Inductive Learning and Ockham's Razor

Konstantin Genin
Kevin T. Kelly

Carnegie Mellon University
kgenin@andrew.cmu.edu

Justifying Inductive Methods

Figure : Rudolf Carnap, 1891-1970

Our system of inductive logic ... is intended as a rational reconstruction ... of inductive thinking as customarily applied in everyday life and science. ... An entirely different question is the problem of the validity of our or any other proposed system of inductive logic ... **This is the genuinely philosophical problem of induction** *(Carnap, 1945).*

Justifying Inductive Methods

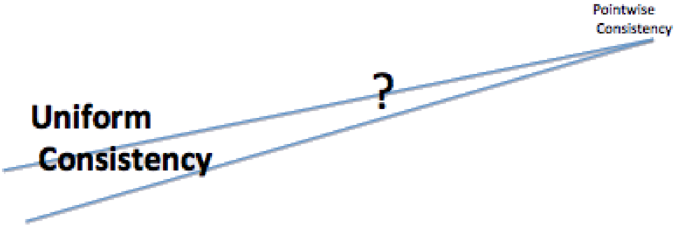
A justification of an inductive procedure

1. must refer to its *success* in some sense;
2. must not require that the truth of its predictions be *guaranteed* in the short-run.

Justifying Inductive Methods

Reichenbach is right ... that any procedure, which does not [converge in the limit] is inferior to his rule of induction. However, his rule ... is far from being the only one possessing that characteristic. The same holds for an infinite number of other rules of induction. ... **Therefore we need a more general and stronger method for examining and comparing any two given rules of induction ...** *(Carnap, 1945)*

Justifying Inductive Methods



Is there something in between?

Ockham's Razor

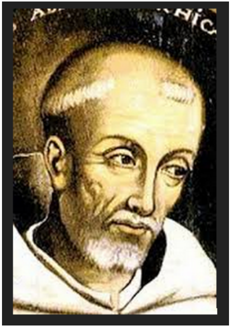


Figure : William of Ockham, 1287-1347

All things being equal, one ought to prefer simpler theories.

Simplicity and Ockham's Razor

Two fundamental questions:

1. How is simplicity to be defined?

Simplicity and Ockham's Razor

Two fundamental questions:

1. How is simplicity to be defined?
2. Given simplicity, what is *Ockham's Razor*?

Simplicity and Ockham's Razor

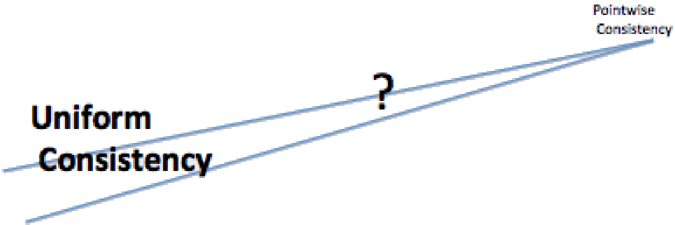
Two fundamental questions:

1. How is simplicity to be defined?
2. Given simplicity, what is *Ockham's Razor*?
3. How does Ockham's Razor help you find the truth?

Simplicity: Epistemic or Methodological?

"Justifying an epistemic principle requires answering an epistemic question: why are parsimonious theories more likely to be true? Justifying a methodological principle requires answering a pragmatic question: why does it make practical sense for theorists to adopt parsimonious theories?" (Baker, SEP).

Justifying Inductive Methods



Ockham's razor cannot provide a short-run guarantee. A justification of Ockham's Razor is tied up with what could be in between these two extremes.

Section 2

Topology as Epistemology

Topological as Epistemology

Related Approaches:

1. Vickers (1996)
2. Kelly (1996)
3. Luo and Schulte (2006)
4. Yamamoto and de Brecht (2010)
5. Baltag, Gierasimczuk, and Smets (2014)

Propositions and Possible Worlds

Let W be a set of possible worlds. A *proposition* is a set $P \subseteq W$. The contradictory proposition is $\varnothing$ and the necessary proposition is W.

$P \wedge Q = P \cap Q$, $P \vee Q = P \cup Q$, $\neg P = W \setminus P$ and P entails Q iff $P \subseteq Q$.

Observable Propositions

Let $\mathcal{O} \subseteq \mathfrak{P}(W)$ be the set of observable propositions. Then the set of all propositions observable in world w is:

$$\mathcal{O}(w) = \{O \in \mathcal{O} : w \in O\}.$$

$\mathcal{O}$ is a *topological basis* iff the following are both satisfied:

O1. $\bigcup \mathcal{O} = W$;

O2. If $A, B \in \mathcal{O}(w)$ then there is $C \in \mathcal{O}(w)$ such that $C \subseteq A \cap B$.

Verifiable Propositions

Say that a proposition P is *verifiable* iff for every world $w \in P$ there is some observation $O \in \mathcal{O}(w)$ such that O entails P.

The following four thesis about verifiability follow from this definition:

V1. The contradictory proposition $\varnothing$ is verifiable.

V2. The trivial proposition W is verifiable.

V3. The verifiable propositions are closed under finite conjunction.

V4. The verifiable propositions are closed under arbitrary disjunction.

The possible worlds and verifiable propositions $(W, \mathcal{V})$ form a *topology*.

Verifiable Propositions

You can verify finitely many sunrises,

   ...

But not that it will rise every morning.

Falsifiable Propositions

Say that a proposition P is *falsifiable* iff $\neg P$ is verifiable.

The following four thesis about falsifiability follow from this definition:

F1. The contradictory proposition $\varnothing$ is falsifiable.

F2. The trivial proposition W is falsifiable.

F3. The falsifiable propositions are closed under finite disjunction.

F4. The verifiable propositions are closed under arbitrary conjunction.

A Translation Key

To translate between topology and epistemology:

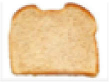
1. basic open set $\equiv$ observable proposition.
2. open set $\equiv$ verifiable proposition.
3. closed set $\equiv$ falsifiable proposition.
4. clopen set $\equiv$ decidable proposition.
5. locally closed set $\equiv$ conditionally refutable proposition.

The Topology of the Problem of Induction

The bread, which I formerly ate, nourished me ... but does it follow, that other bread must also nourish me at another time, and that like sensible qualities must always be attended with like secret powers? The consequence seems nowise necessary (Enquiry Concerning Human Understanding).

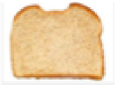
Sierpinski Space

Suppose we have two worlds.



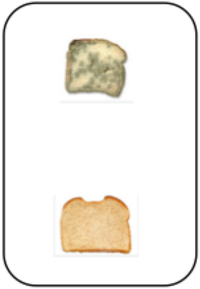
Sierpinski Space

Suppose we have two worlds.



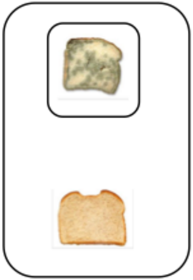
Sierpinski Space

If bread always nourishes, we can never rule out that one day it will stop nourishing.



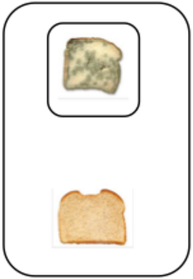
Sierpinski Space

If someday bread will cease to nourish, this will be verified.



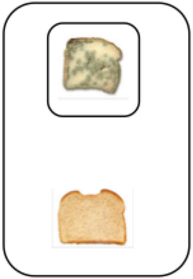
Sierpinski Space

This simple structure defines the *Sierpinski space*, a simple topological space.



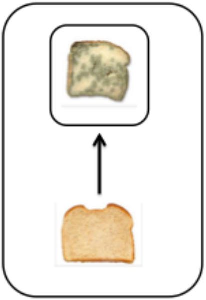
Sierpinski Space

Note that all information compatible with the bottom world is compatible with the top world, but the converse is not true.



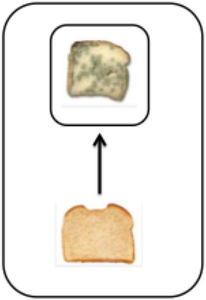
The specialization order

Let $w \preceq v$ iff $\mathcal{O}(w) \subseteq \mathcal{O}(v)$ i.e. all information consistent with w is consistent with v.



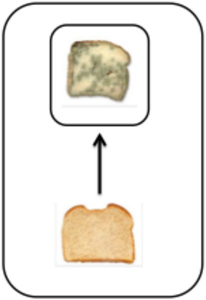
The specialization order

Let $w \prec v$ if $w \preceq v$ but $v \npreceq w$.



The specialization order

This defines the *specialization order* over points in the space.



The specialization order

A topology is T_0 if for all w, v, if $w \neq v$ then $\mathcal{O}(w) \neq \mathcal{O}(v)$ i.e. if two worlds are distinct, then there is some observational difference between them. The T_0 axiom rules out "metaphysical" distinctions between worlds.

For T_0 spaces, the specialization order is a *partial* order.

The specialization order

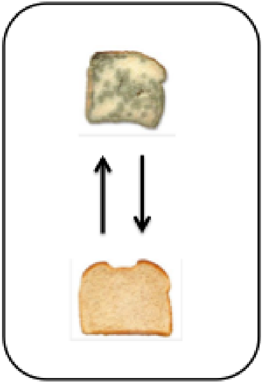


Figure : A "metaphysical" distinction.

The specialization order

A topology is T_d (Aull and Thron, 1962) iff for all w, if $\{v : v \prec w\}$ is closed.

Section 3

Empirical Simplicity

Topological Closure

The *closure* of a proposition A is the set of all worlds where A is never refuted:

$$\overline{A} = \{w : \text{Every } O \in \mathcal{O}(w) \text{ is consistent with } A\}.$$

Furthermore,

$$\overline{\{w\}} = \{v : v \leq w\}.$$

Popper and Simplicity

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 (Popper, 1959).

Popper and Simplicity

A proposition P is *more falsifiable* than Q if and only if every observation that falsifies Q falsifies P.

Equivalently, every observation consistent with P is consistent with Q.

Popper and Simplicity

A proposition P is *more falsifiable* than Q if and only if every observation that falsifies Q falsifies P.

Equivalently, every observation consistent with P is consistent with Q.

So if P is true, Q will never be refuted. Therefore $P \subseteq \overline{Q}$.

Empirical Simplicity

Definition (The Simplicity Order)

P is simpler than Q, written $P \preceq Q$,

1. iff $P \subseteq \overline{Q}$,
2. iff P entails Q will never be refuted,
3. iff P has a problem of induction with Q,
4. iff P is more falsifiable than Q.

Section 4

Learning

Empirical Problem Contexts

An *empirical problem context* is a triple $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$.

- W is the set of possible worlds.
- $\mathcal{O}$ is a countable set of observables.
- $\mathcal{Q}$ is a *question* that partitions W into countably many *answers*.

Let $\mathcal{Q}(w)$ be the answer true at w.

Learning

An *empirical method* is a function $\lambda : \mathcal{O} \to \mathcal{P}(W)$.

Say that λ is *solves* $\mathfrak{P}$ *in the limit* iff for all $w \in W$, there is $E \in \mathcal{O}(w)$ such that for all $F \in \mathcal{O}(w)$, $w \in \lambda(E \cap F) \subseteq \mathcal{Q}(w)$.

Say that $\mathfrak{P}$ is *solvable in the limit* iff there exists λ that solves it in the limit.

Learning

Proposition (Yamamoto and de Brecht (2010))

If $|W| \leq \omega$ *then* $\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is solvable in the limit iff* $(W, \mathcal{O})$ *is* T_d.

Proposition (Baltag, Gierasimczuk, and Smets (2014))

$\mathfrak{P} = (W, \mathcal{O}, \mathcal{Q})$ *is solvable in the limit iff each* $Q \in \mathcal{Q}$ *is a countable union of locally closed sets.*

A Special Case

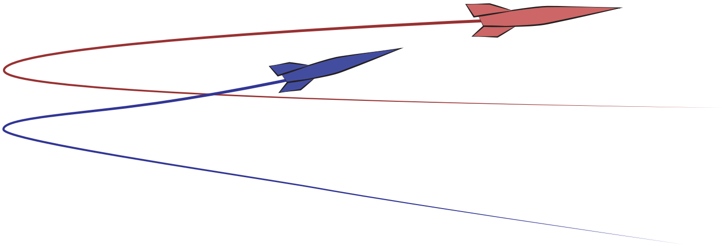
For simplicity we restrict our attention to the case where $|W| \leq \omega$ and $\mathcal{Q} = \mathcal{Q}_{\perp}$.

Section 5

Efficient Convergence

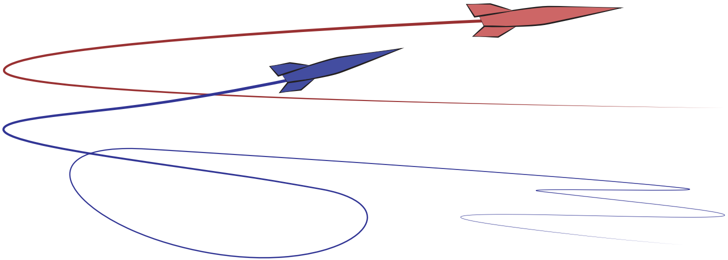
Refining Convergence

Pursuit of truth ought to be as direct as possible.



Refining Convergence

Needless cycles and reversals in opinion ought to be avoided.



Cycles and Reversals

Definition (Reversals)

$(\lambda(E), \lambda(F))$ is a *reversal* iff $F \subset E$ and $\lambda(F) \subseteq \lambda(E)^c$.

Definition (Cycles)

$(\lambda(E), \lambda(F), \lambda(G))$ is a *cycle* iff $(\lambda(E), \lambda(F))$ and $(\lambda(F), \lambda(G))$ are reversals and $\lambda(G) \subseteq \lambda(E)$.

Monotonicity Principles

A learner λ is

1. **rationally monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when F meets $\lambda(E)$.

Monotonicity Principles

A learner λ is

1. **rationally monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when F meets $\lambda(E)$.
2. **cautiously monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when $\lambda(E) \subseteq F$.

Monotonicity Principles

A learner λ is

1. **rationally monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when F meets $\lambda(E)$.
2. **cautiously monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when $\lambda(E) \subseteq F$.
3. **reversal monotone** if $\lambda(E \cap F)$ meets $\lambda(E) \cap F$ when F meets $\lambda(E)$.

Monotonicity Principles

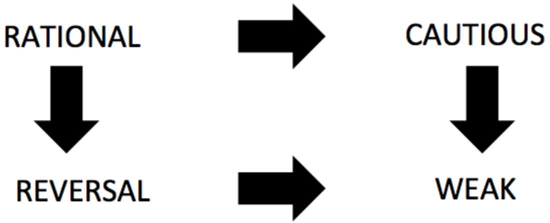
A learner λ is

1. **rationally monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when F meets $\lambda(E)$.
2. **cautiously monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when $\lambda(E) \subseteq F$.
3. **reversal monotone** if $\lambda(E \cap F)$ meets $\lambda(E) \cap F$ when F meets $\lambda(E)$.
4. **weakly monotone** if $\lambda(E \cap F)$ meets $\lambda(E) \cap F$ when $\lambda(E) \subseteq F$.

Monotonicity Principles

A learner λ is

1. **rationally monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when F meets $\lambda(E)$.
2. **cautiously monotone** if $\lambda(E \cap F) \subseteq \lambda(E) \cap F$ when $\lambda(E) \subseteq F$.
3. **reversal monotone** if $\lambda(E \cap F)$ meets $\lambda(E) \cap F$ when F meets $\lambda(E)$.
4. **weakly monotone** if $\lambda(E \cap F)$ meets $\lambda(E) \cap F$ when $\lambda(E) \subseteq F$.



Monotonicity Principles

Proposition

A learner λ is reversal monotone iff there are no $E, F, G \in \mathcal{O}$ such that $(\lambda(E), \lambda(F), \lambda(G))$ is a cycle.

Section 6

Ockham's Razor

Ockham's Razor

Definition (The Vertical Razor)

Say that a learner λ is vertical Ockham iff

1. iff for all $w \in E$, $\{w\} \leq \lambda(E) \Rightarrow w \in \lambda(E)$;
2. iff $\lambda(E)$ is closed in E;
3. iff $\lambda(E)$ is downward-closed in $\leq$.

Vertical Ockham and Monotonicity

Proposition

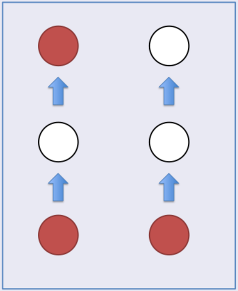
A learner λ is reversal monotone **only if** *λ is vertical Ockham.*

Vertical Ockham and Monotonicity

Proposition

A learner λ is reversal monotone **only if** *λ is vertical Ockham.*

Sketch.



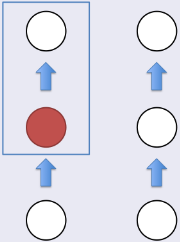
Suppose you violate the vertical razor. □

Vertical Ockham and Monotonicity

Proposition

A learner λ is reversal monotone **only if** *λ is vertical Ockham.*

Sketch.



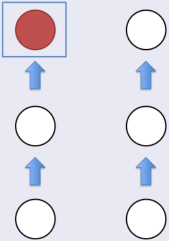
You reverse on further information, though your first conjecture is not refuted. □

Vertical Ockham and Monotonicity

Proposition

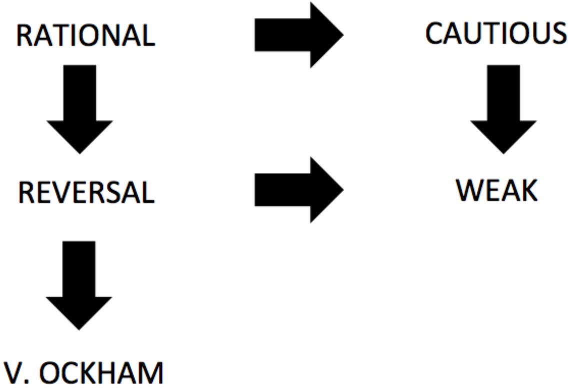
A learner λ is reversal monotone **only if** *λ is vertical Ockham.*

Sketch.



On even further information, you are forced into a cycle. $\square$

Vertical Ockham and Monotonicity



Definition (The Weak Razor)

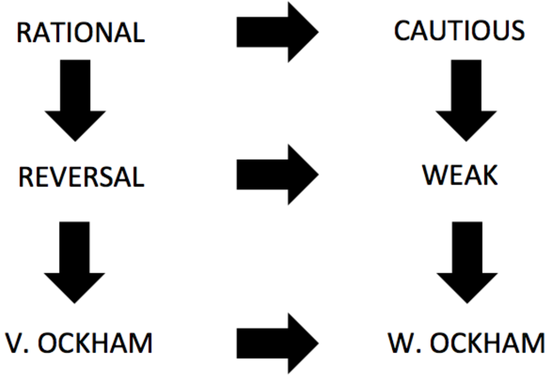
Say that a learner λ is weak Ockham iff for all E,

$$\left(\bigcap_{v \in \lambda(E)} \overline{\{v\}} \right) \cap E \subseteq \lambda(E).$$

Proposition

A learner λ is weakly monotone **only if** *λ is weak Ockham.*

Ockham and Monotonicity



Definition (The Horizontal Razor)

Say that a learner λ is horizontal Ockham iff for all E, $\lambda(E)$ is co-initial in $\leq$.

Reversal Optimality

Definition

Say that λ is *reversal-optimal iff for every reversal sequence* $(\lambda(E), \lambda(F))$ *and learner* λ' *there is a reversal sequence* $(\lambda'(G), \lambda'(H))$ *such that* $\lambda'(G) \subseteq \lambda(E)$ *and* $\lambda'(F) \subseteq \lambda(H)$.

Proposition

A learner λ *is reversal-optimal iff* λ *is horizontal Ockham.*

Retractions

Definition (Retractions)

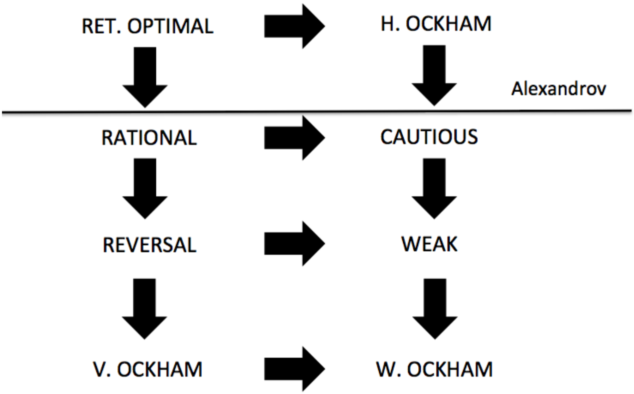
$(\lambda(E), \lambda(F))$ is a *retraction* iff $F \subset E$ and $\lambda(F) \not\subseteq \lambda(E)$.

Retraction Optimality

Definition

Say that λ is *retraction-optimal iff for every retraction sequence* $(\lambda(E), \lambda(F))$ *and learner* λ' *there is a retraction sequence* $(\lambda'(G), \lambda'(H))$ *such that* $\lambda'(G) \subseteq \lambda(E)$ *and* $\lambda'(F) \subseteq \lambda(H)$.

Ockham and Monotonicity



Thank you!

Supported by a grant from the John Templeton Foundation.

Manuscript: http://www.andrew.cmu.edu/user/kk3n/simplicity/bulletin-12.pdf

Works Cited

CE Aull and WJ Thron. Separation axioms between t0 and t1. *Indag. Math*, 24:26–37, 1962.

Alexandru Baltag, Nina Gierasimczuk, and Sonja Smets. Epistemic topology (to appear). *ILLC Technical Report*, 2014.

Rudolf Carnap. On inductive logic. *Philosophy of Science*, 12(2):72, 1945.

Kevin Kelly. *The Logic of Reliable Inquiry*. 1996.

Wei Luo and Oliver Schulte. Mind change efficient learning. *Information and Computation*, 204(6):989–1011, 2006.

Karl R. Popper. *The Logic of Scientific Discovery*. London: Hutchinson, 1959.

Steven Vickers. *Topology Via Logic*. Cambridge University Press, 1996.

Akihiro Yamamoto and Matthew de Brecht. Topological properties of concept spaces (full version). *Information and Computation*, 208(4):327–340, 2010.