E) \cap F \neq \emptyset$, but since *L* preserves non-refuted reasons, $L(A | F) \subseteq L(A | E)$.

Definition 3.20. We say that a reasoned method *L* **avoids self-defeating induction** if it is deductive as possible; that is, for all $A \in Q, E$,

$$L(A \mid E) \subseteq \operatorname{int} A \implies E \subseteq A$$

3.2 Topological Characterizations of Inductive Solvability

Recall the complexity classes from Definition 1.40. We augment these by defining Σ_2° to be the set of countable disjunctions of locally closed sets and \prod_2° to be the same for co-locally closed sets.

Proposition 3.21. *The following are equivalent.*

1. Some reasoned method has all of the inductive virtues.

- 2. Some reasoned method answers the question and eliminates false reasons.
- *3. Each answer to* Q *is* Σ_2° *in* \mathcal{E}^* *.*
- 4. For each answer $A \in Q$, frnt $\neg A$ is Σ_2° in \mathcal{E}^* .

Proof. (1) \implies (2): trivial. (2) \implies (3): Fix $w \in W$, $A \in Q_w$. Since *L* answers the question, there exists $T_w \in \mathcal{E}_w$ such that $L(A | F) \subseteq L(A | T_w) \subseteq A$. for all $F \in \mathcal{E}_w$ with $F \subseteq T_w$. Now let $\mathcal{D}_w = \{ D \in \mathcal{E} : D \subseteq T_w; L(A \mid F) \notin L(A \mid T_w) \}$ and $\mathcal{D}_w = \bigcup \mathcal{D}_w$. Note that: i. T_w , $D_w \in \mathcal{E}$; $D_w \subseteq T_w$ by construction. ii. $w \in T_w \setminus D_w$, by choice of T_w . iii. $T_w \setminus D_w \subseteq A$, since if not, there would be $v \in T_w \setminus D_w \cup A$, which then implies the existence of $E_v \in \mathcal{E}_v$ with $L(A \mid F) \not\subseteq L(A \mid T_w)$ for all *F* as before, since *L* eliminates false reasons. Then letting $G_v \in \mathcal{L}_v$ be such that $G_v \subseteq E_v \cap T_w$, we would have $v \in$ $G_v \in \mathcal{L}_w$, so $v \in \cup \mathcal{D}_w = D_w$, contradiction. Now let $\mathcal{T}_A = \{(T_w, D_w) : w \in A\}$. \mathcal{T}_A is countable, since \mathcal{E} is and \mathcal{D}_w depends only on $T_w \in \mathcal{E}$, Then $A = \bigcup_{(T,D)\in \mathcal{T}_A} T \setminus D$, by (ii) and (iii), so we're done. (3) \implies (4): Suppose that A is Σ_2° in \mathcal{E}^* , so that there exists $A = \bigcup_{i < \omega} O_i \cap C_i$ for open O_i and closed C_i . Then $A = \operatorname{int} A \cup \operatorname{frnt} \neg A = \operatorname{int} A \cup \bigcup_{i < \omega} O_i \cap C_i \cap \neg \operatorname{int} A$. $O_i \cap C_i \cap \neg \operatorname{int} A$ is locally closed. So frnt $\neg A$ is Σ_2° in \mathcal{E}^* as desired.

Proposition 3.22. The following are equivalent.

- 1. Some method answers the question, eliminates false conclusions, eliminates false answers, is consistent, remembers the data, preserves truth, and produces only locally closed conclusions.
- 2. Some method answers the question and eliminates false answers.
- *3. Each answer to* Q *is* Δ_2° *in* \mathcal{E}^* *.*
- 4. Each answer to Q^{\pm} is Σ_2° in \mathcal{E}^* .
- 5. Each answer to Q^{\pm} has a cover by propositions of form $T \setminus D$, where $T \in \mathcal{E}$, $D \in \mathcal{E}^*$, $D \subseteq T$.

3.3 Ockham's Razor and its Epistemic Justification

Definition 3.23. We call a propositional method **Ockham** if it always draws an Ockham conclusion; that is, that it doesn't exclude answers simpler than non-excluded answers.

Definition 3.24. We call a propositional method **patient** if every Ockham answer is left open by *L*; that is,

 $(\forall E \in \mathcal{E})(\forall A \in \mathcal{Q} : \operatorname{Ock}(A \mid E))A \cap E \subseteq LE.$

This is a nice property that we'd like to have, but patience prevents you from actually answering some questions (such as two paradigms or the cofinite topology on the natural numbers).

Definition 3.25. We call a propositional method **gonzo** if it always rules out any complex answer; that is, for all $E \in \mathcal{E}$,

$$(\forall A \in Q : \neg \operatorname{Ock}(A \mid E))A \cap E \subseteq \neg LE.$$

Proposition 3.26. *Patient gonzo methods are Ockham, but the other implications do not hold.*

Why use Ockham's razor? Here's a simple argument as to why enforcing it might not be productive: If you already know the world is simple, then you don't need Ockham's razor, but if you don't already know the world is simple, then how could a simplicity bias help you in complex worlds?

An answer to this can be given through the following theorem:

Theorem 3.27 (Ockham Necessity). *Any propositional method that answers the question, eliminates error, and is inductively monotonic satisfies Ockham's razor.*

Since satisfying the first two properties is possible only when all three are, you ought to satisfy the last one as well and, hence, you ought to follow Ockham's razor!

Proof. Suppose that answer B = LE is not Ockham given E, so there exists answer A such that $A \triangleleft_E B \neq \emptyset$. Let $w \in E \cap A \backslash B \cap \partial B \backslash A$. Since L eliminates error and answers the question, there exists $F \subseteq E$ such that $w \in F, LF \subseteq A$, and $LF \nsubseteq B$. But there also exists $v \in F \cap B \backslash A$, since $w \in \partial B \backslash A$. So from E to F in v, L both retracts B and falls into error A.

3.4 Negligibility

When doing science, it's often the case that we want to ignore conclusions that are "too unlucky" to be true in the name of making progress. As a motivating example, consider the planet Venus, which previously was thought by the Egyptians to be two separate entities during the sunset/sunrise (i.e., the morning and evening star). After measuring the size and speed of these two entities and finding that they were the same, Pythagoras reasonably concluded that the two were in fact the same object, reasoning that two different objects with identical sizes and speeds would be "too unlucky" to be possible. There are thus two types of negligibility that we'll be considering:

- Credal (Bayes). What you are sure won't happen doesn't matter to you. But then why are you so sure? That invites a regress argument.
- Ontic (Lewis). World proximity is presupposed in the truth conditions of subjunctive conditionals. So negligibility in terms of world-proximity is semantic rather than credal.

In order to formalize a notion of negligibility, then, we'll be turning again to the language of set theory.

Definition 3.28. A propositional structure N is called an **ideal** if

- $\emptyset \in \mathcal{N}$ (i.e., that the empty set is negligible).
- $A \subseteq B, B \in \mathcal{N} \implies A \in \mathcal{N}$ (less than negligible is negligible).
- $A, B \in \mathcal{N} \implies A \cup B \in \mathcal{N}$ (finite unions of negligible things remain negligible).

We can similarly define σ -ideals to be structures where the unions in the third point are allowed to be countable.

EXAMPLE 3.29 **Probabilistic Neglibility**

Suppose $(W, \mathcal{F}, \mathbb{P})$ is a complete measure space (i.e. so that all subsets of null sets are measurable). Then $\{A \subseteq W \mid \mathbb{P}(A) = 0\}$ is a σ -ideal.

EXAMPLE 3.30 Geometric Neglibility

When \mathbb{P} from above is set to be Hausdorff measure, the ideal above can be considered the collection of "geometrically negligible sets."

EXAMPLE 3.31 Topological Neglibility

Call $A \subseteq W$ nowhere dense if int cl $A = \emptyset$. Then the set of nowhere dense propositions is an ideal.

Proof. The only non-trivial thing to check is closure under union. Indeed, given *A*, *B* nowhere dense, we have

$$\operatorname{int} \operatorname{cl}(A \cup B) = \operatorname{int}(\operatorname{cl} A \cup \operatorname{cl} B) \subseteq \operatorname{int} \operatorname{cl} A \cup \operatorname{int} \operatorname{cl} B = \emptyset$$

as desired.

Definition 3.32. Call $A \subseteq W$ meager if $A = \bigcup_{i \in \mathbb{N}} A_i$ with each A_i nowhere dense, and say that a space is **Baire** if every non-empty open set is non-meager.

Theorem 3.33 (Baire Category). *Every completely metrizable space is Baire*.

To prove this result, we'll first need some auxiliary claims.

Lemma 3.34. Let (X, ρ) be a complete metric space with A nowhere dense. Then $(\overline{A})^c$ is open and dense in X.

Proof. It suffices to demonstrate that $(\overline{A})^c$ is dense in X since this set is open by fiat. Suppose towards a contradiction that the above set is not dense in X. Then there exists $x \in X$, $\varepsilon > 0$ such that $B(x, \varepsilon) \cap (\overline{A})^c = \emptyset \implies B(x, \varepsilon) \subseteq \overline{A} \implies x$ is an interior point of $(\overline{A}) = \overline{A}$, a contradiction.

Lemma 3.35. Let (X, ρ) be a complete metric space, and suppose that for each $n \in \mathbb{N}$, $U_n \subseteq X$ is open and dense in X. Then $\bigcap_{n=0}^{\infty} U_n$ is dense in X.

Proof. Let $x \in X$, ε be arbitrary. It suffices to find $u \in B(x, \varepsilon) \cap (\bigcap_{n=0}^{\infty} U_n)$. By the shrinking closed set property in complete metric spaces, it suffices to find x_0, x_1, x_2, \cdots and $\varepsilon_0, \varepsilon_1, \varepsilon_2, \cdots$ such that

- $B(x,\varepsilon) \supseteq \overline{B(x_0,\varepsilon_0)} \supseteq \overline{B(x_1,\varepsilon_1)} \supseteq \overline{B(x_2,\varepsilon_2)} \supseteq \overline{B(x_3,\varepsilon_3)} \supseteq \cdots$
- $\varepsilon_n \to 0$ as $N \to \infty$.
- $B(x_n, \varepsilon_n) \subseteq U_n$ for all $n \in \mathbb{N}$.

since it will then follow that there exists some u such that

$$u \in \{u\} = \bigcap_{n=0}^{\infty} \overline{B(x_n, \varepsilon_n)} \subseteq B(x, \varepsilon) \cap \left(\bigcap_{n=0}^{\infty} U_n\right)$$

as desired.

We define such a sequence inductively as follows.

First set $x_{-1} = x$, $\varepsilon_{-1} = \varepsilon$. Now for $n \in \mathbb{N} \cup \{-1\}$, observe that by density of U_{n+1} , there must exist $x_{n+1} \in B(x_n, \varepsilon_n) \cap U_{n+1}$.

Since $B(x_n, \varepsilon_n) \cap U_{n+1}$ is open (as the intersection of two open sets), we may find $\varepsilon_{n+1} > 0$ such that $\varepsilon_{n+1} \le 2^{-n-1}$ and $B(x_{n+1}, 2\varepsilon_{n+1}) \subseteq B(x_n, \varepsilon_n) \cap U_{n+1}$. Then certainly $\overline{B(x_{n+1}, \varepsilon_{n+1})} \subseteq B(x_n, \varepsilon_n) \cap U_{n+1}$ as desired, and we can always proceed in this way. Given this definition, the above subset inclusions follow by construction, and since $0 < \varepsilon_n \le 2^{-n}$ for all $n \in \mathbb{N}$, we also have that $\varepsilon_n \to 0$ as $n \to \infty$ as desired.

With this in hand, the proof of Baire category then actually isn't too bad:

Proof (Baire). Suppose towards a contradiction that *X* is meager in itself. Then we may fix $\{U_i\}_{i \in \mathbb{N}}$ nowhere dense such that

$$X = \bigcup_{i \in \mathbb{N}} U_i \implies X = \bigcup_{i \in \mathbb{N}} \overline{U_i} \implies \bigcap_{\alpha \in A} \left(\overline{U_\alpha} \right)^c = \emptyset$$

But by the first lemma, each $(\overline{U_{\alpha}})^c$ is open and dense, so by the second lemma, \emptyset must be dense in *X*, a clear contradiction.

Justifying Ockham's razor again, we see that in the case of miracles, one must fail no matter what. Given the choice then, it's better to fail on a negligible rather than a non-negligible set.

Proposition 3.36. *Fix a* σ *-ideal over W, and suppose that L answers the question almost everywhere, eliminates error almost everywhere, and is inductively monotonic everywhere. Then L is almost everywhere Ockham.*