Simplicity and Scientific Questions

{Kevin T. Kelly, Konstantin Genin}

Ockham’s Razor

- **Ockham:** “Pluralitas non est ponenda sine neccesitate.”
- **Science:** “Presume no more complexity than necessary.”

Ockham's Razor

- **Ockham:** "Pluralitas non est ponenda sine neccesitate."
- **Science:** "Presume no more complexity than necessary."
- But what is simplicity?
- And why rely on it?

Indispensable for:

- Theory choice
- Inductive inference, e.g. language learning.
- Statistical model selection
- Causal discovery from non-experimental statistical data.

Epistemic Justification

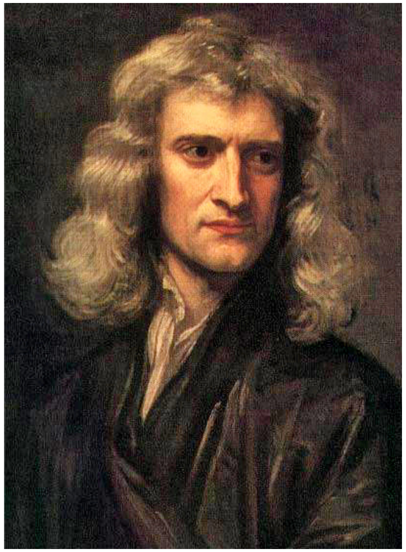
Traditionally, epistemic justification has been too strictly conceived:

> . . . justifying an epistemic principle requires answering an epistemic question: why are parsimonious theories more likely to be true? (Baker, 2013)

Epistemic Justification

When your standards are too high, you are led either to metaphysics,

Nature is pleased with simplicity and affects not the pomp of superfluous causes (Newton, 1833).



Epistemic Justification

… or despair.

> [N]o one has shown that any of these rules is more likely to pick out true theories than false ones. It follows that none of these rules is epistemic in character (Laudan, 2004).

Epistemic Justification

Theoretical virtues do not indicate the truth the way litmus paper indicates pH.

Epistemic Justification

Theoretical virtues do not indicate the truth the way litmus paper indicates pH. We can make progress if we don't demand the impossible:

> The fact that the truth of the predictions reached by induction cannot be guaranteed does not preclude a justification in a weaker sense (Carnap, 1945).

Epistemic Justification

Truth-indicativeness is too strong a standard. But mere convergence to the truth in the limit is too weak to mandate any behavior in the short run.

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)

Epistemic Justification

Truth-
Indicative

?

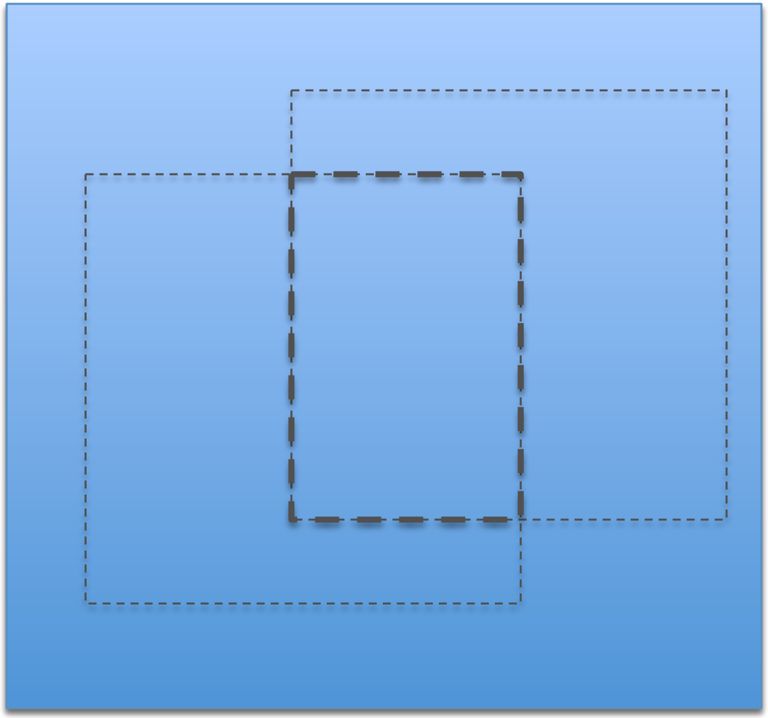
Converges
In the limit

Is there something in between?

1. EMPIRICAL PROBLEMS

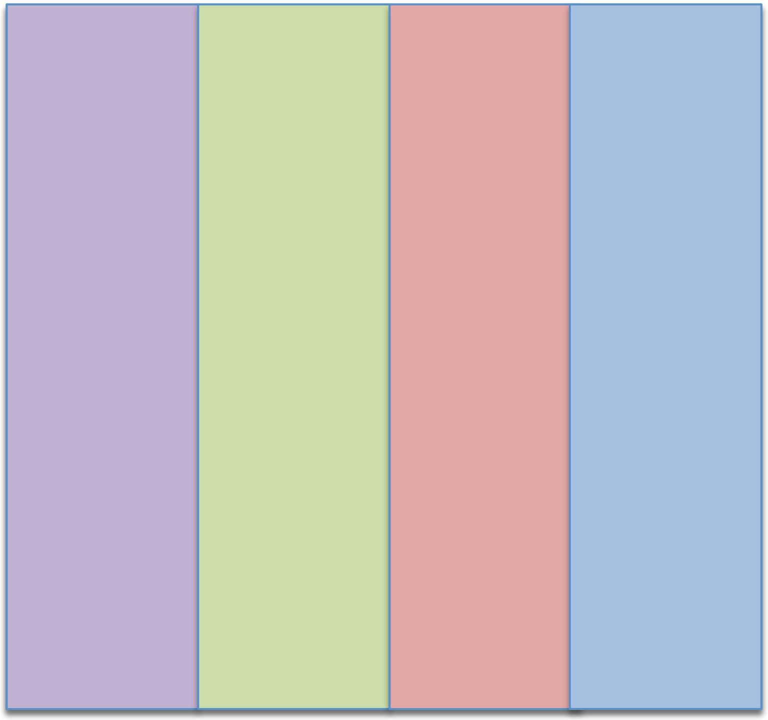
Information Spaces

- W is a set of possible worlds.
- $\mathcal{I}$ is a set of propositional information states.
  1. Covers W.
  2. Closed under finite conjunction.



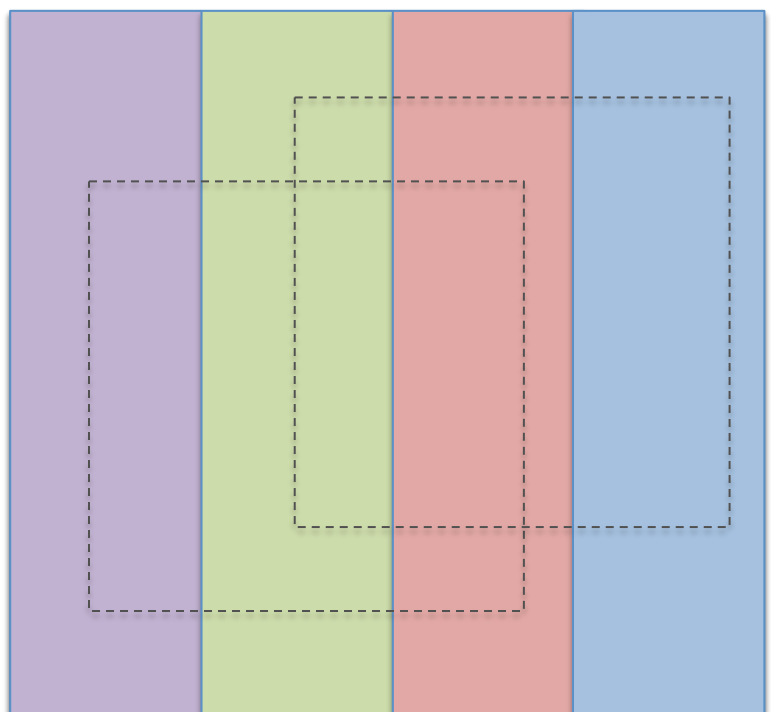
Questions

- A question partitions W into countably many possible answers (Hamblin 1958)
- Relevant responses are disjunctions of answers.

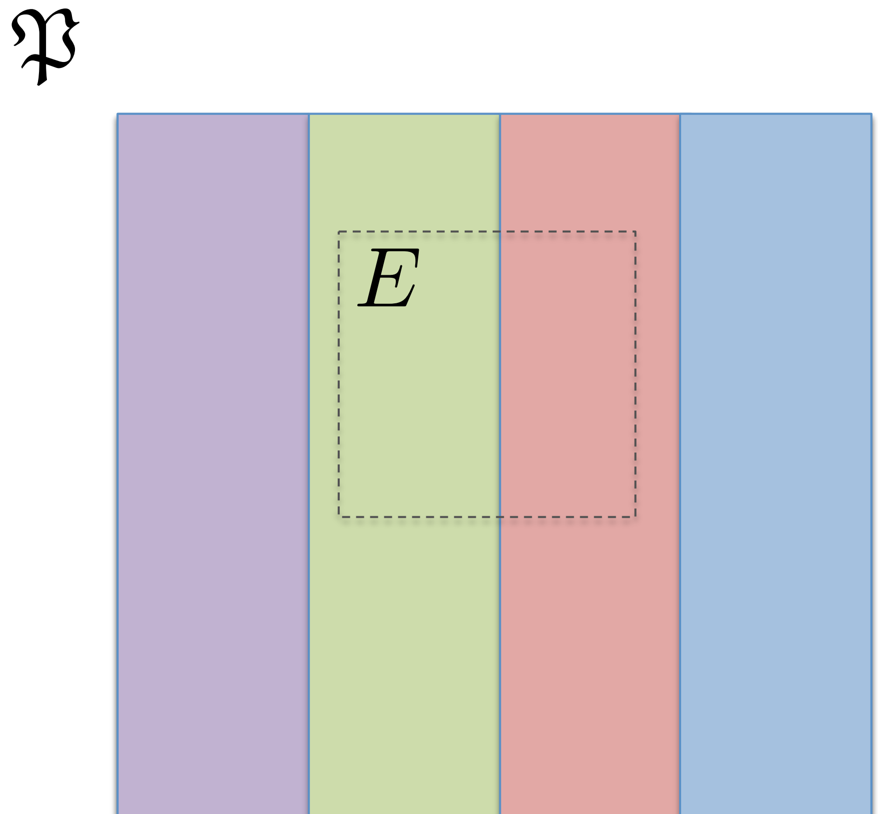


Empirical Problem

$$\mathfrak{P} \; = \; (W, \mathcal{I}, \mathcal{Q}).$$

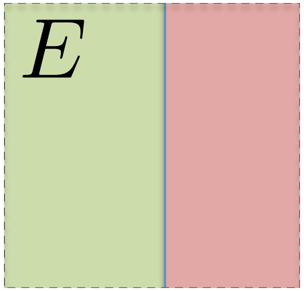


Problem Restriction



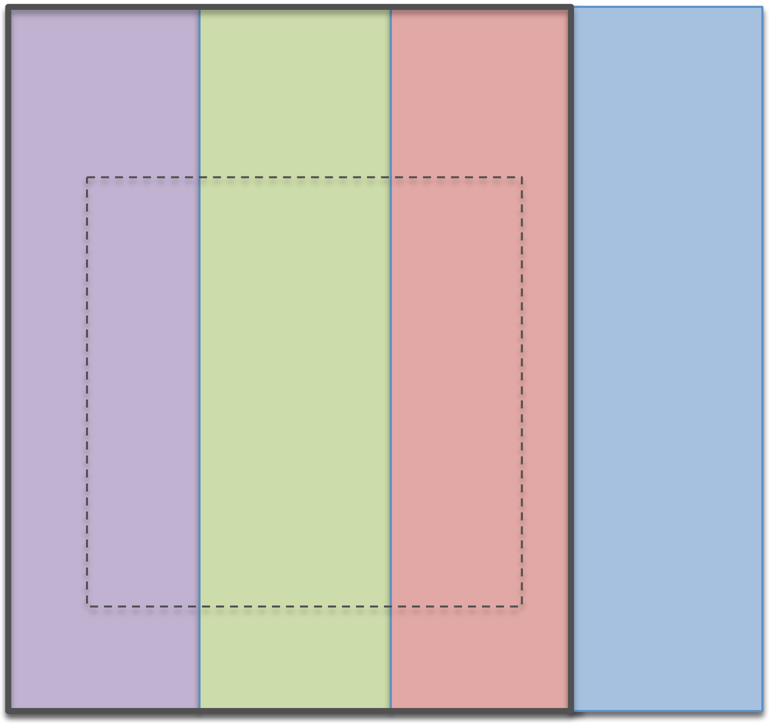
Problem Restriction

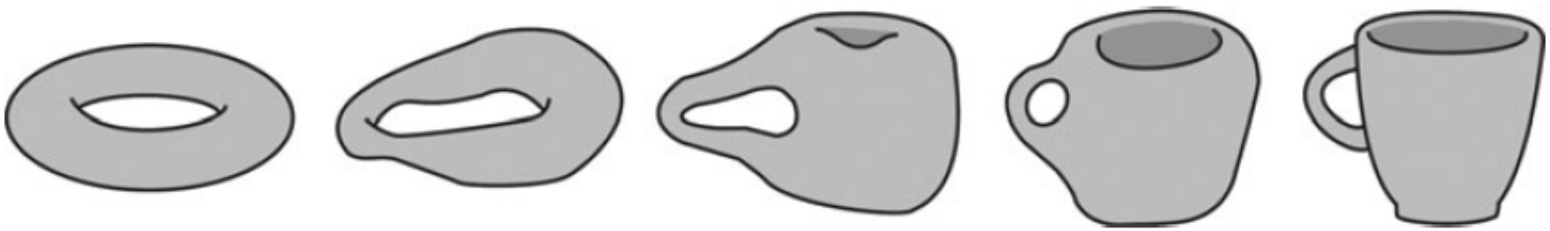
$\mathfrak{P}|_E$



Relevant Response Given E

Disjunction of answers compatible with E.

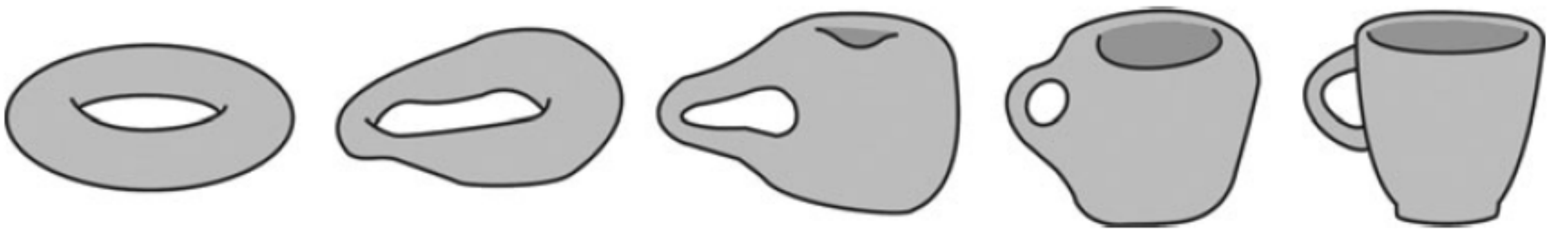




2. INFORMATION TOPOLOGY

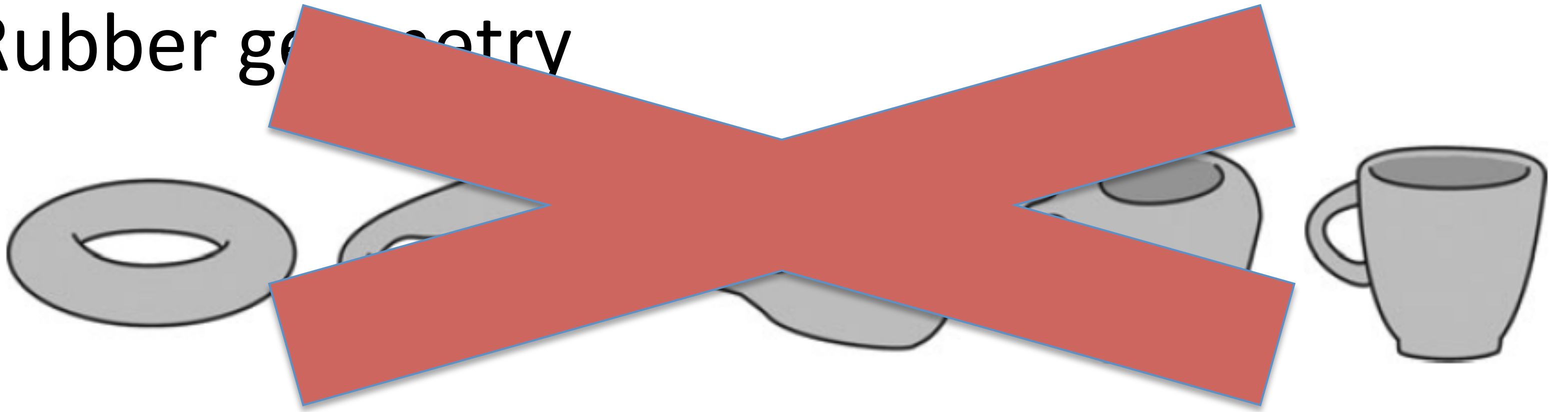
Topology

Rubber geometry



Topology

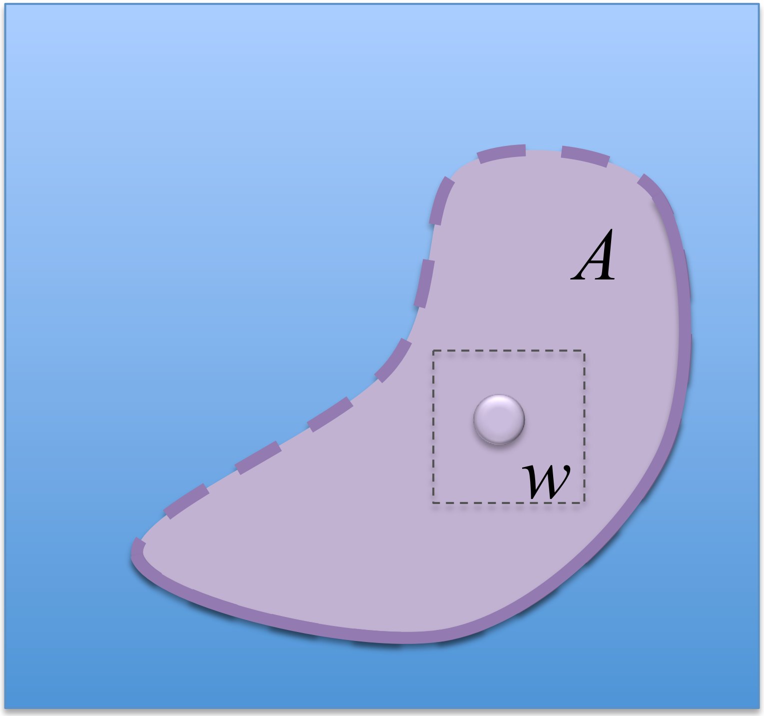
Rubber geometry



The logic of verification.

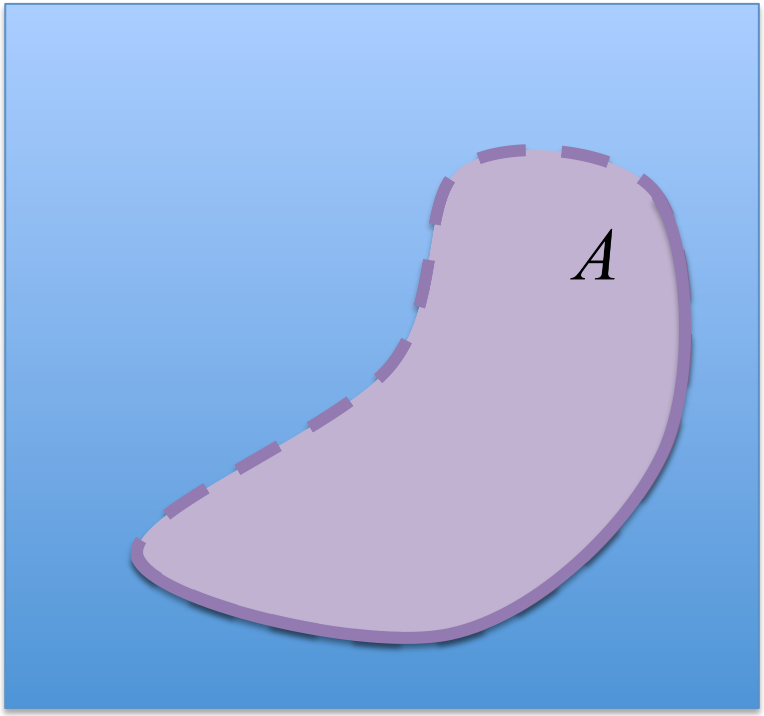
w is an Interior Point of A

W presents information that verifies A.



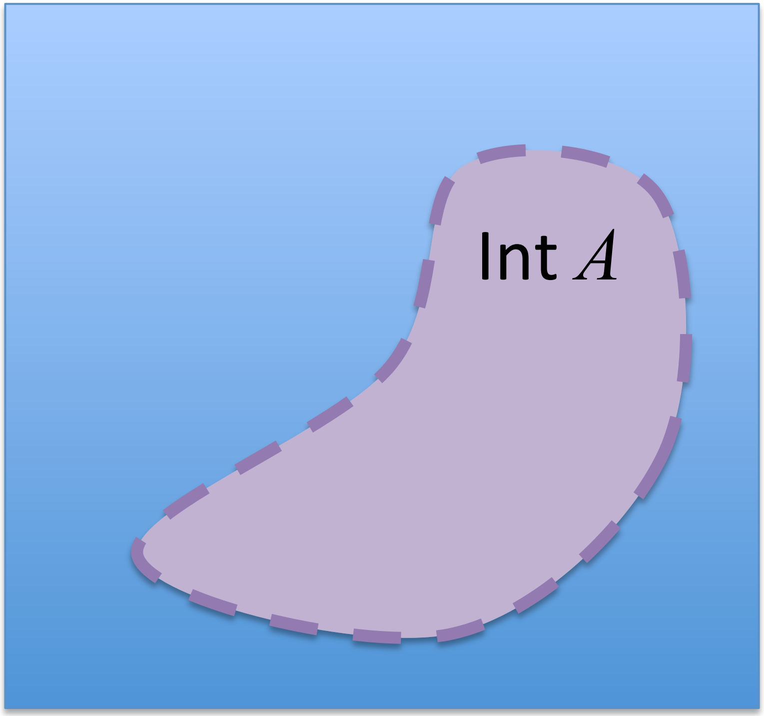
Interior of A

Int A = it will be verified that A.



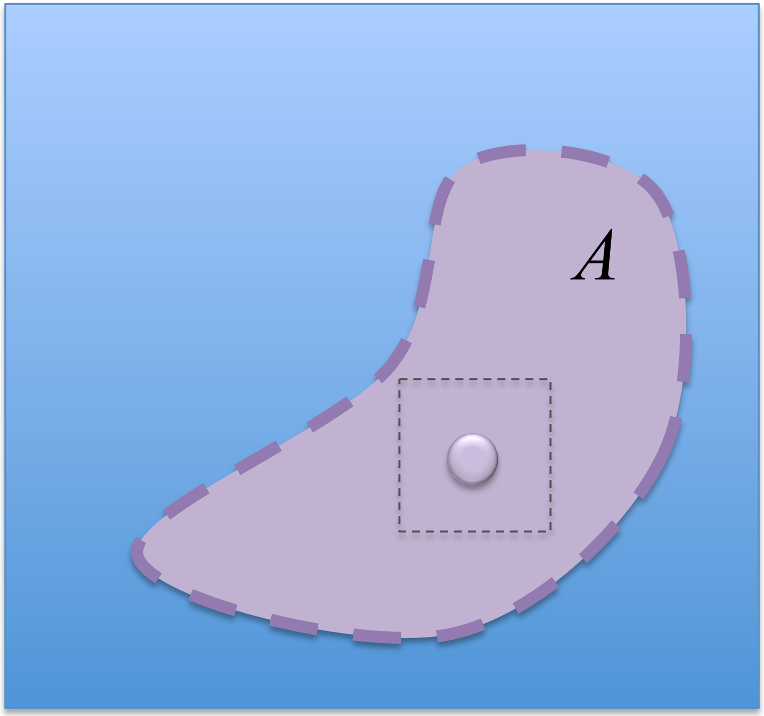
Interior of A

Int A = it will be verified that A.



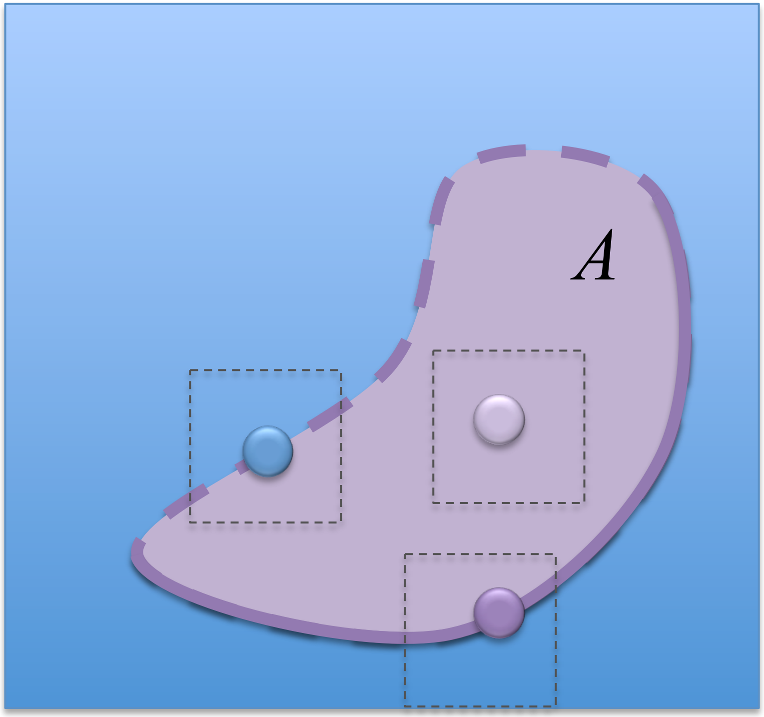
Open = Verifiable

A is open iff A entails that A will be verified.



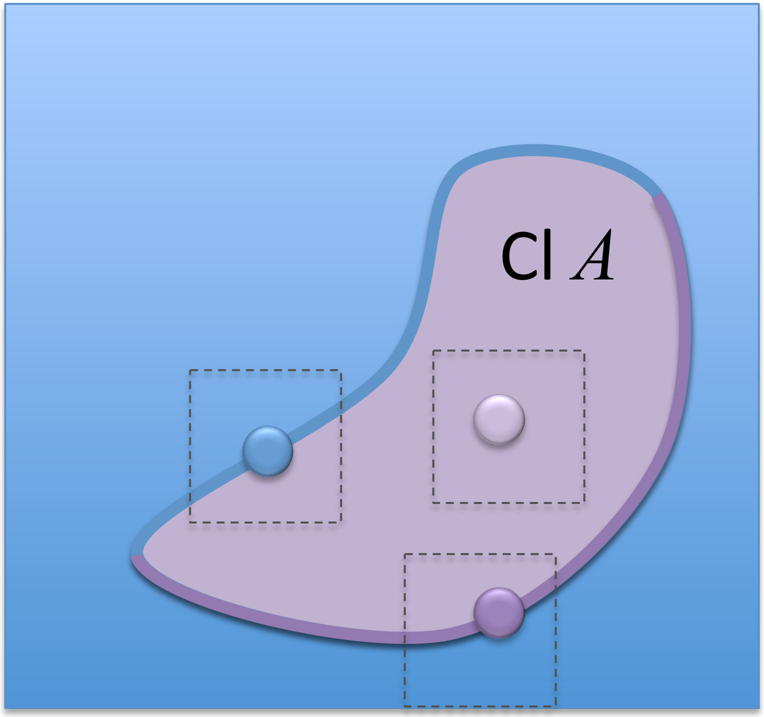
Closure of A

Cl A = A will never be refuted.



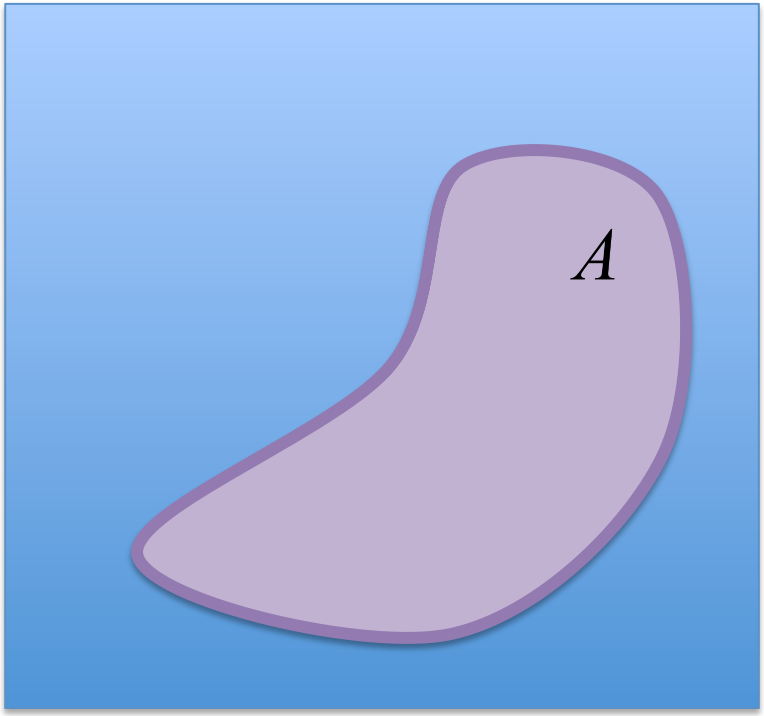
Closure of A

Cl A = A will never be refuted.



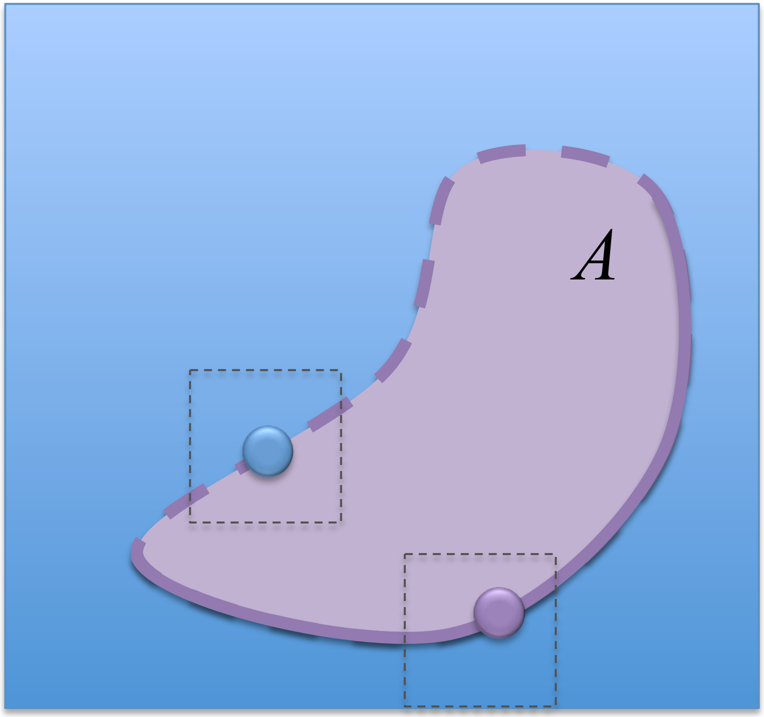
Closed = Refutable

A is closed iff not-A entails that A will be refuted.



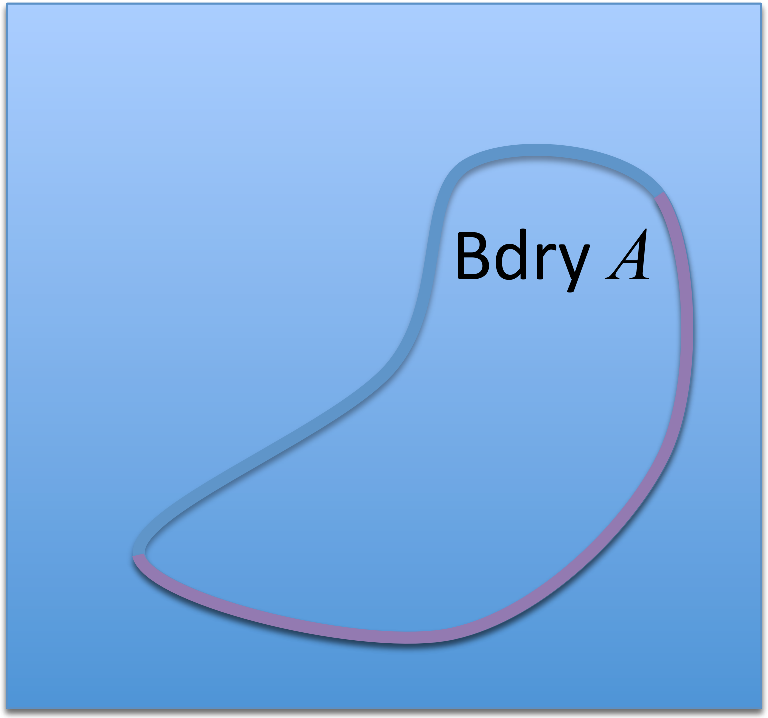
Boundary of A

$\text{Bdry}A$ = "A will never be decided".



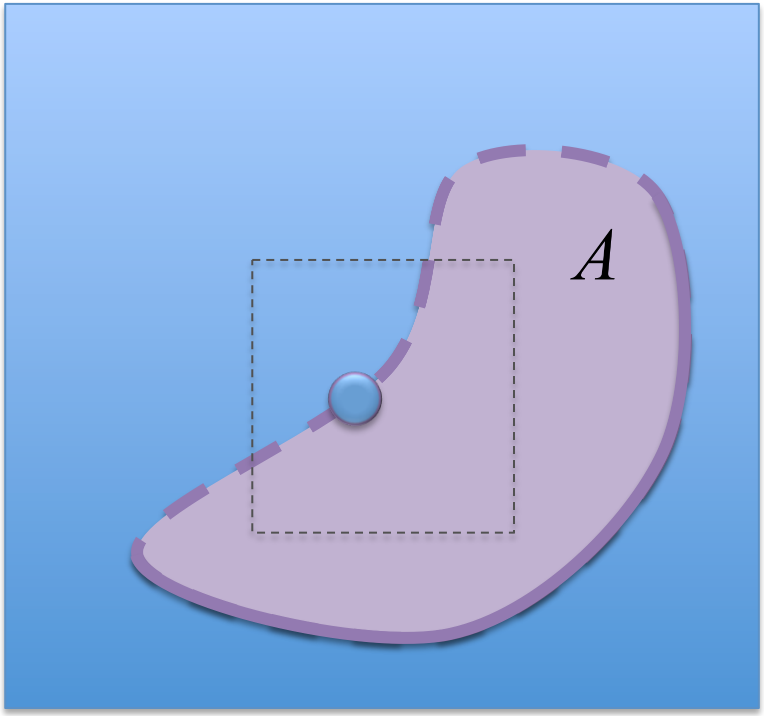
Boundary of A

$$\text{Bdry}\,A \;=\; \text{“}A \text{ will never be decided”}.$$



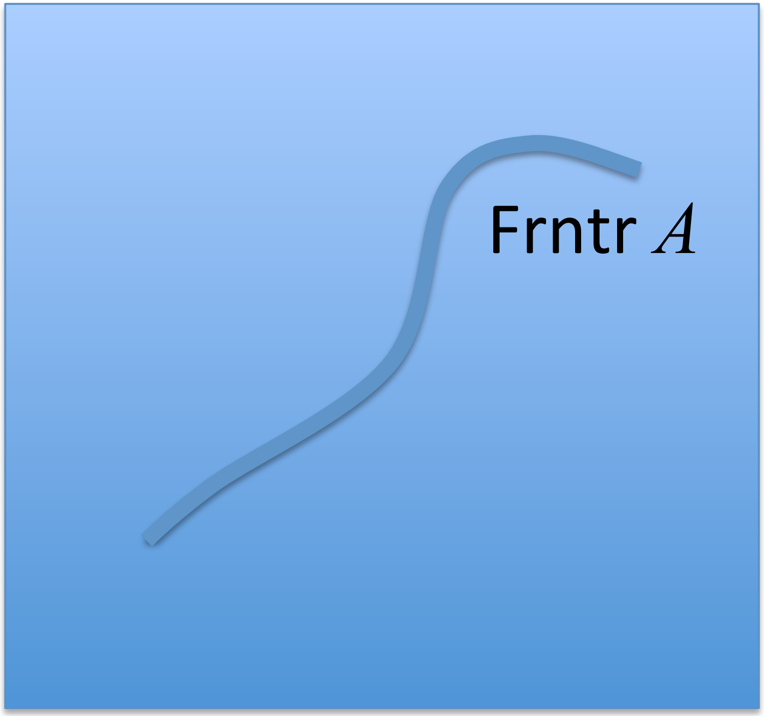
Frontier of A

Frntr A = A is false, but will never be refuted.



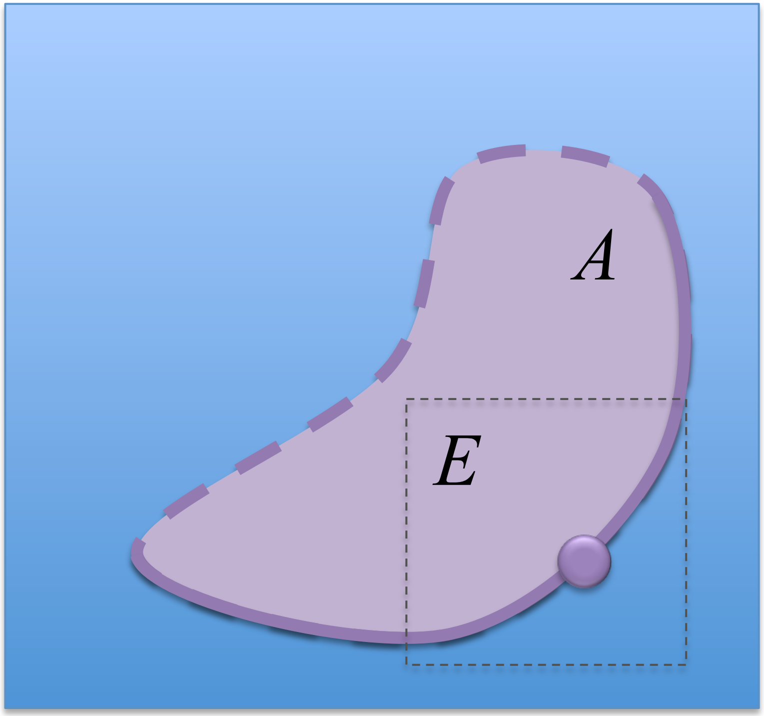
Frontier of A

Frntr A = A is false, but will never be refuted.



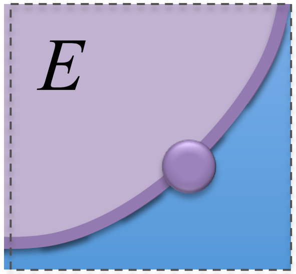
Locally Closed

A is locally closed iff A entails that A will become refutable.



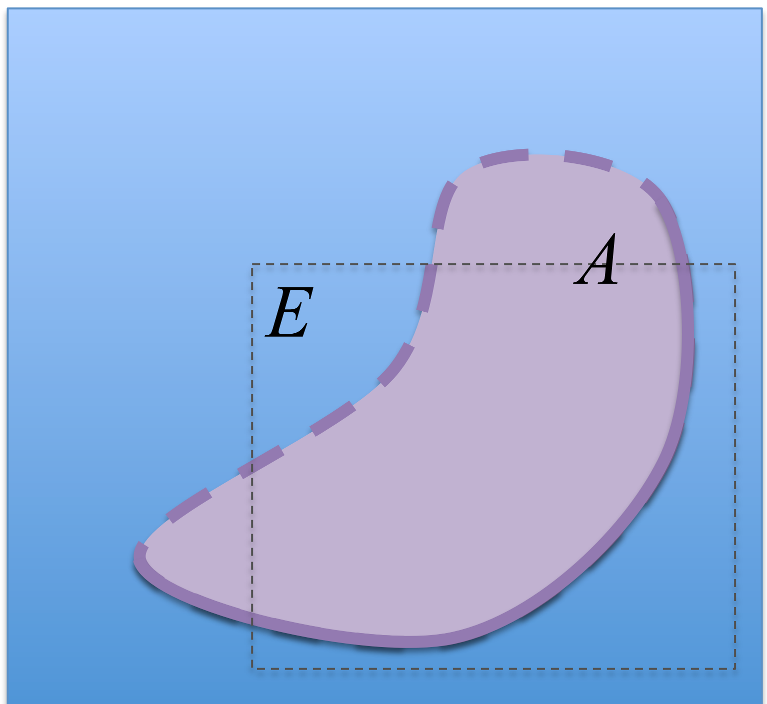
Locally Closed

A is locally closed iff A entails that A will become refutable.



Local Closure and Default Reasoning

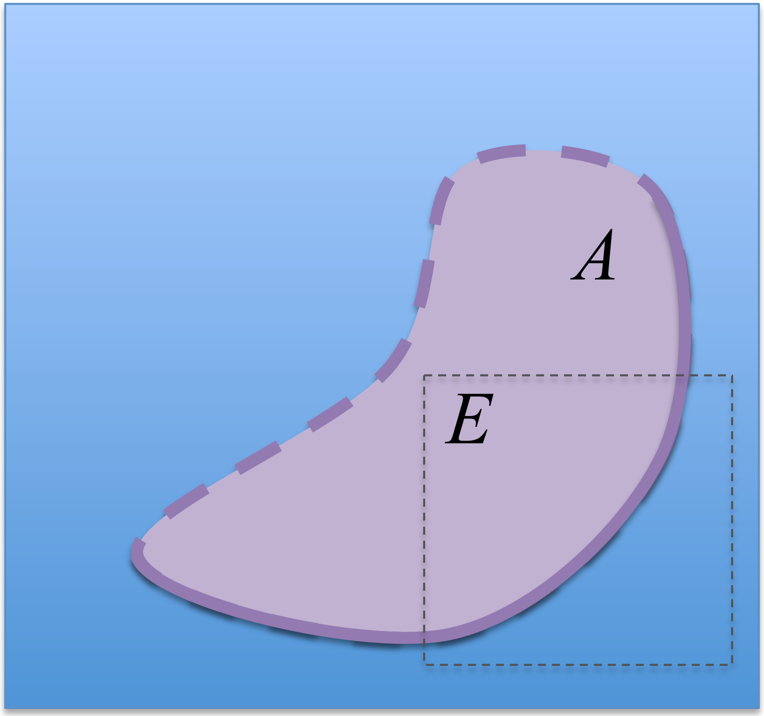
E is a default reason for A iff
A is refutable but not refuted given E.



E is not a default reason for A.

Locally Closed

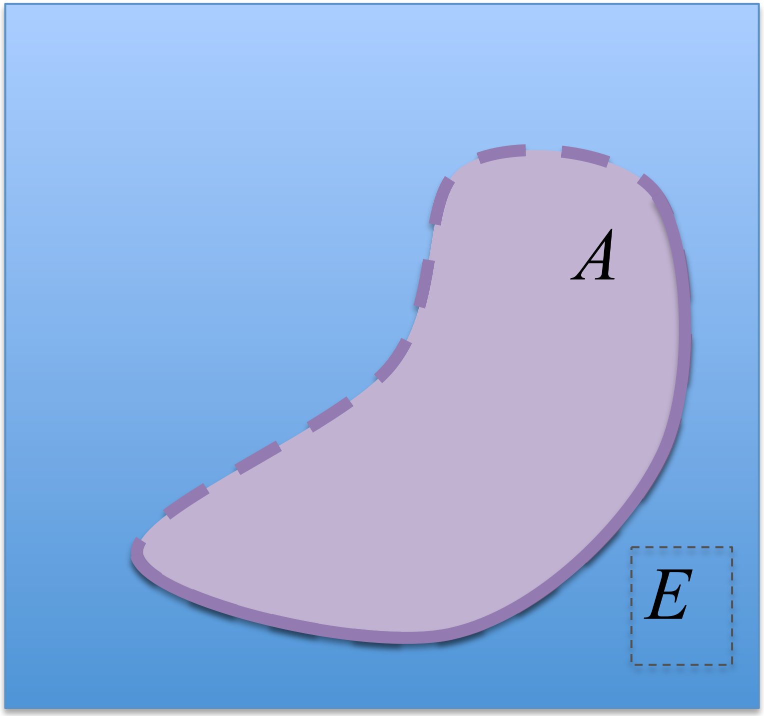
E is a default reason for A iff
A is refutable but not refuted given E.



E is a default reason for A.

Locally Closed

E is a default reason for A iff
A is refutable but not refuted given E.



E refutes A.

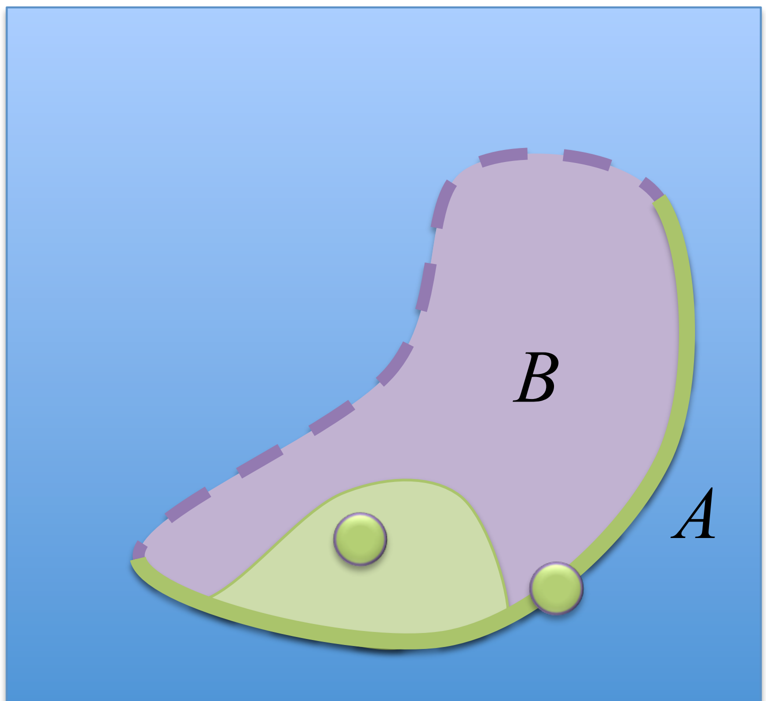
3. EMPIRICAL SIMPLICITY

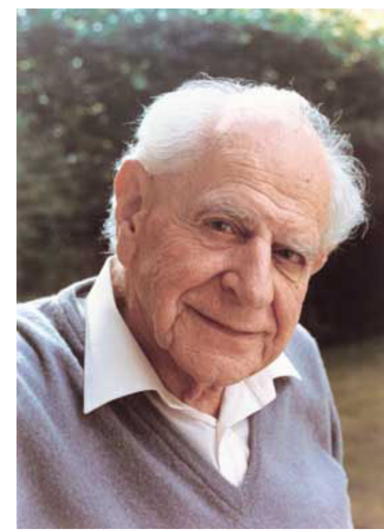
Simplicity = Specialization Preorder

$A \preceq B$ iff $A \subseteq \mathrm{cl} B$

iff all information compatible with A is compatible with B

iff all information refuting B also refutes A.

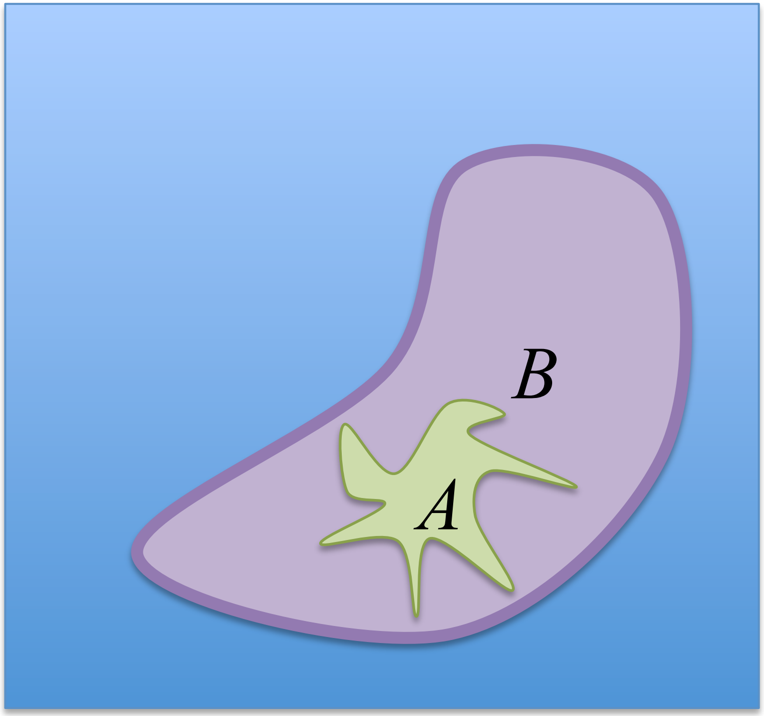


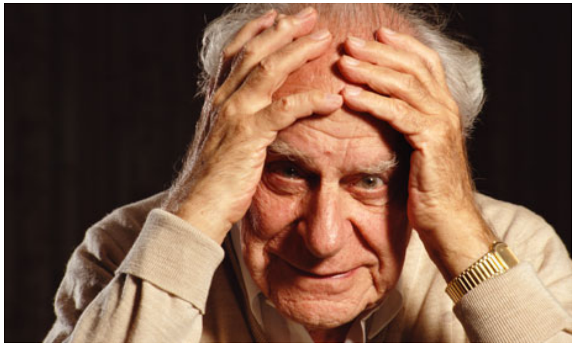


Sir Karl Popper

The "Tack-on" Objection

- Adding complex principles to a simple theory doesn't make it simpler (Glymour 1980).

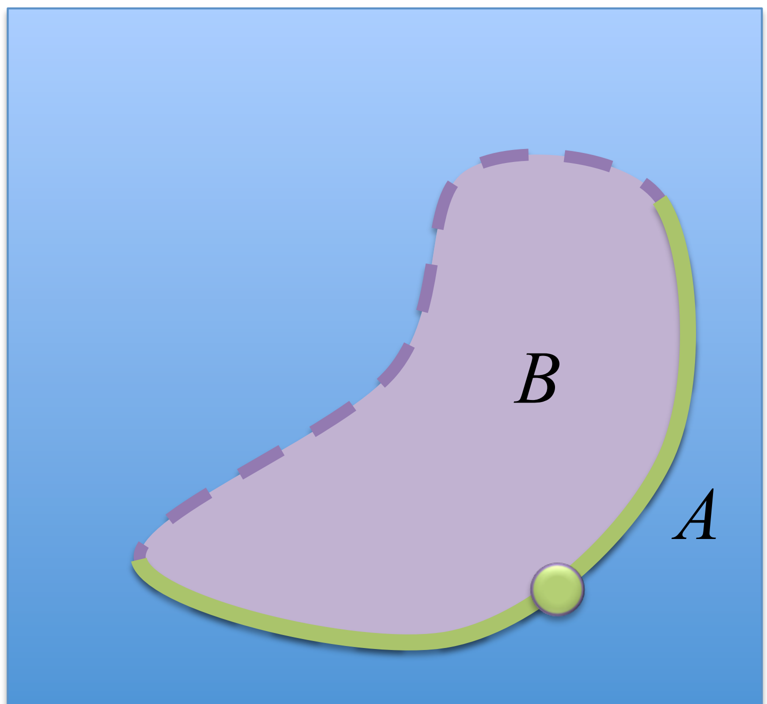




Empirical Simplicity

$A \lhd B$ iff $A \cap \text{Frntr} B \neq \varnothing.$

iff A is consistent with: B is false, but will never be refuted.

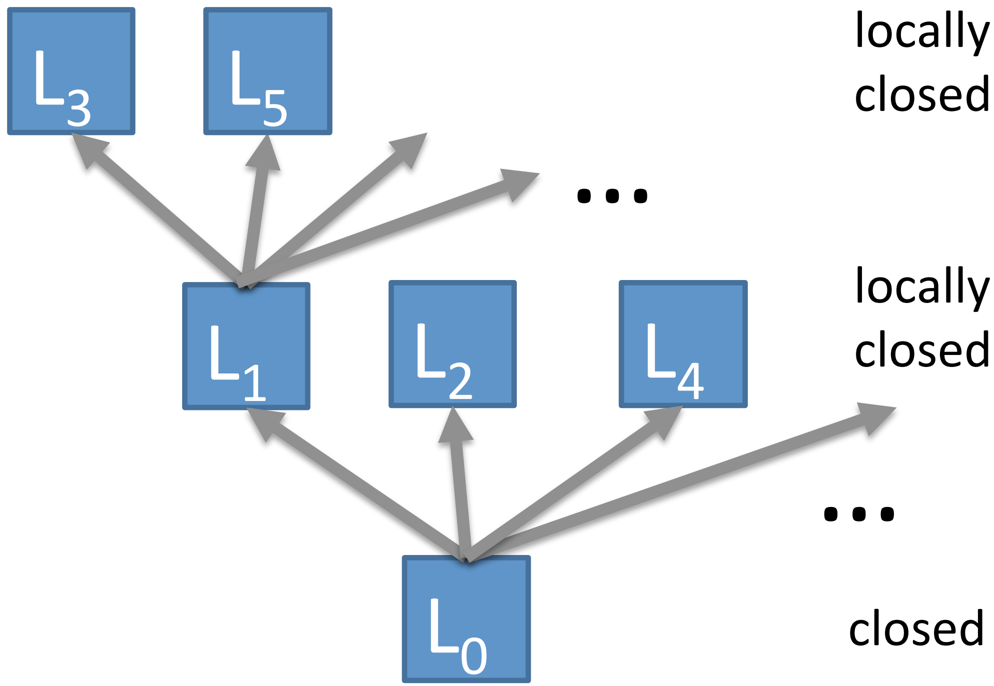


Example: Language Learning (Gold)

$\mathcal{Q}$ = Which language?

$\mathcal{I}$ = finitely many grammatical sentences.

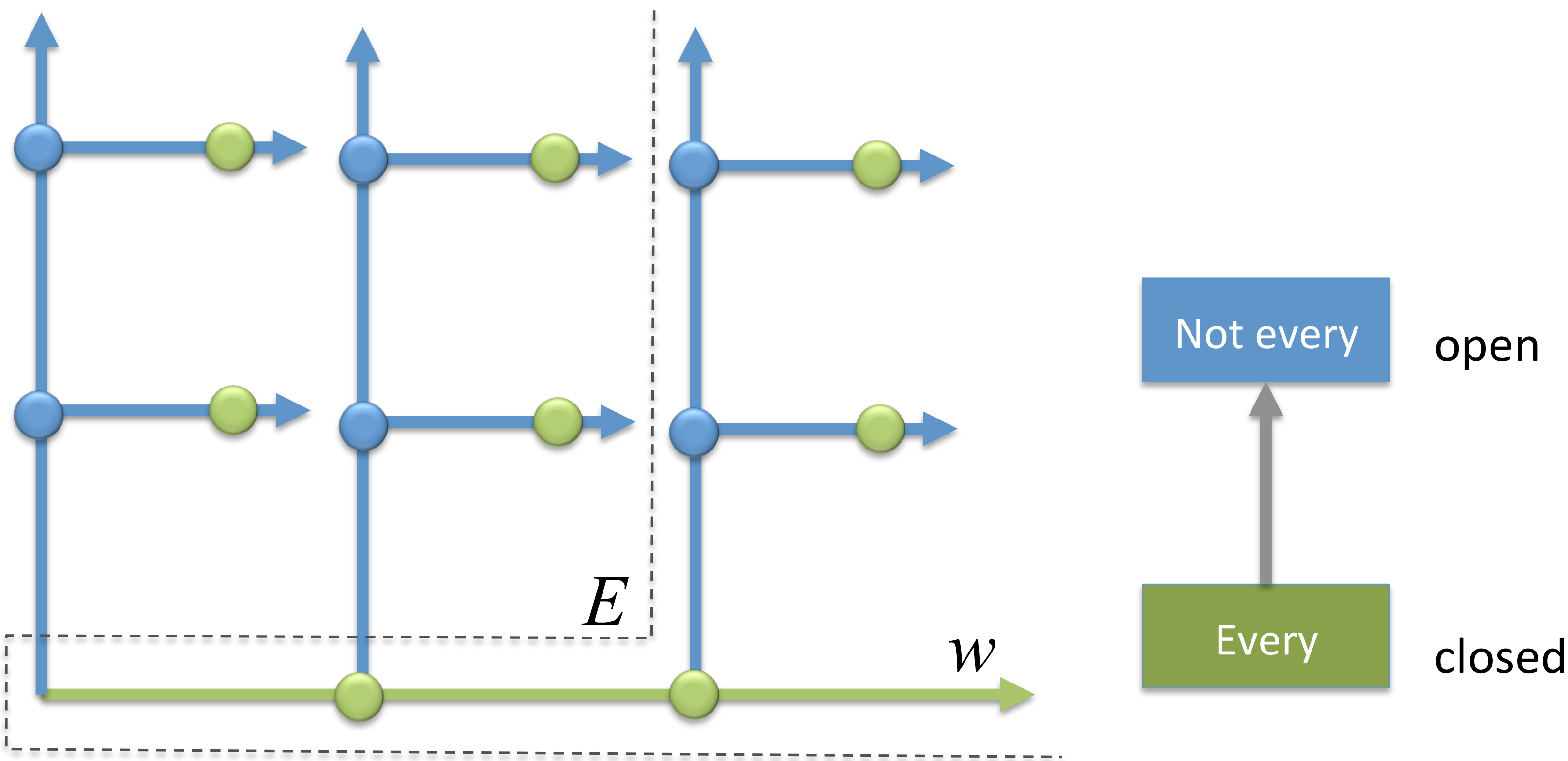
$s_1, s_2, s_3, \ldots$



Example: The Baire Space

$\mathcal{Q}$ = Will every outcome be green?

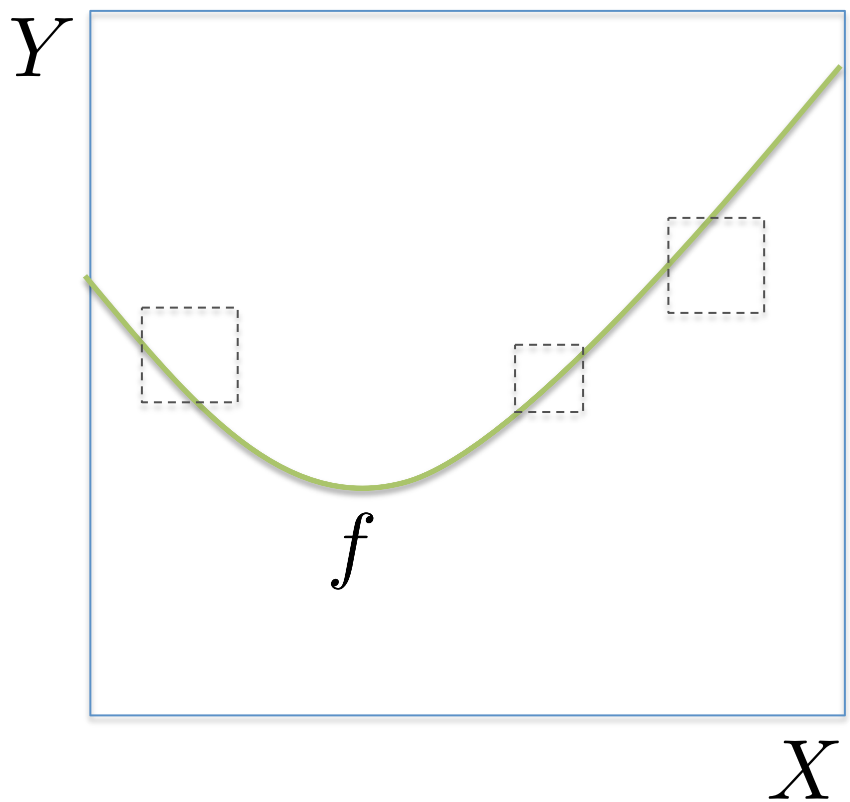
$\mathcal{I}$ = observation histories.

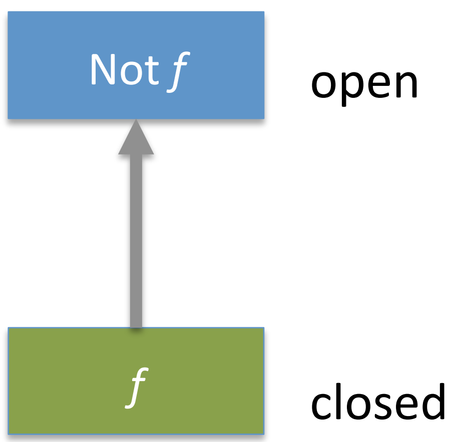


Example: Continuous Laws

$\mathcal{Q} =$ Does $Y = f(X)$?.

$\mathcal{I} =$ finitely many inexact measurements.

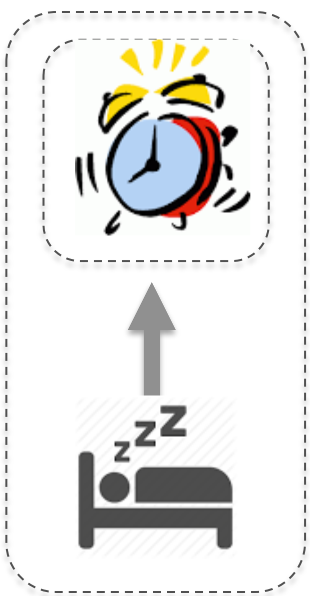


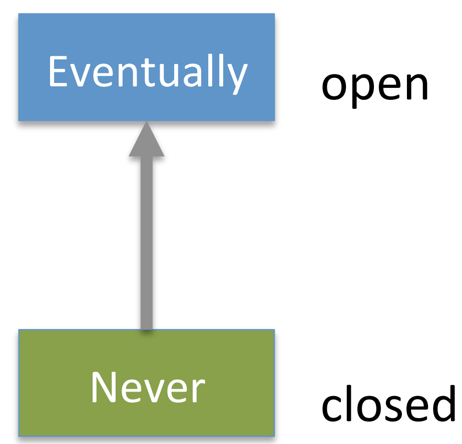


Example: Sierpinski Space

$\mathcal{Q} =$ Will it ever ring?

$\mathcal{I} =$ alarm or no alarm yet.

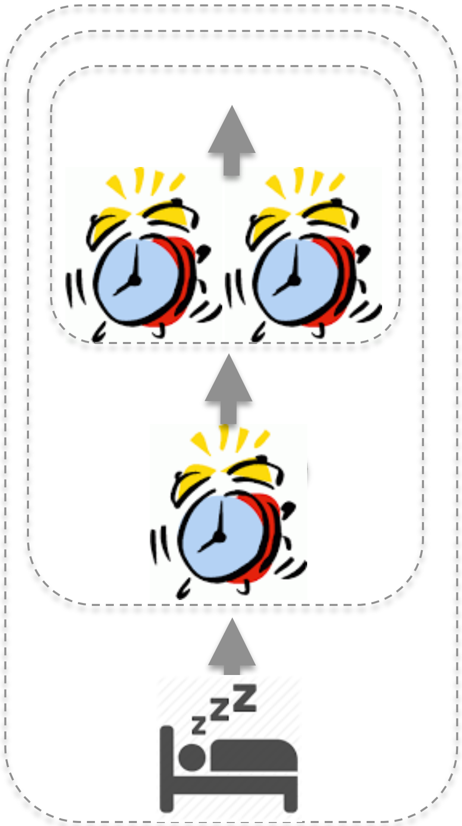


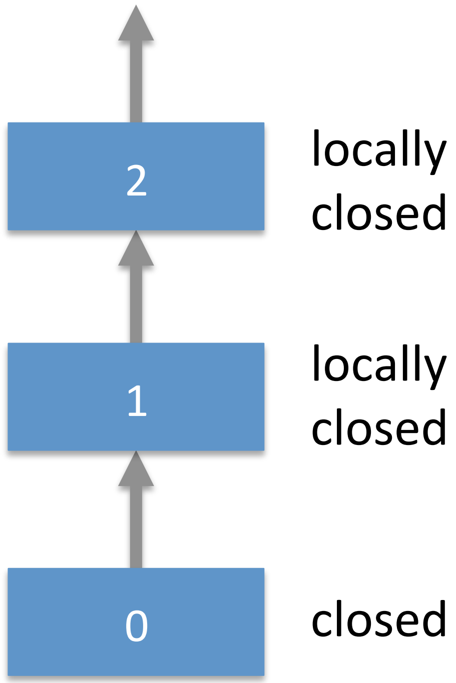


Example: Upward Topology

$\mathcal{Q} =$ How many alarms?

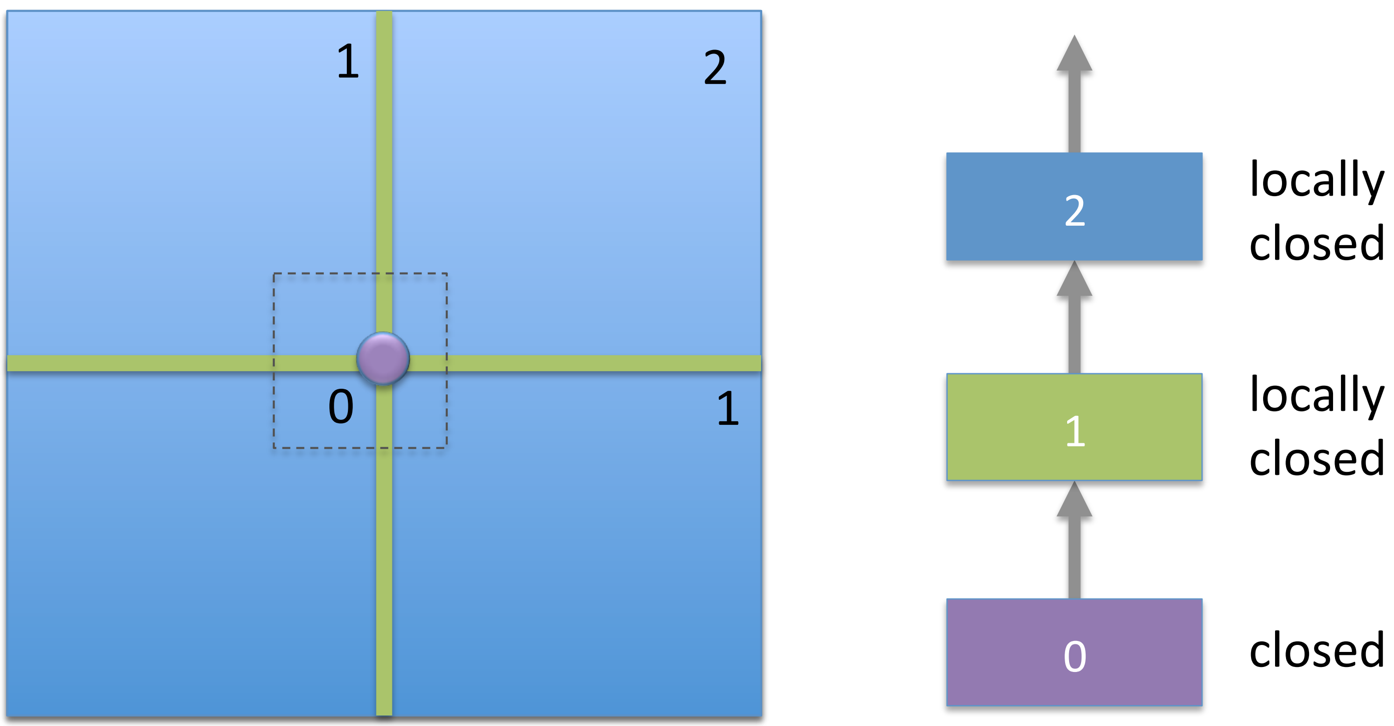
$\mathcal{I} =$ cumulative alarms.





Example: Euclidean Metric Topology

$\mathcal{Q} =$ How many parameters are free?



Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

$\mathcal{I} =$ finitely many inexact measurements.

degree

$$Y = \sum_{i=0}^{N} a_i X^i.$$

$$Y = a_0.$$

$$Y = a_0 + a_1 X.$$

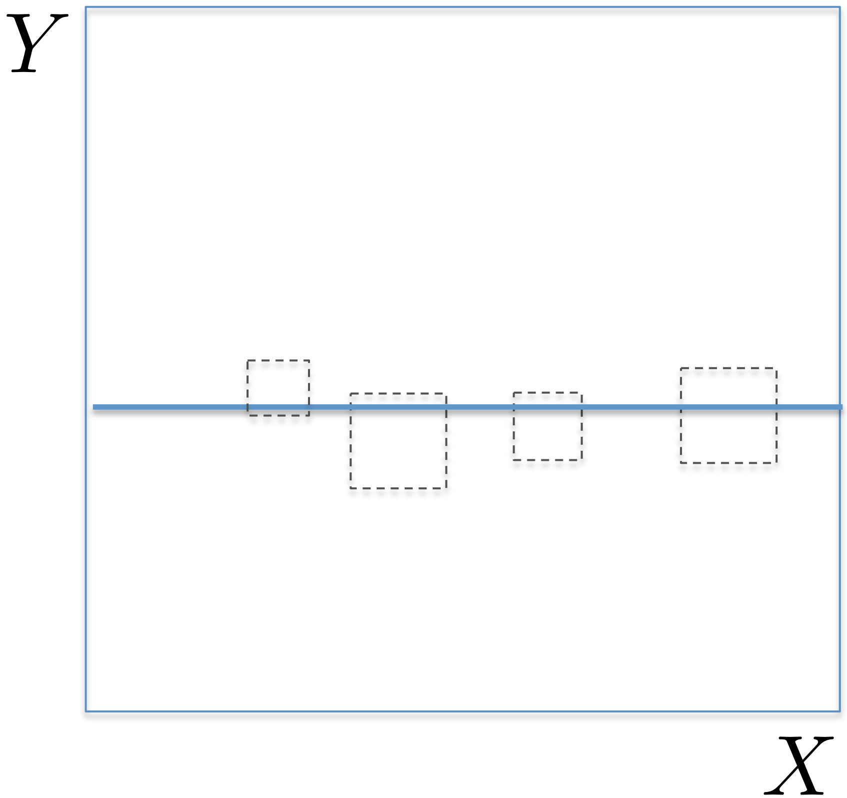
$$Y = a_0 + a_1 X + a_2 X^2.$$

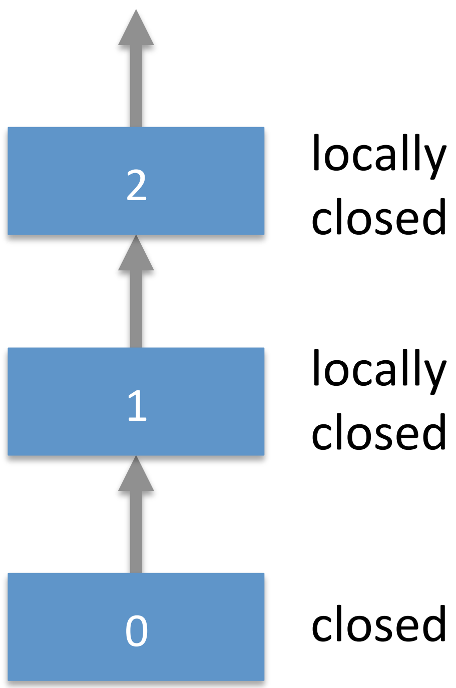
$$\vdots$$

Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

$\mathcal{I} =$ finitely many inexact measurements.

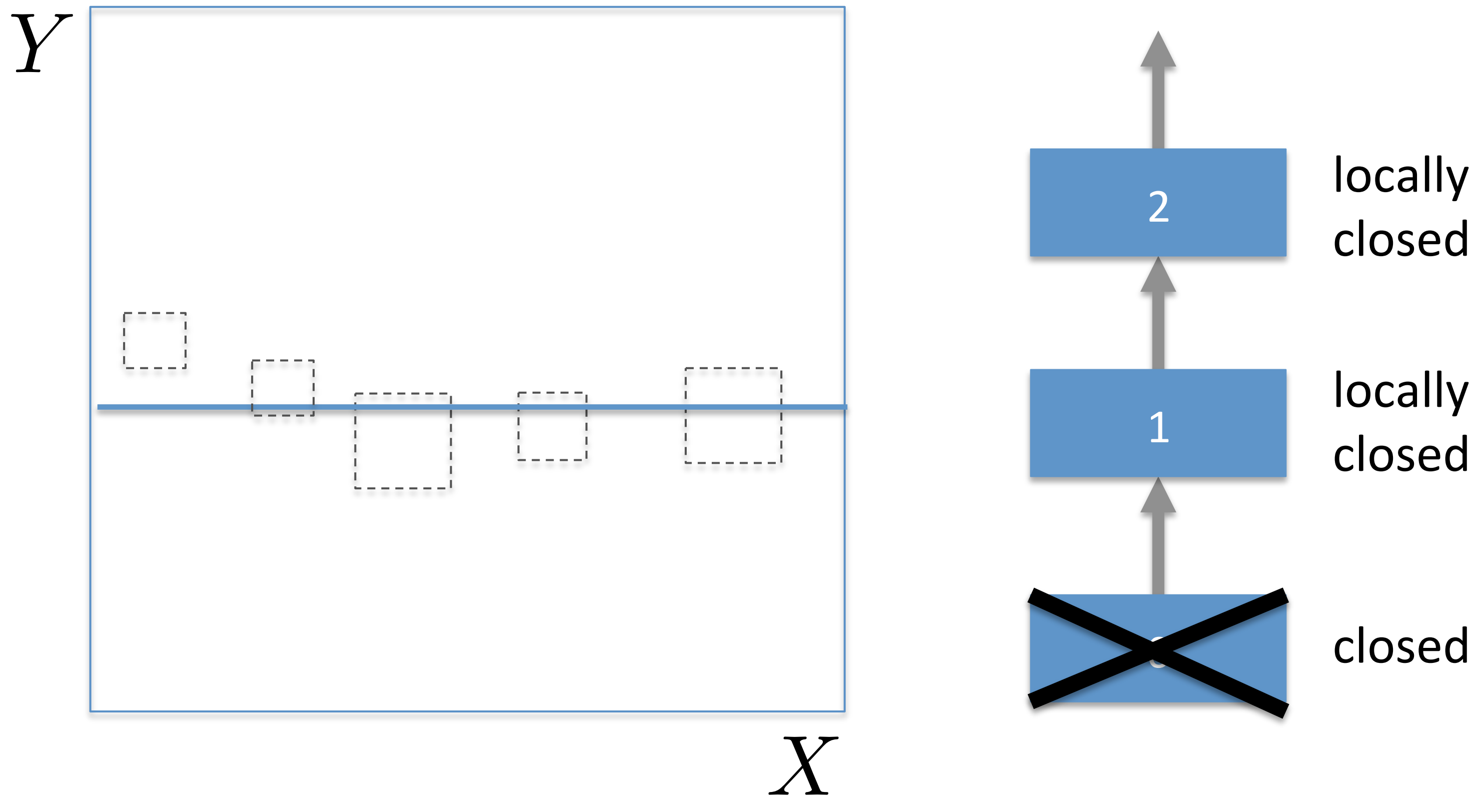




Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

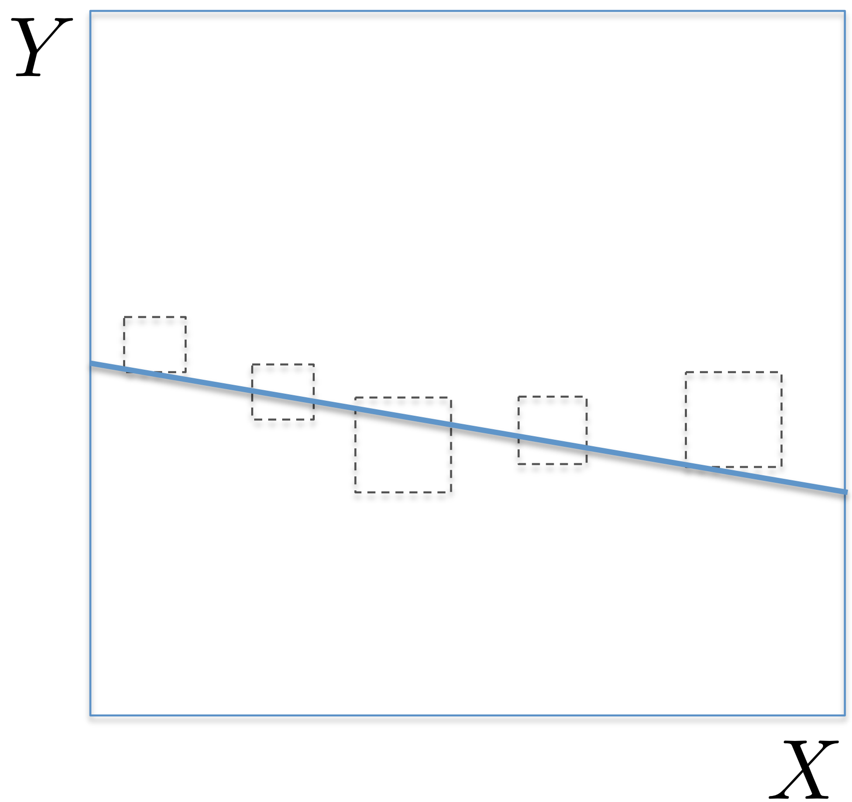
$\mathcal{I} =$ finitely many inexact measurements.

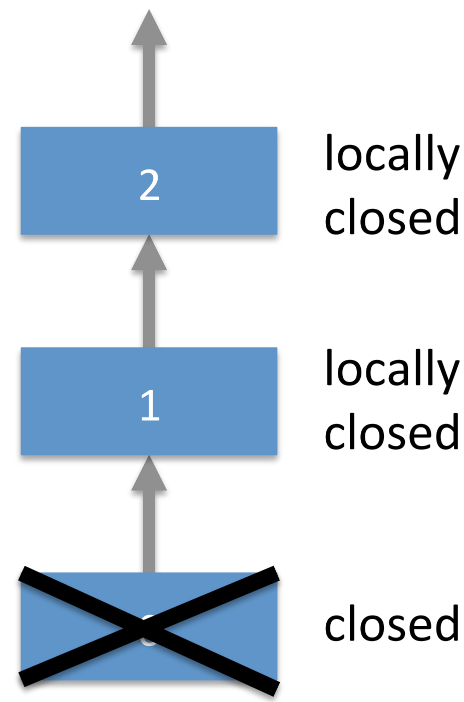


Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

$\mathcal{I} =$ finitely many inexact measurements.

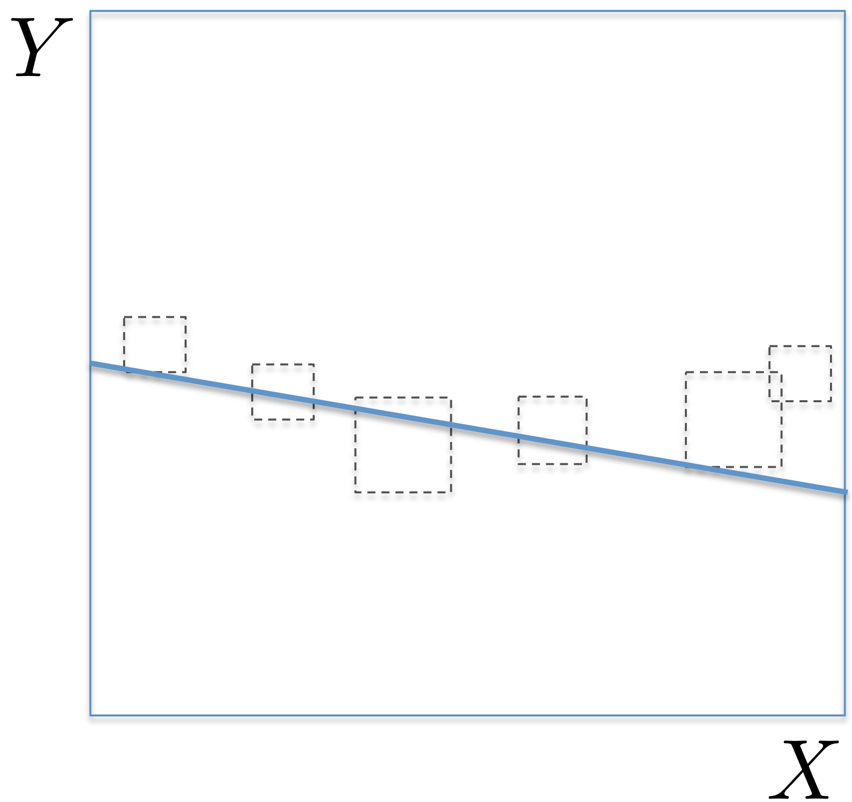


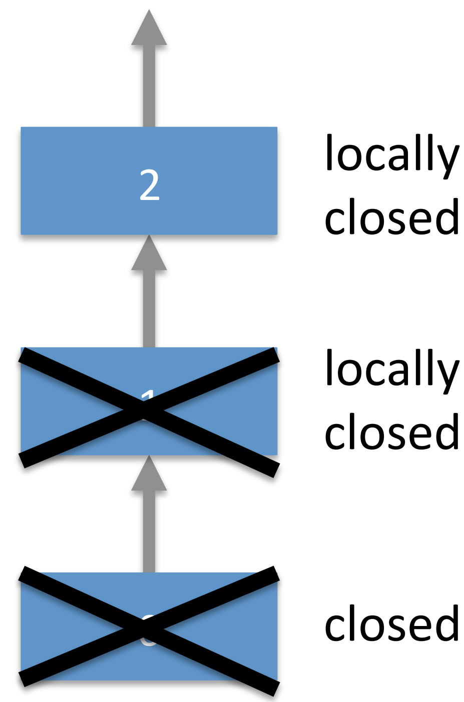


Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

$\mathcal{I} =$ finitely many inexact measurements.

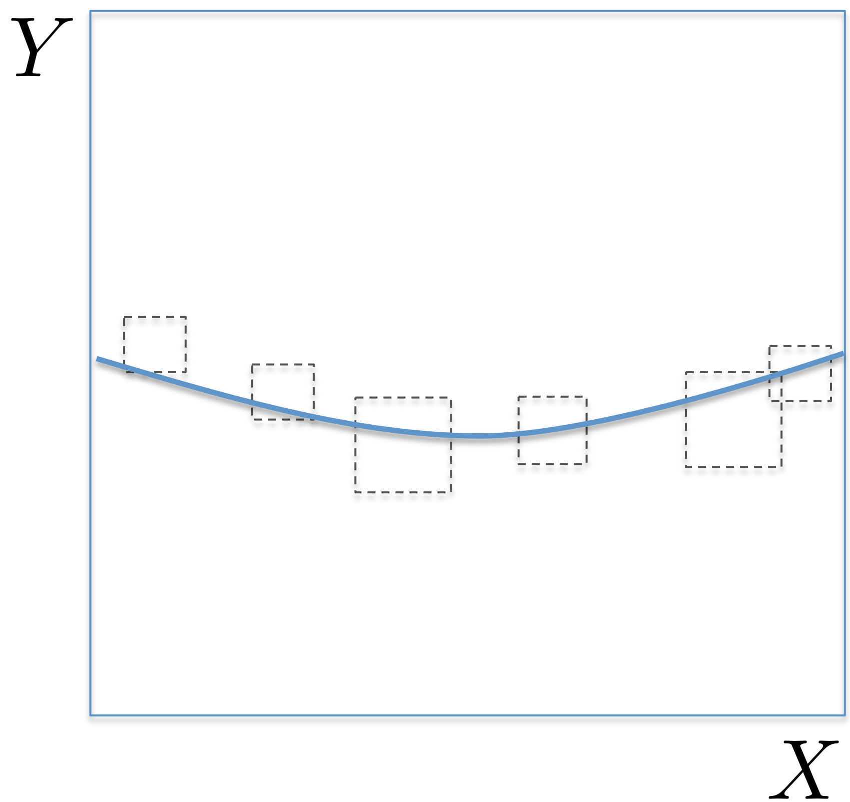


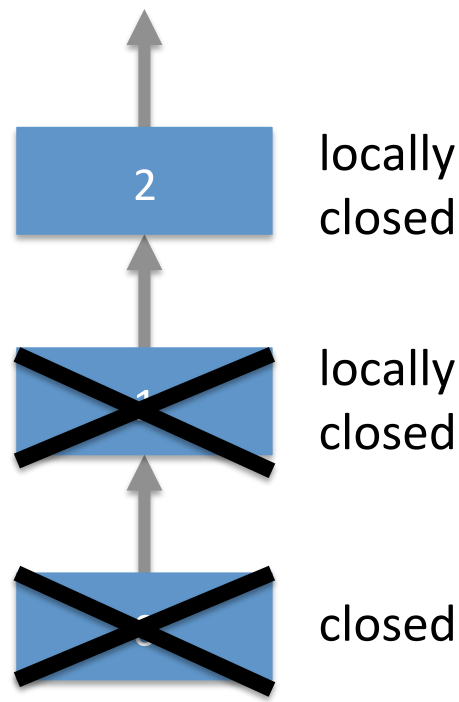


Example: Quantitative Laws

$\mathcal{Q} =$ What is the true polynomial degree?

$\mathcal{I} =$ finitely many inexact measurements.





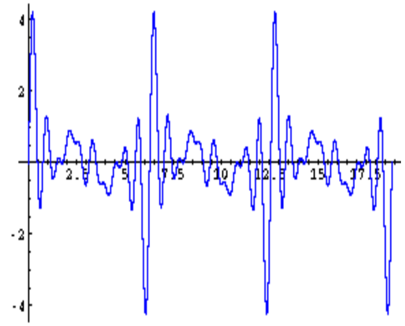
Example: Competing Paradigms

Polynomial paradigm

$$Y = \sum_{i=0}^{N} a_i X^i.$$

Trigonometric polynomial paradigm

$$Y = \sum_{i=0}^{N} a_i \sin(iX) + b_i \cos(iX).$$



Example: Competing Paradigms

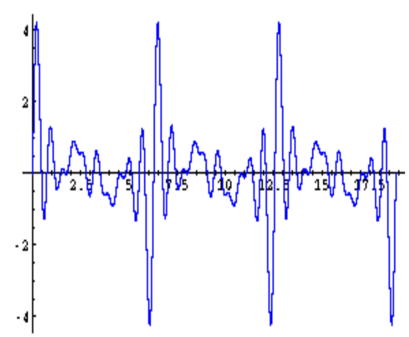
Polynomial paradigm

$$Y = \sum_{i=0}^{N} a_i X^i.$$

degree

Trigonometric polynomial paradigm

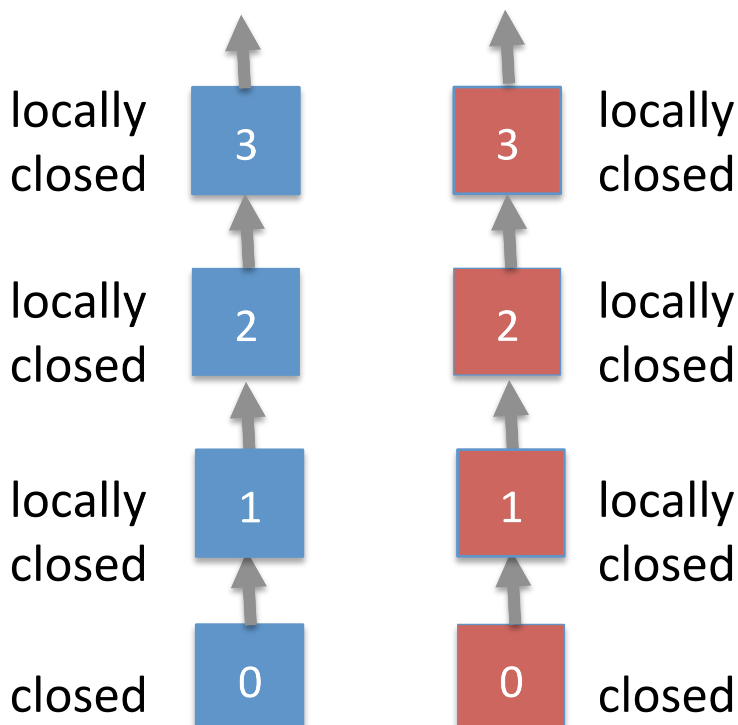
$$Y = \sum_{i=0}^{N} a_i \sin(iX) + b_i \cos(iX).$$

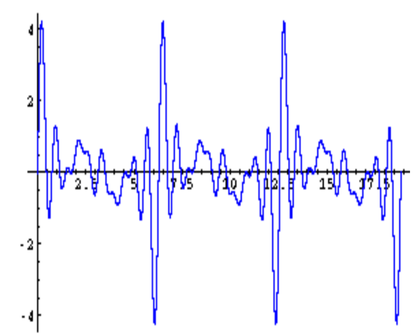


Example: Competing Paradigms

$\mathcal{Q}$ = which degree and which paradigm is true?

$\mathcal{I}$ = finitely many inexact measurements.





4. INDUCTIVE METHODS

Reasoning

Deductive

- Monotonic
- Non-ampliative

Reasoning

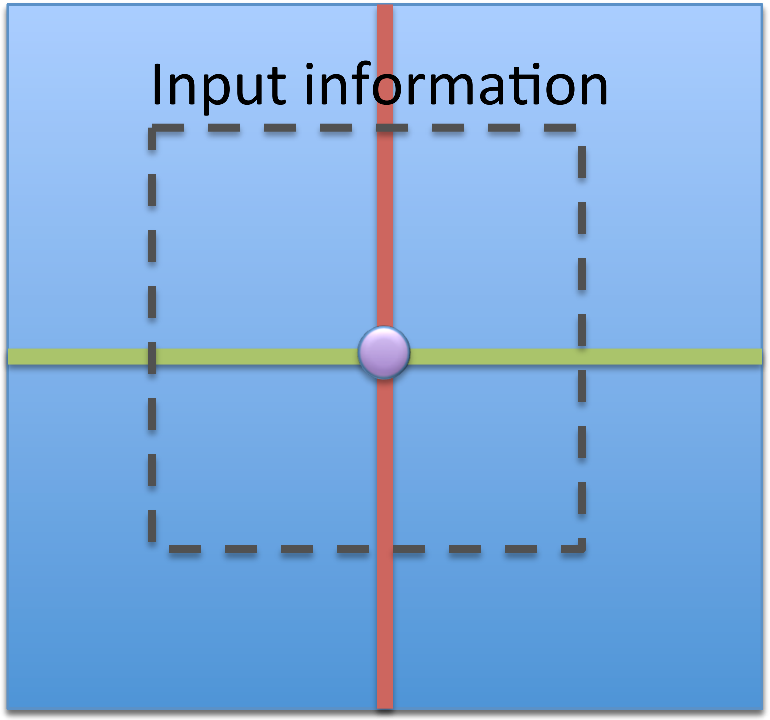
Deductive

- Monotonic
- Non-ampliative

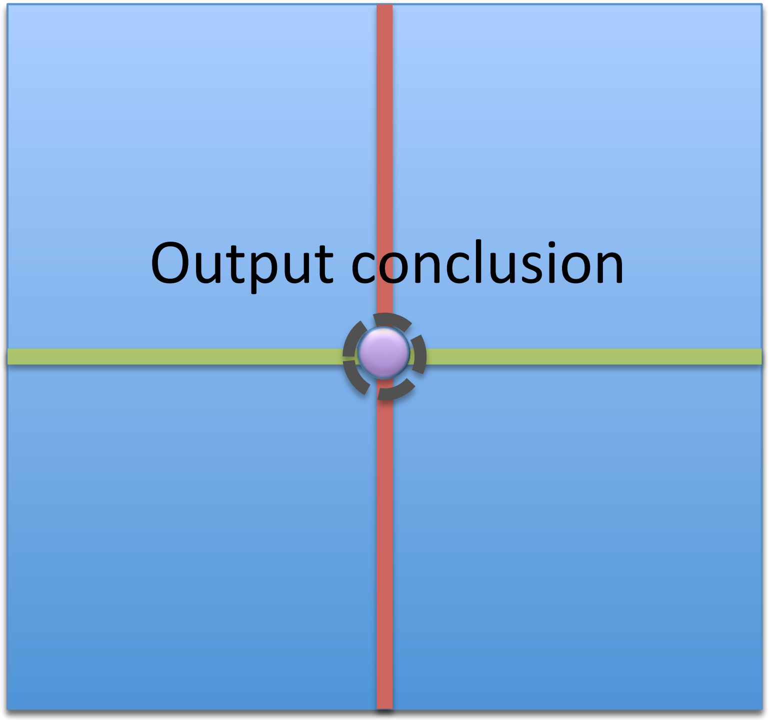
Inductive

- Non-monotonic
- Ampliative

Inductive Inference

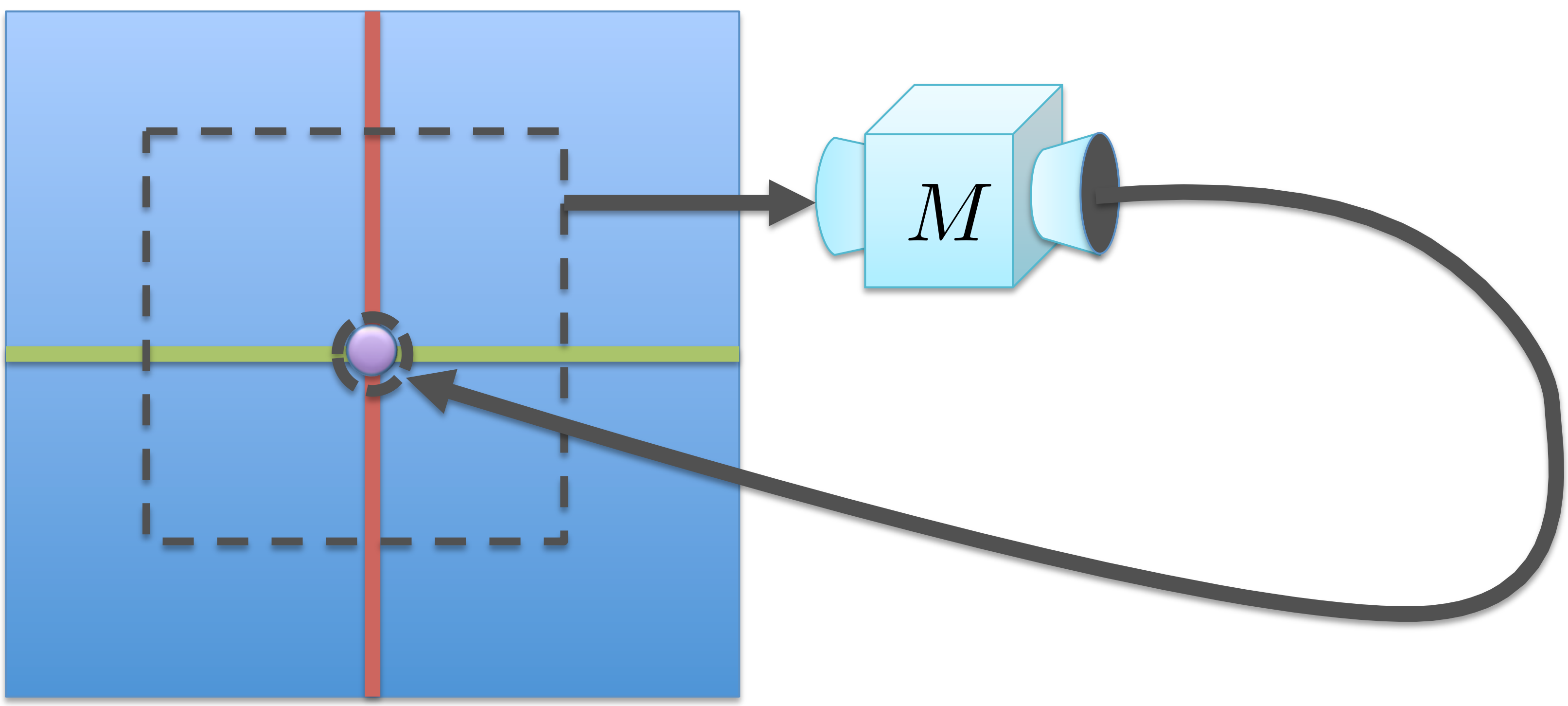


Inductive Inference



Inductive Methods

Information in, relevant response out.

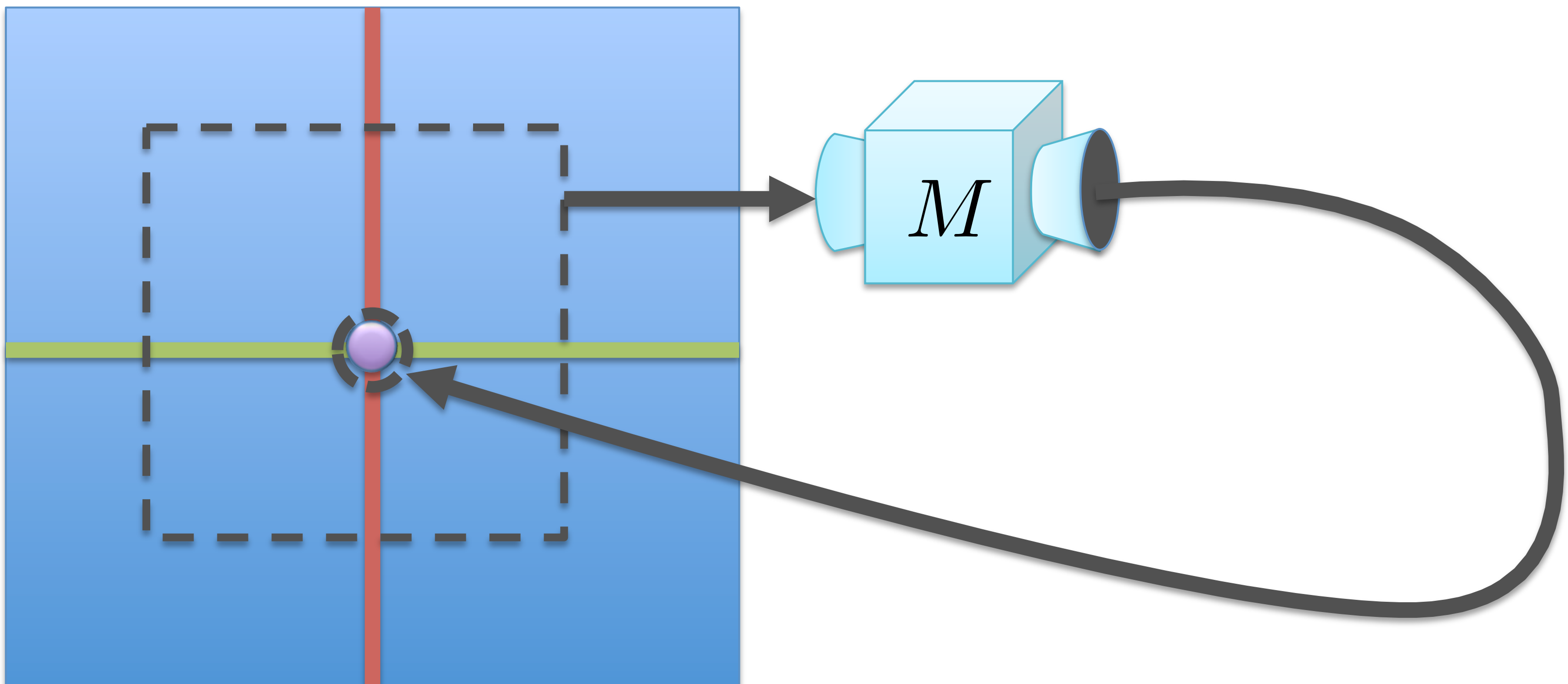


Solution in the Limit

M solves a problem in the limit iff

each world w presents information such that

M produces the true answer in w on any further information in w.

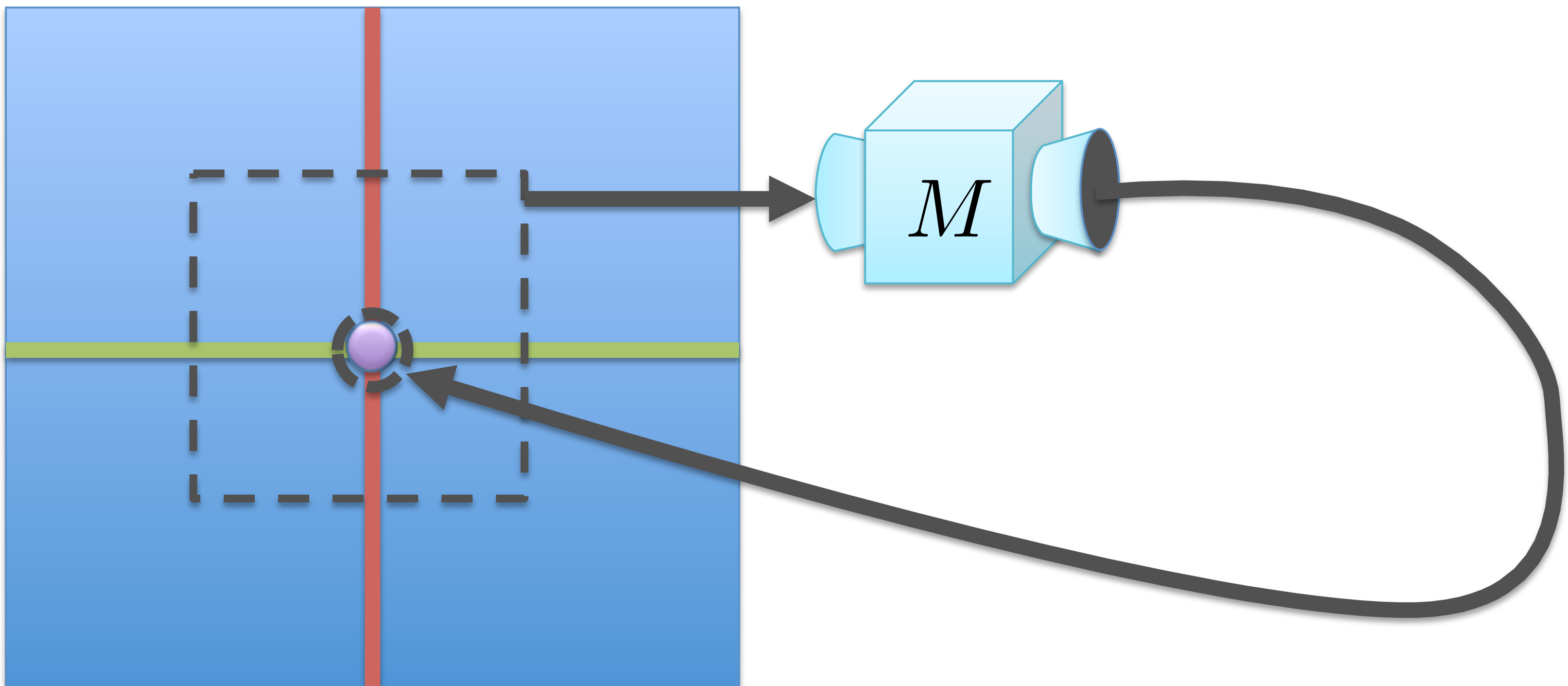


Solution in the Limit

M solves a problem in the limit iff

each world w presents information such that

M produces the true answer in w on any further information in w.

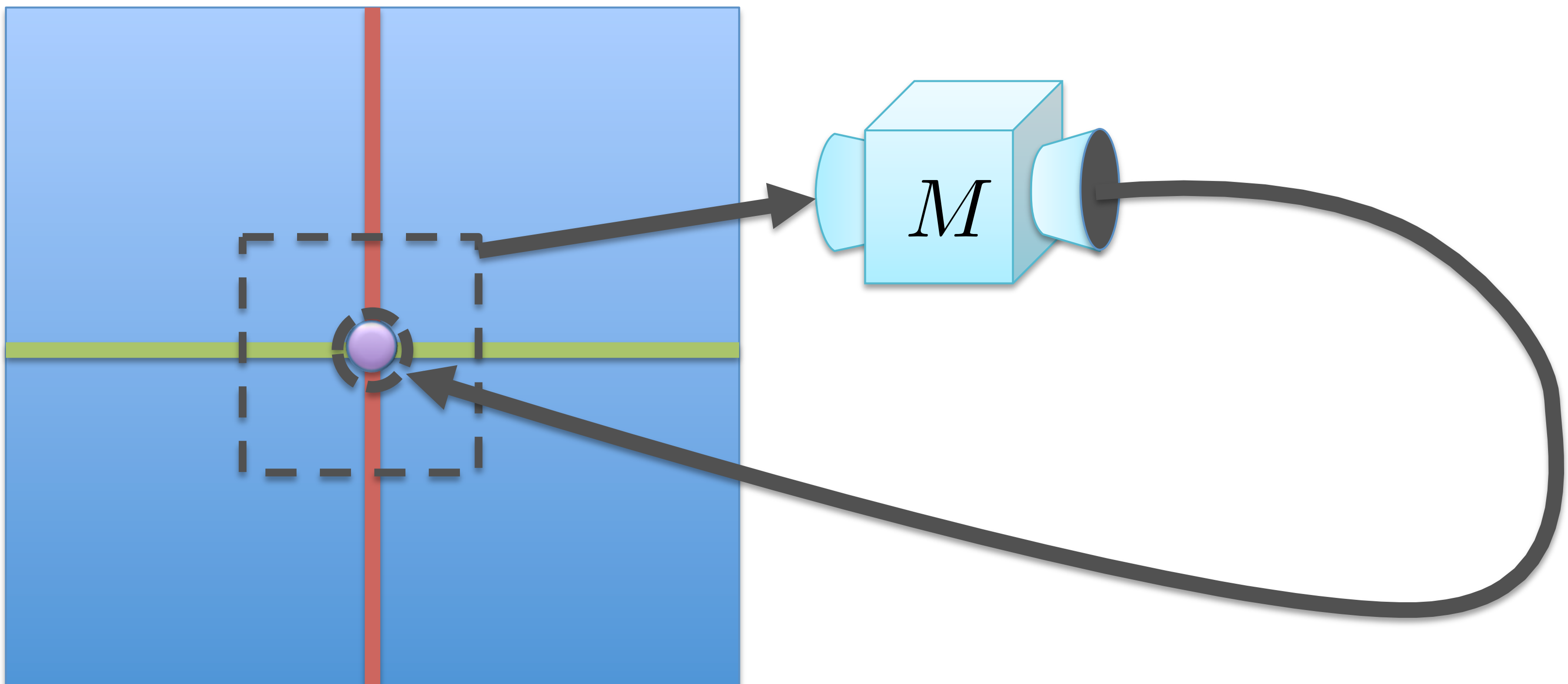


Solution in the Limit

M solves a problem in the limit iff

each world w presents information such that

M produces the true answer in w on any further information in w.

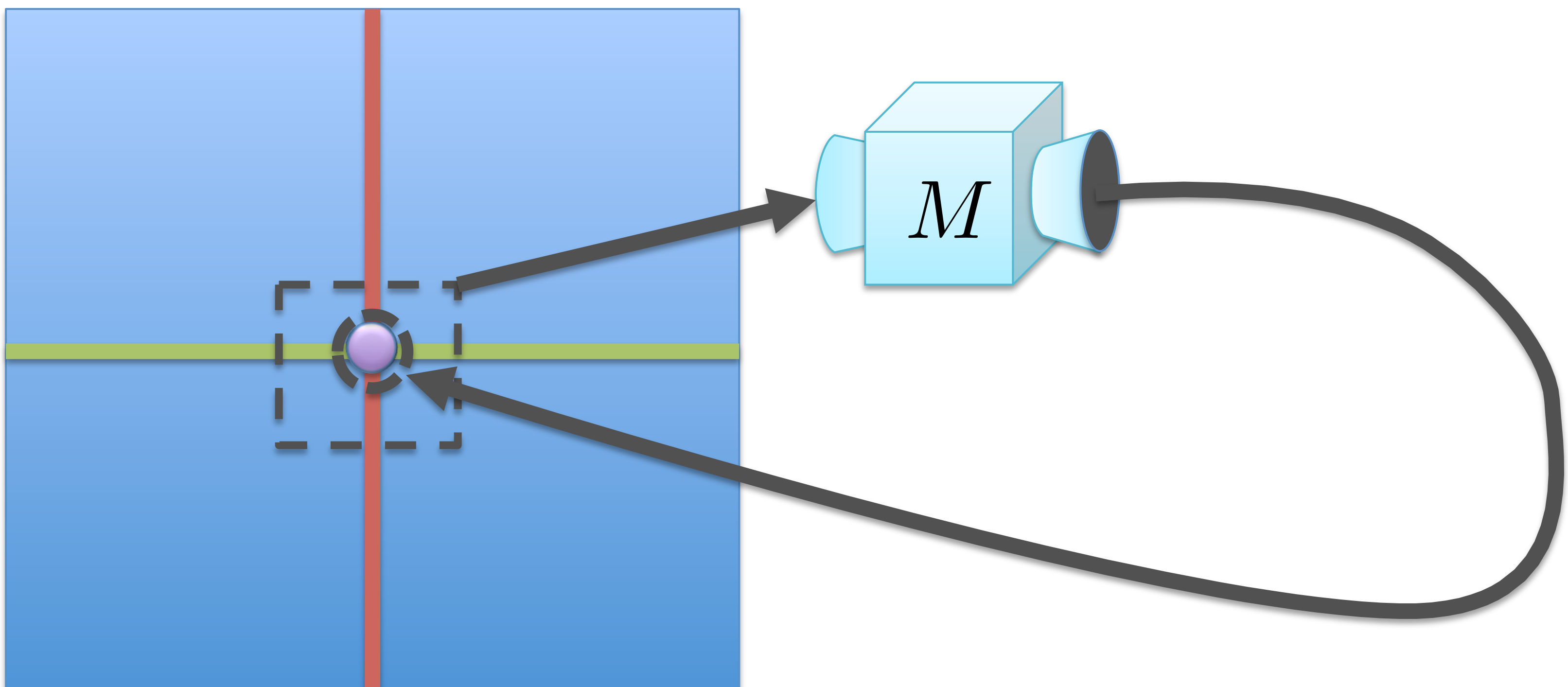


Solution in the Limit

M solves a problem in the limit iff

each world w presents information such that

M produces the true answer in w on any further information in w.



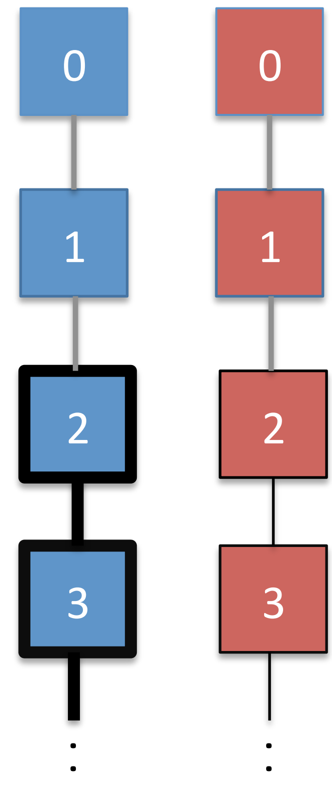
Reliability and Knowledge

If knowing the answer to a scientific question entails that your method is a solution, then whether you know the answer depends intrinsically on the question.

5. OCKHAM’S RAZOR

Ockham's Razor

- Output a simplest relevant response given E.
  - Allows for suspension of judgment.
  - Makes sense for infinite descending chains.



Popper's Razor

- Output a relevant response that is refutable (closed) given E.

Error Razor

- "Err on the side of simplicity".
- In arbitrary world w, never produce a relevant response B such that the true answer A_w is strictly simpler than B.

Equivalence

Proposition.

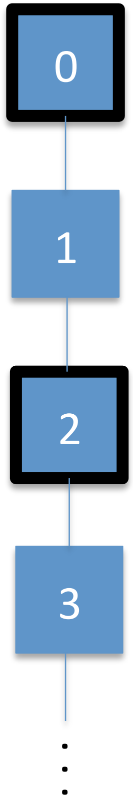
Ockham's razor = Popper's razor = error razor.

=

=

Patience

- Never rule out a simplest relevant response given E.
  - Says that simplicity is the only reason for inductive leaps beyond experience.
  - Logically independent of Ockham's razor.



Patient but not Ockham

Patience

- Never rule out a simplest relevant response given E.
  - Says that simplicity is the only reason for inductive leaps beyond experience.
  - Logically independent of Ockham's razor.

Ockham but not patient

Error patience

- In arbitrary world w, never output a relevant response that rules out all answers as simple as A_w.

Equivalence

- **Proposition:** Error patience is equivalent to patience.

6. OCKHAM'S RAZOR JUSTIFIED

Reliability

Deductive

- Converge to the truth directly

Reliability

Deductive

- Converge to the truth directly

Inductive

- Converge to the truth indirectly

Goldilocks Philosophy

Straight

Too strong!

Just right?

Arbitrarily crooked

Too weak!

Straightest Possible Convergence

Straight

Too strong!

Couldn’t be straighter

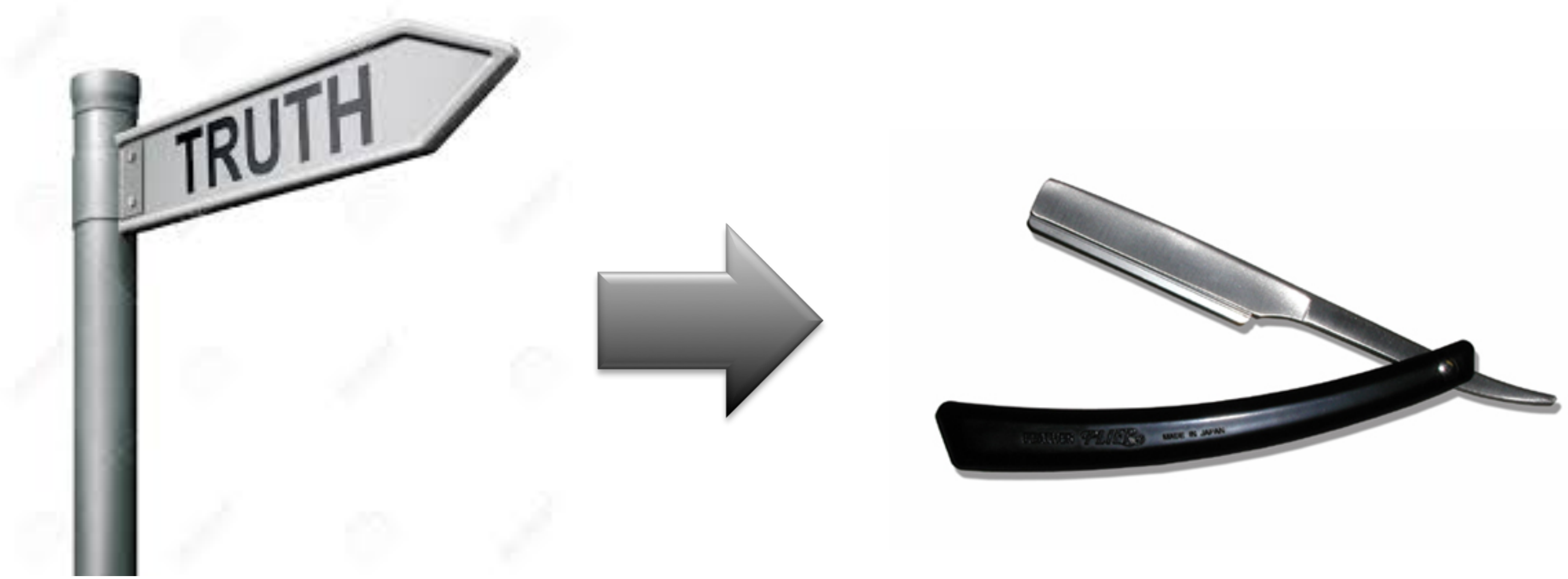
Just right!

Arbitrarily crooked

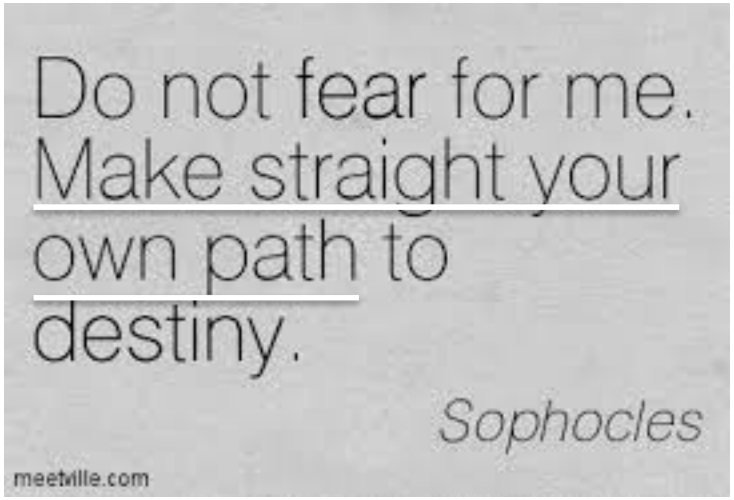
Too weak!

Thesis

- Ockham's razor is necessary for straightest convergence to the truth.



The Straightest Path



Fools dwelling in darkness, but thinking themselves wise, go round and round, by tortuous paths, like the blind led by the blind.

Katha Upanishad

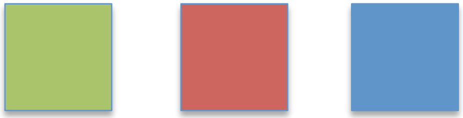
Two Departures from Straightness

Course-reversals

Cycles

Doxastic Reversal Sequence

- A finite sequence of relevant responses in which each entry contradicts its predecessor.



Doxastic Cycle Sequences

- A reversal sequence whose terminal entry entails its first entry.

Straightest Convergence

- M produces reversal [cycle] sequence

$$\mathbf{s} \;=\; (R_1, \ldots, R_k)$$

iff there exist information states

$$\mathbf{e} \;=\; (E_1 \supset \ldots \supset E_k)$$

such that

$$M(\mathbf{e}) = (M(E_1), \ldots, M(E_k))$$

is a reversal [cycle] sequence such that $M(E_i) \subseteq R_i$, for *i* from 1 to *k*.

Straightest Convergence

- Solution M is reversal [cycle] optimal iff:

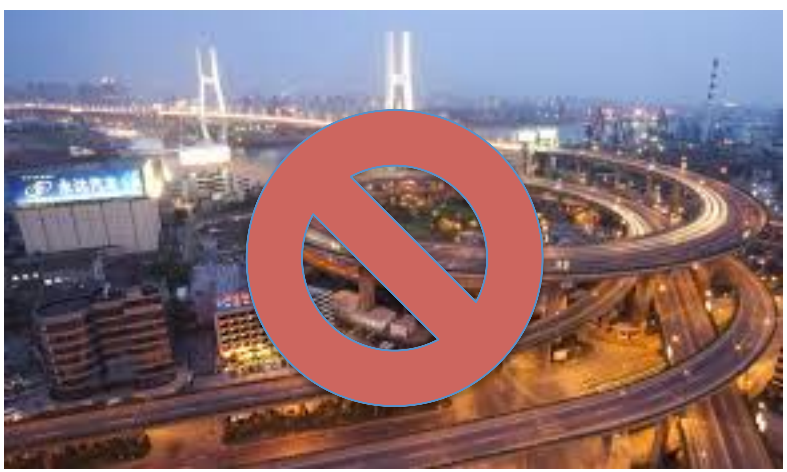
  every solution produces each reversal [cycle] sequence by M.

Main Result 1

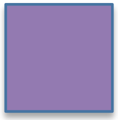
- **Proposition (Baltag, Gierasimczuk, and Smets):** Every solvable problem is refinable to a problem with a cycle-free solution.

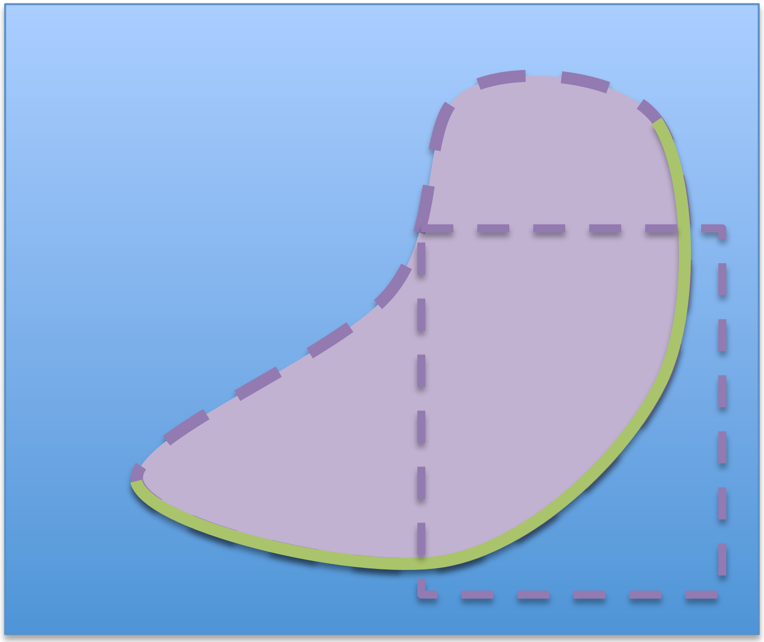
Main Result 2

- **Proposition:** Every cycle-free solution satisfies Ockham's razor.

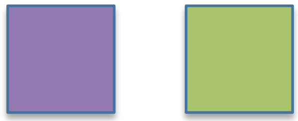


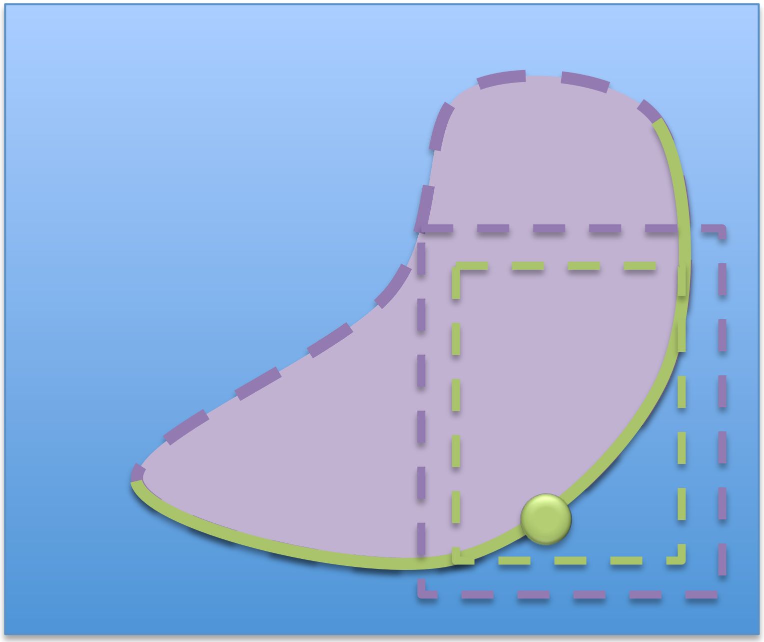
The Idea



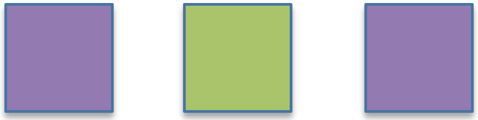


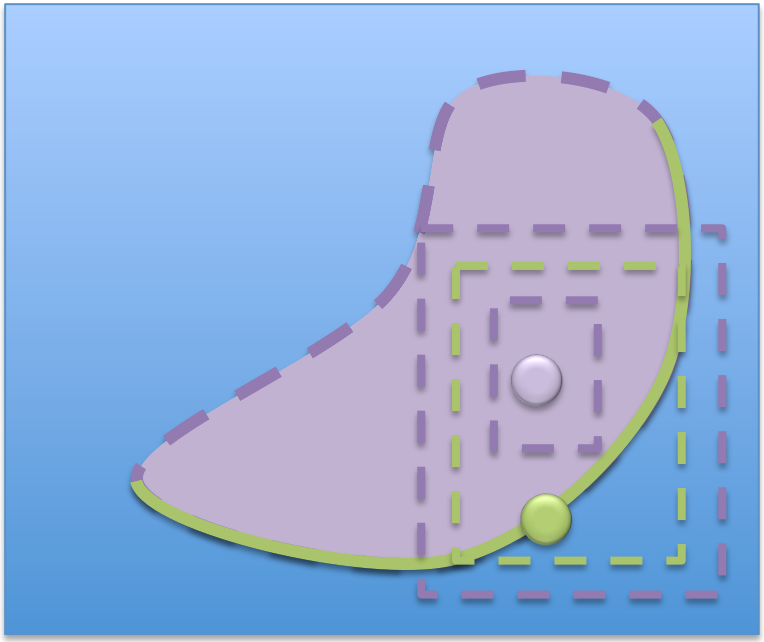
The Idea





The Idea





Main Result 3

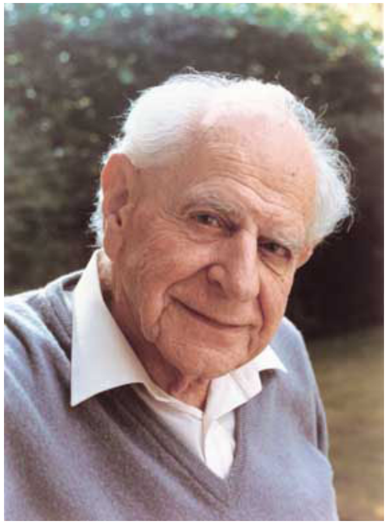
- We can characterize the solvable problems that have reversal optimal solutions.
- The characterization depends on the basis, so it is not topological.

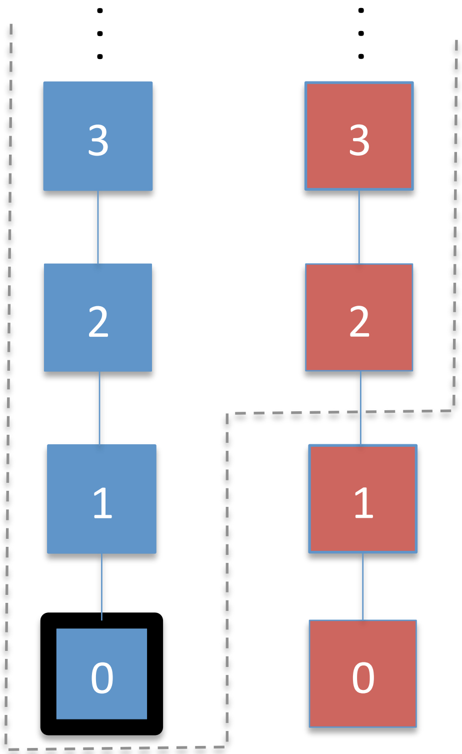
Main Result 3

- We can characterize the solvable problems that have reversal optimal solutions.

No Reversal Optimal Solution

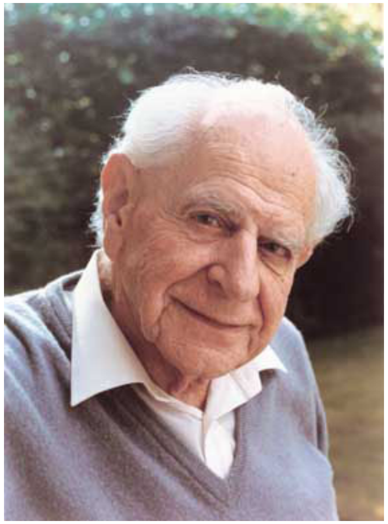
"Popper": choose the paradigm with fewer free parameters.

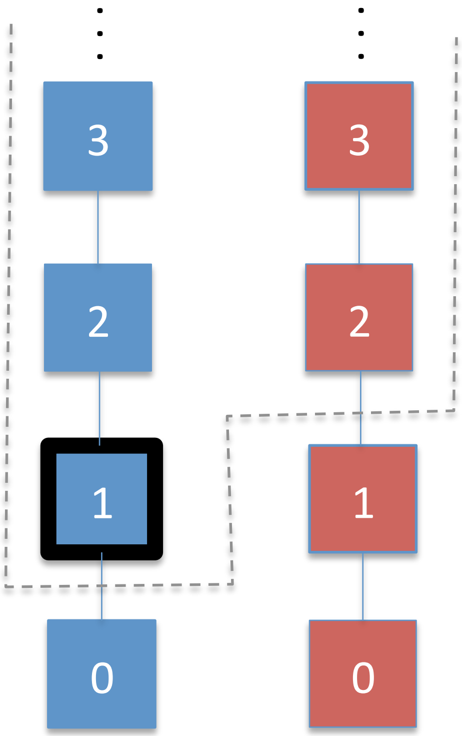




No Reversal Optimal Solution

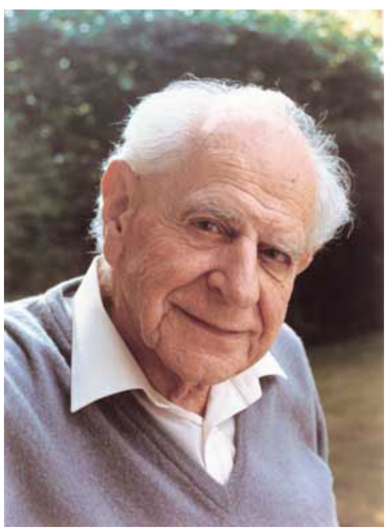
"Popper":
choose the paradigm with fewer free parameters.

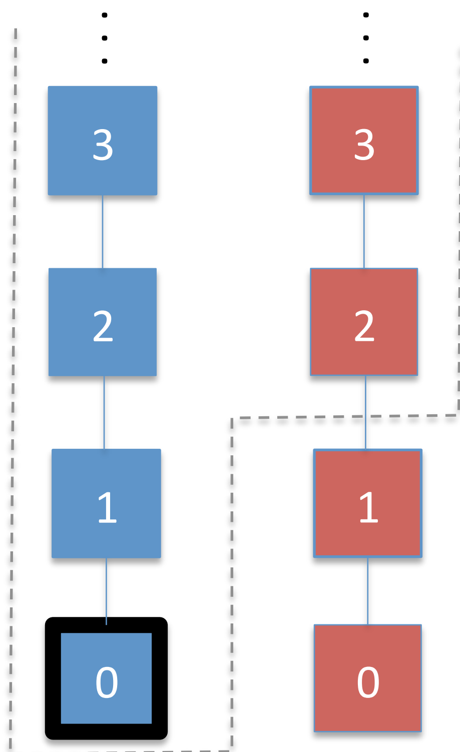




No Reversal Optimal Solution

"Popper":
choose the paradigm with fewer free parameters.

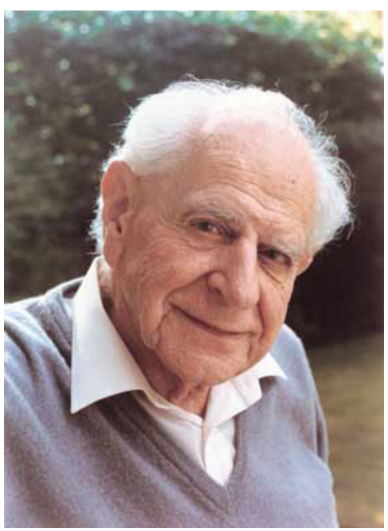


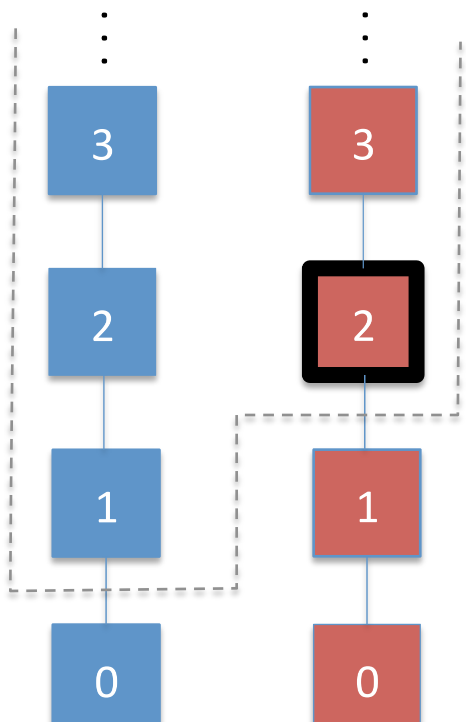


Lakatos:
choose the paradigm that was adjusted least recently.

No Reversal Optimal Solution

"Popper":
choose the paradigm with fewer free parameters.





Lakatos:
choose the paradigm that was adjusted least recently.

Main Result 4

- **Proposition:** a solution is reversal optimal only if it is patient.

Contextual Justification

- If patience is truth-conducive in your problem, its feasibility in some other problem is irrelevant.

Summary and Discussion

1. Simplicity is a topological feature of problems.
2. Ockham's razor is necessary for cycle-optimal convergence to the true answer.
3. Patience is necessary for reversal-optimal convergence to the true answer.
4. Optimally straight convergence is weak, but its implications for scientific method are strong.
5. The same holds for statistical inductive inference.
    1. Significance → a small tolerance for reversals and cycles.
    2. Power → drop theories you are destined to drop sooner.