

Title: How does simplicity help science find true laws?

Date: May 21, 2010 03:10 PM

URL: <http://pirsa.org/10050058>

Abstract:

Simplicity Truth Zen

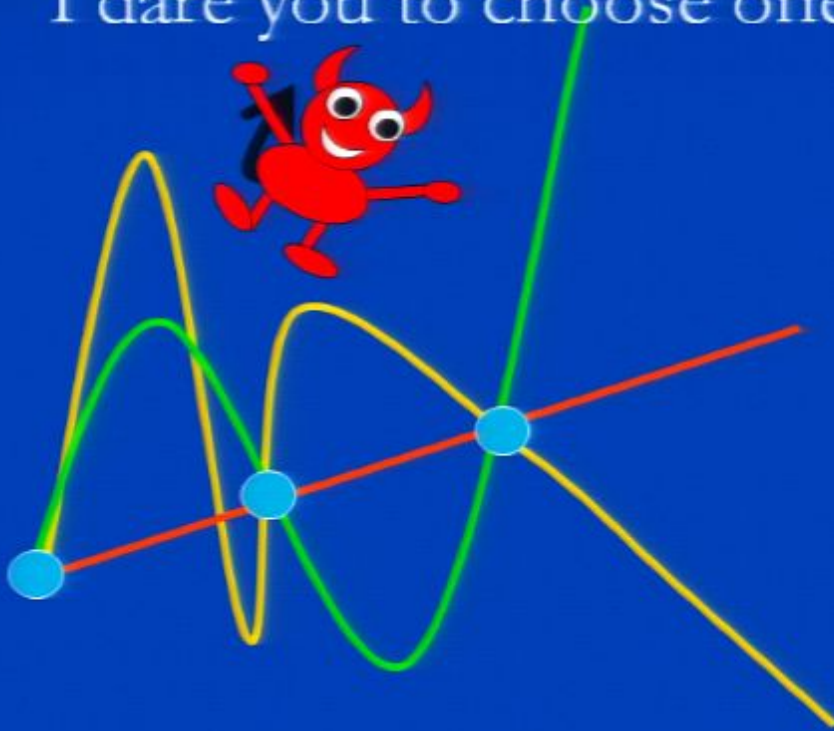
Kevin T. Kelly
Conor Mayo-Wilson
Hanti Lin

Department of Philosophy
Carnegie Mellon University

<http://www.andrew.cmu.edu/user/kk3n/ockham/Ockham.htm>

Theoretical Underdetermination

I dare you to choose one!



Western Epistemology

- The demon **undermines** science, so **defeat** him.

I smite thee with **scientific rationality**!



Western Epistemology

- The demon **undermines** science, so **defeat** him.
- Portray him as **weak**.

You can't fool me!



Zen Epistemology

- The demon justifies science with his strength.
- Stronger demons justify weaker inferences!



Zen Epistemology

How is this supposed to work, eh?



Zen Epistemology

- If you **could** do better, you **should**.
- But because of me, you **can't**.
- So you are **optimal** and, hence, **justified**.



Zen Epistemology

Isn't that how **computer scientists** routinely evaluate **algorithms**?

Yes.



Zen Epistemology

But computing is **deduction**.

Isn't **induction** a completely different animal?

No. That's what
**formal learning
theory** is about.



Zen Epistemology

Cool.

But how does it apply to science?



Ockham's Razor

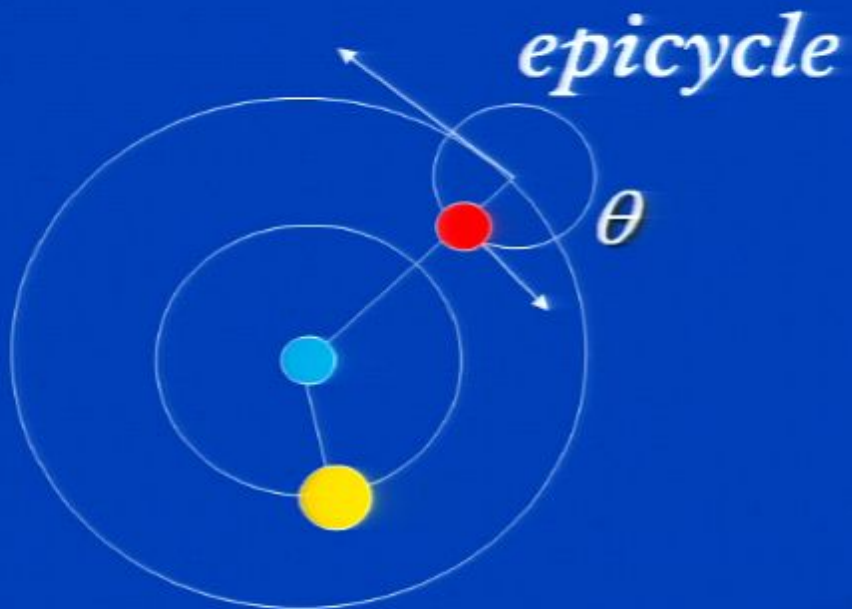
Astronomy 1543

Planetary retrograde motion



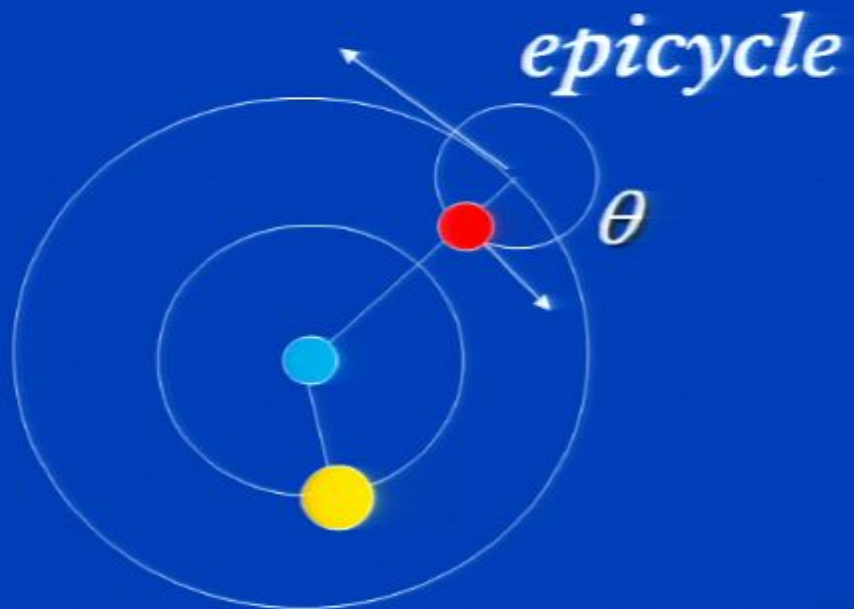
Ptolemaic Explanation

- Retrograde path is due to an epicycle.
- Can happen regardless of the position of the sun



Copernican Explanation

■ “Lapped” competitor appears to backtrack against the bleachers.



Copernicus Victorious

- Explains the data rather than merely accommodating them.
- Has fewer adjustable parameters.
- Is more falsifiable (failure of correlation would rule it out).
- Is more unified.
- Is simpler.

More Victories for Simplicity

- universal gravitation vs. divided cosmos
- wave theory of light vs. particles and aether
- oxygen vs. phlogiston
- natural selection vs. special creation
- special relativity vs. classical electrodynamics
- Dirac's equation
- chaos vs. infinite series of random variables

Puzzle

- An indicator must be **sensitive** to what it indicates.

simple



Puzzle

- An indicator must be **sensitive** to what it indicates.



complex

Puzzle

- An indicator must be **sensitive** to what it indicates.

simple



Puzzle

- An indicator must be **sensitive** to what it indicates.

simple



Puzzle

- But Ockham's razor **always** points at simplicity.

simple



Puzzle

- How can a broken compass help you find something unless you already know where it is?



complex

Metaphysicians for Ockham

Somehow, I don't like the sound of that...



Information Channel

Simplicity bias ← Simple reality

Mystical vision
Natural selection??

Metaphysicians for Ockham

Somehow, I don't like the sound of that...



Information Channel

Simplicity bias ← Simple reality

Mystical vision
Natural selection??

Pre-established Harmony



Idealism

Simplicity bias  Simple reality

Metaphysicians for Ockham

So before you can use your razor to eliminate
Metaphysical causes, you have to assume metaphysical causes.



Theoretical “Virtues”

Simpler theories:

- . Are more unified;
- . Explain better;
- . Are more falsifiable.

But the truth might **not** be virtuous.

Wishful thinking to assume otherwise.

Statistical Explanations

1. Prior Simplicity Bias

Bayes, BIC, MDL, MML, etc.

2. Risk Minimization

SRM, AIC, cross-validation, etc.

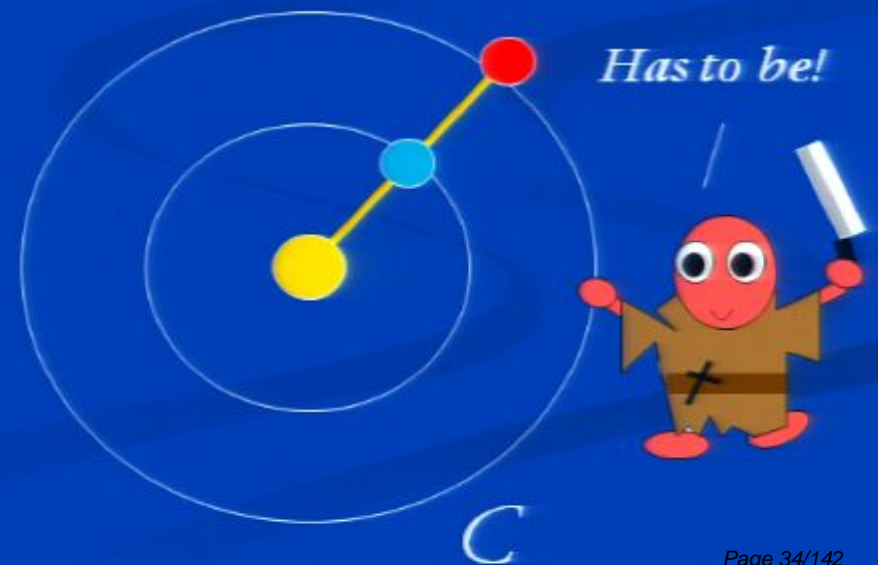
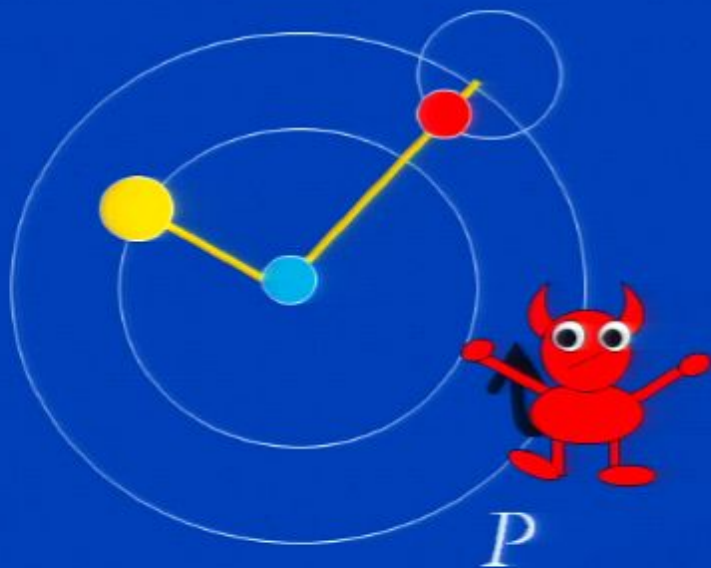
Prior Simplicity Bias

I am inclined to believe the **simpler** theory
because I am inclined to believe that the
world is **simple**.



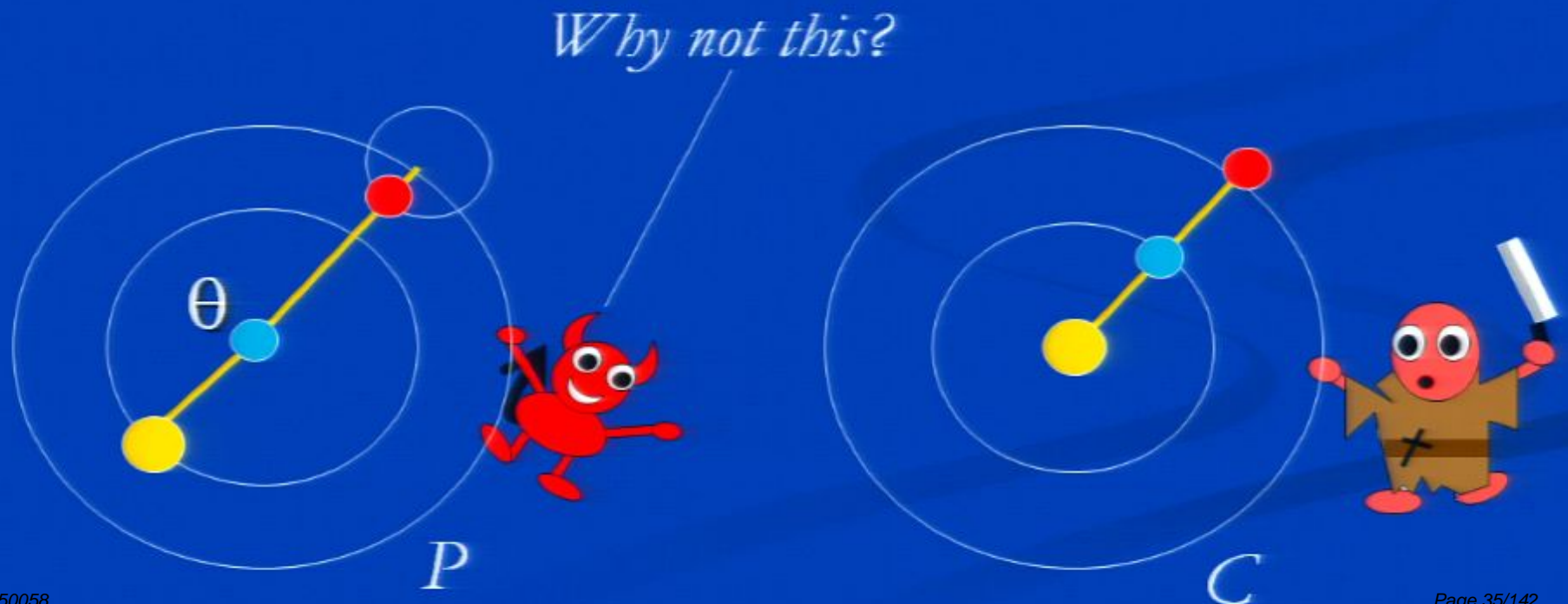
More Subtle Version

- Simple data are a *miracle* in the complex theory but not in the simple theory.



However...

- correlation would **not** be a miracle given $P(\theta)$;



The Real Miracle

Ignorance about model:


$$p(C) \approx p(P);$$

+ Ignorance about parameter setting:

$$p'(P(\theta) \mid P) \approx p(P(\theta') \mid P).$$

= **Knowledge** about worlds.

$$p(P(\theta)) \ll p(C).$$

| CP | |
|--|----------|
|  | θ |
| | θ |
| | θ |
| | θ |
| | θ |
| | θ |
| | θ |
| | θ |



Ignorance is knowledge.

= Paradox of Indifference

Ignorance of *red* vs. not-*red*

+ Ignorance over not-*red*:

= **Knowledge** about *red* vs. *white*.



Knognorance =

All the **priveleges** of knowledge

With none of the **responsibilities**

I'm for it!



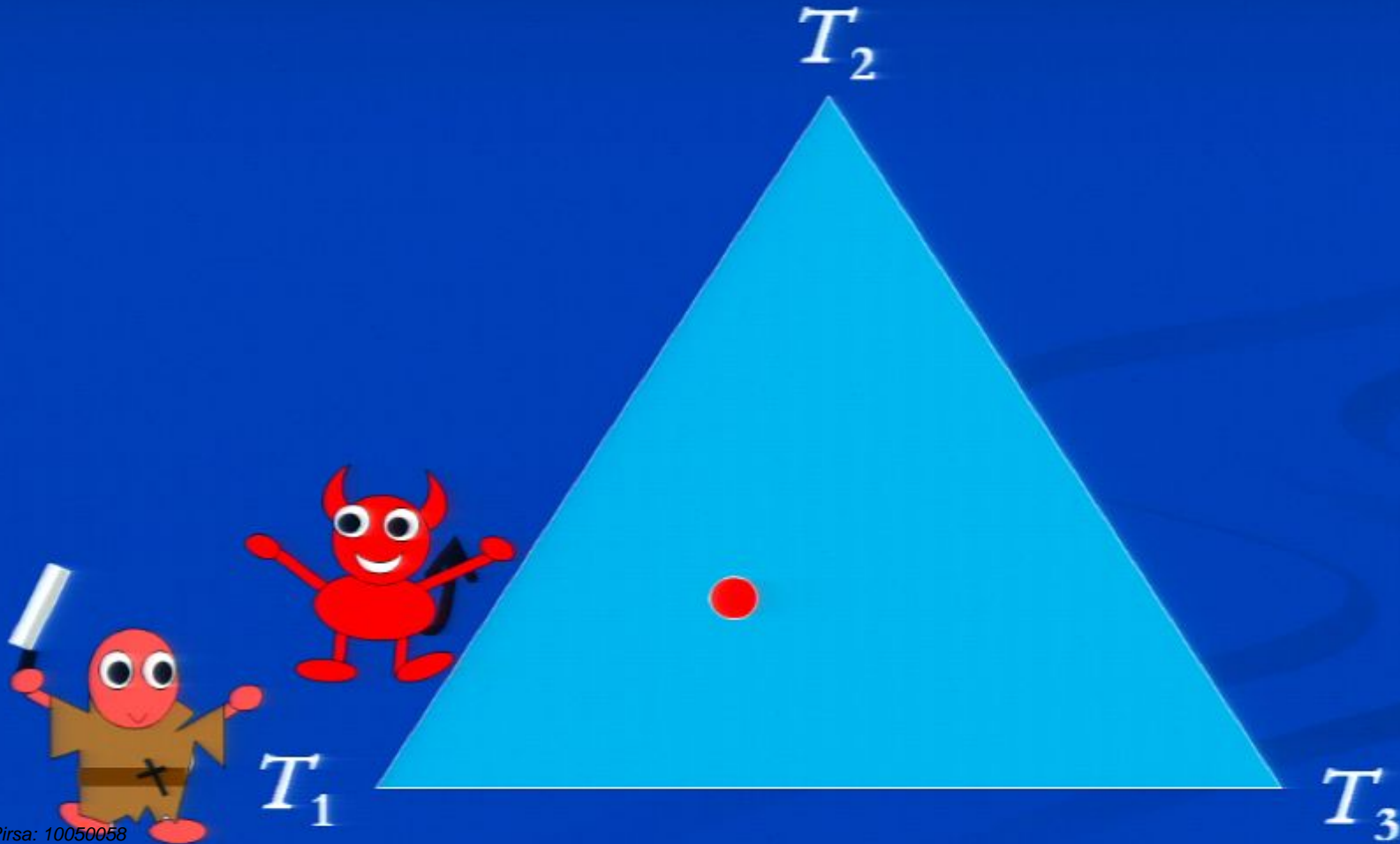
In Any Event

The coherentist foundations of Bayesianism
have **nothing to do with short-run truth-
conduciveness.**

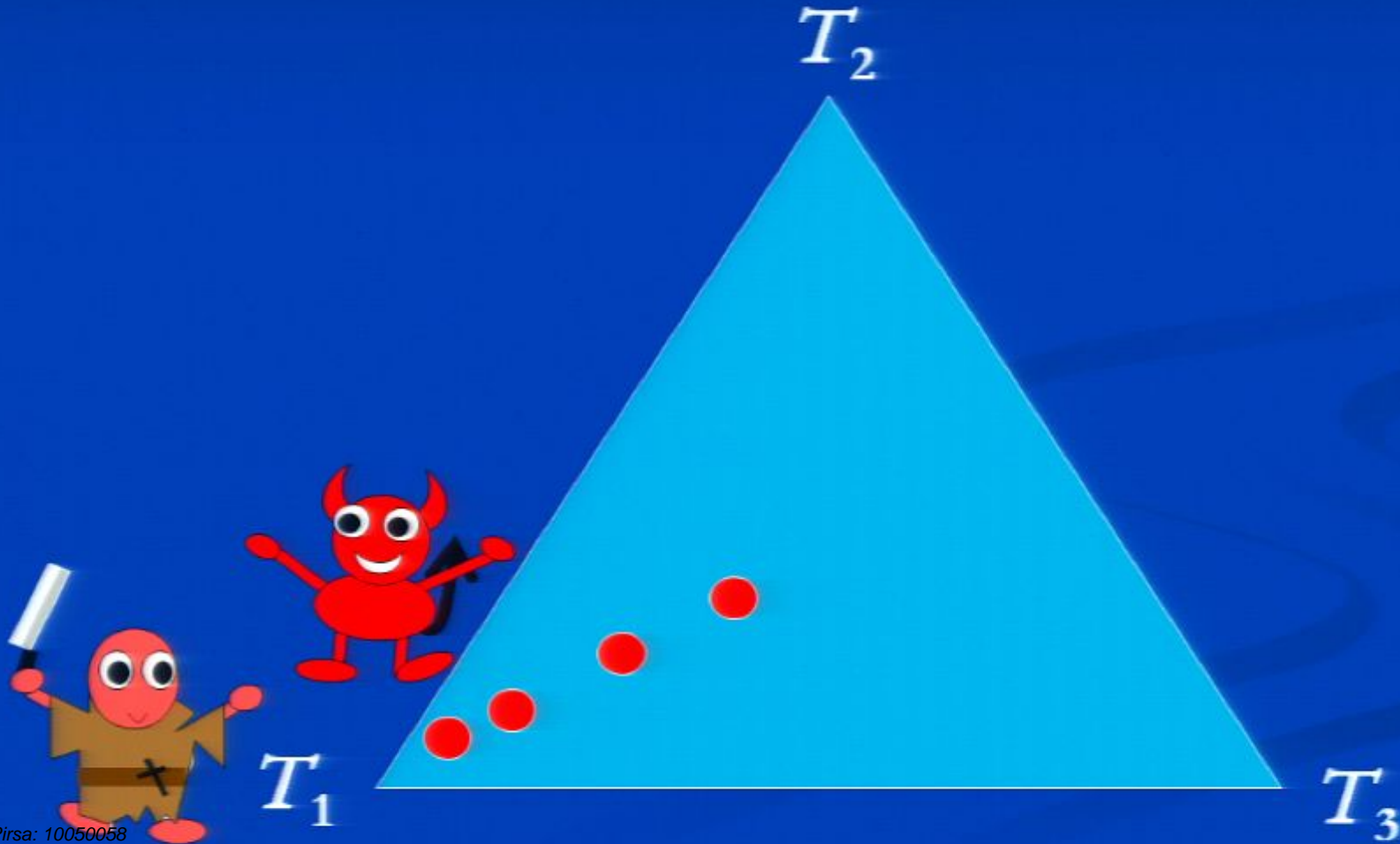
Not so loud!



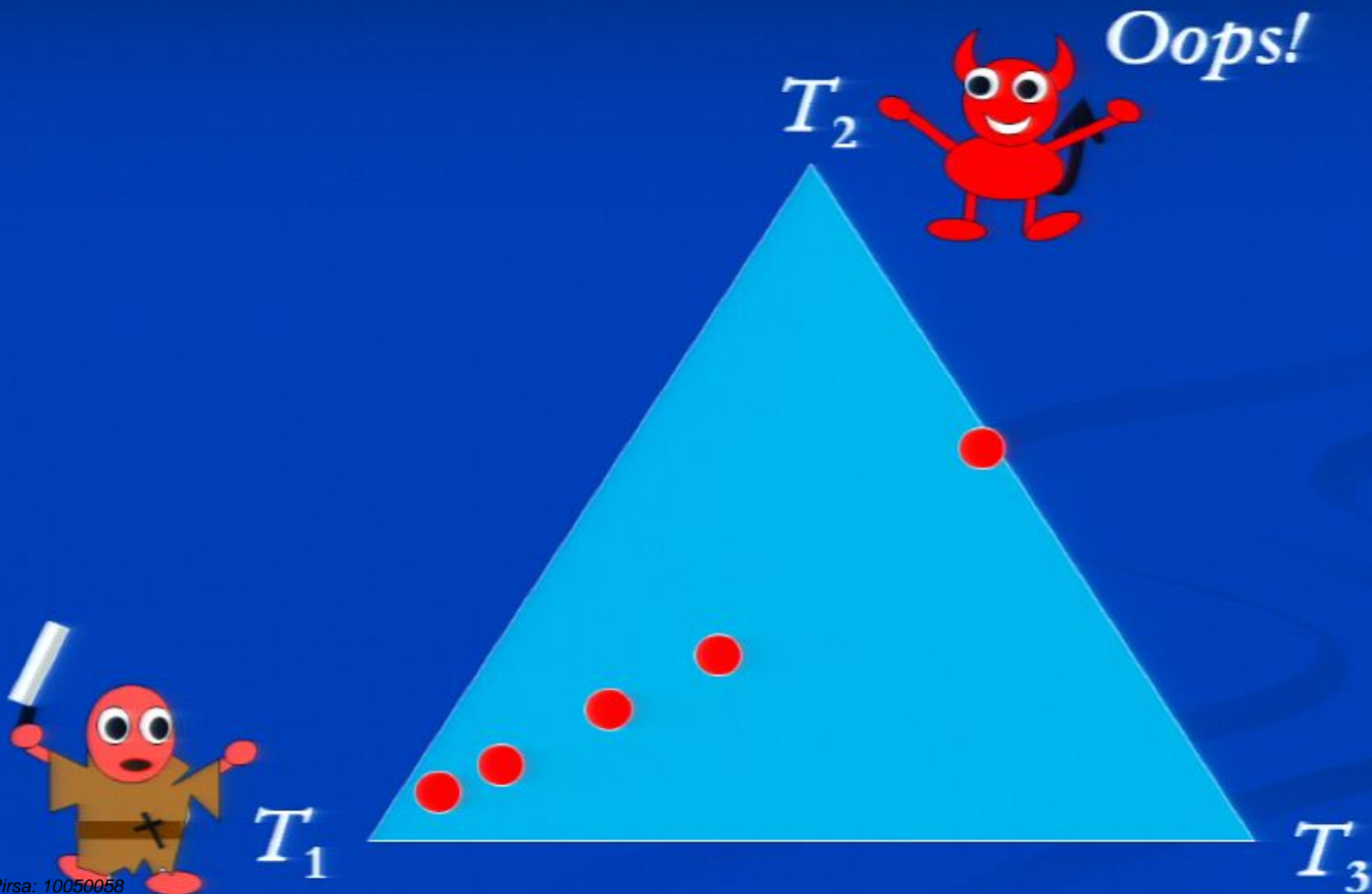
Ockham Bayesians Converge



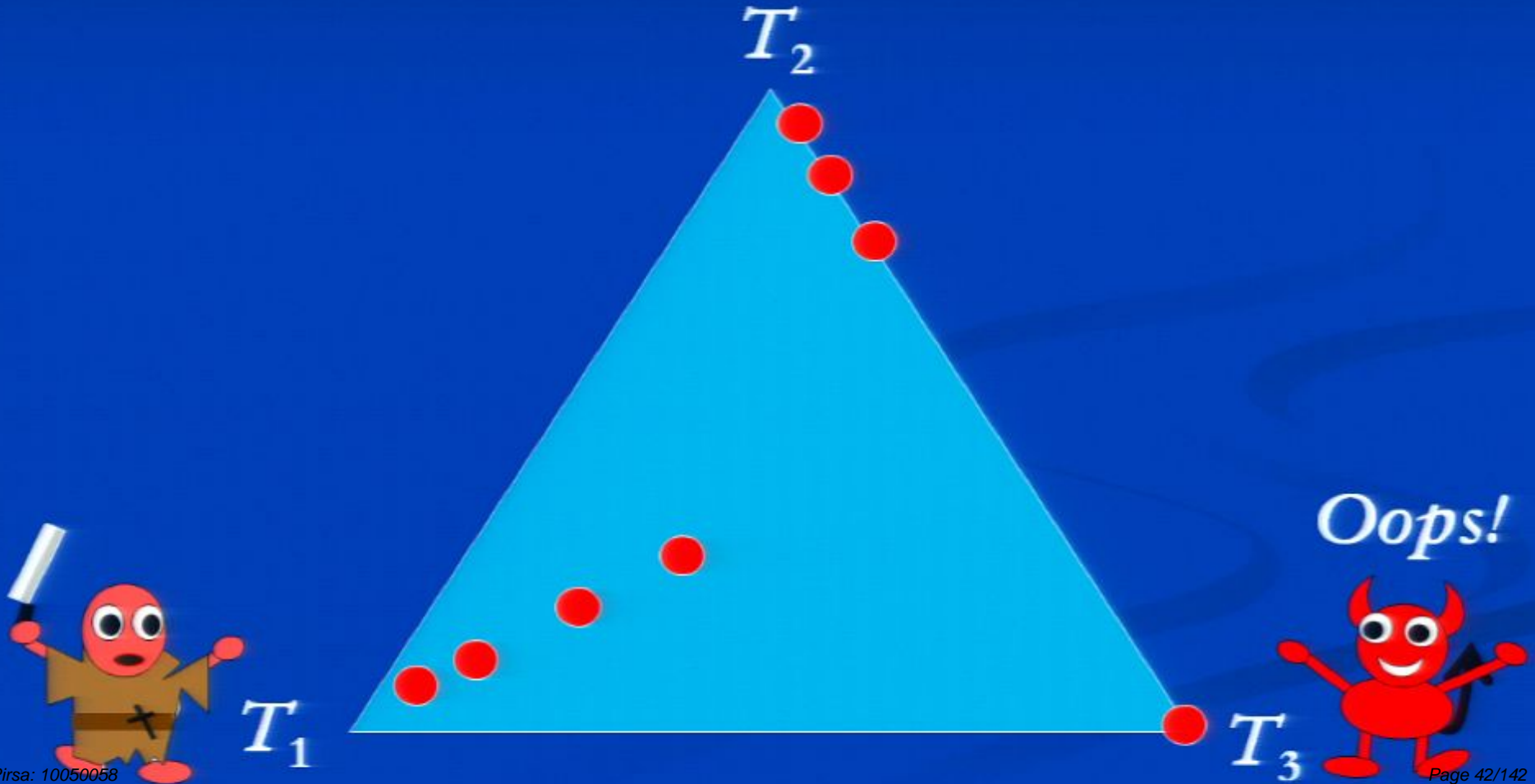
Ockham Bayesians Converge



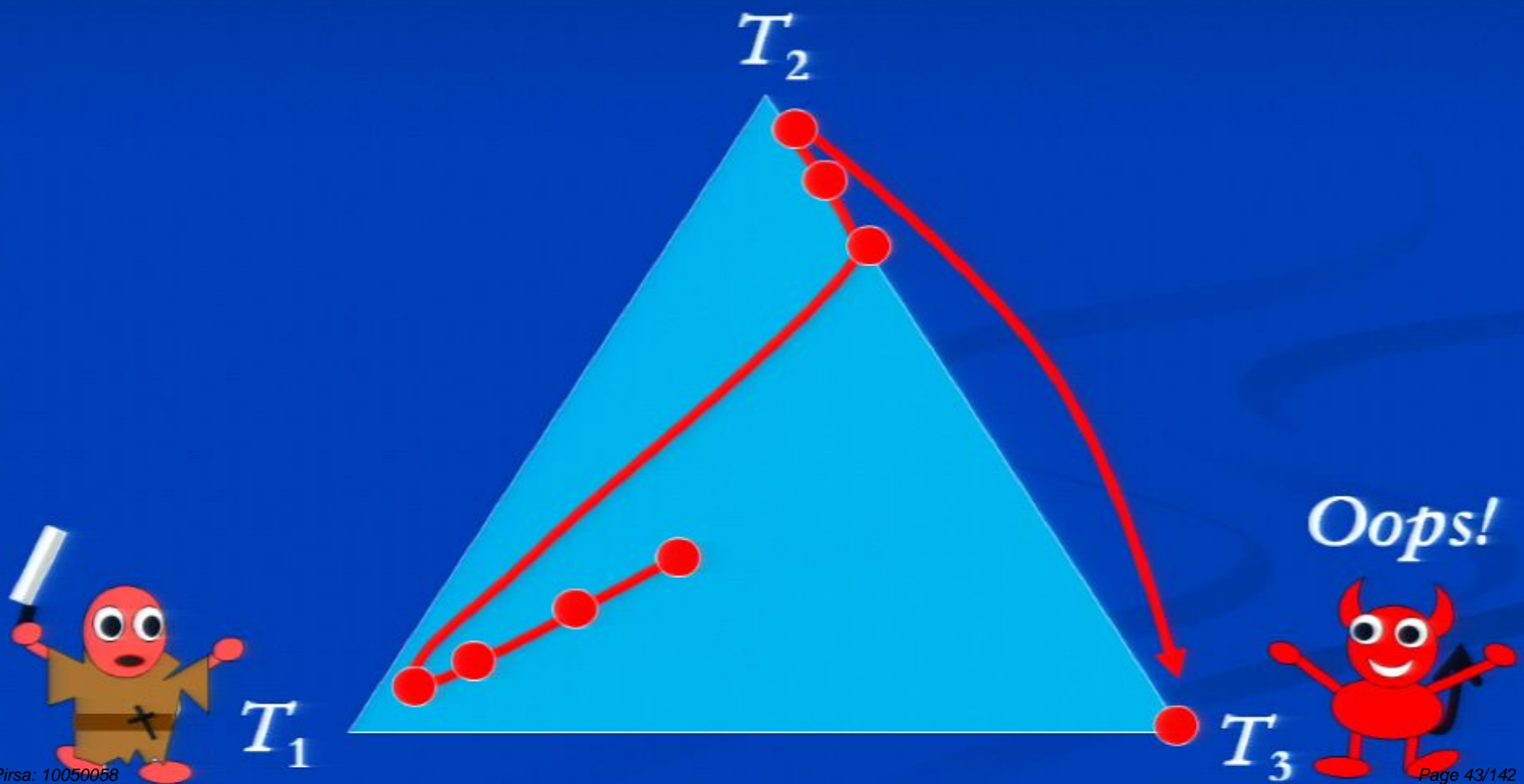
Ockham Bayesians Converge



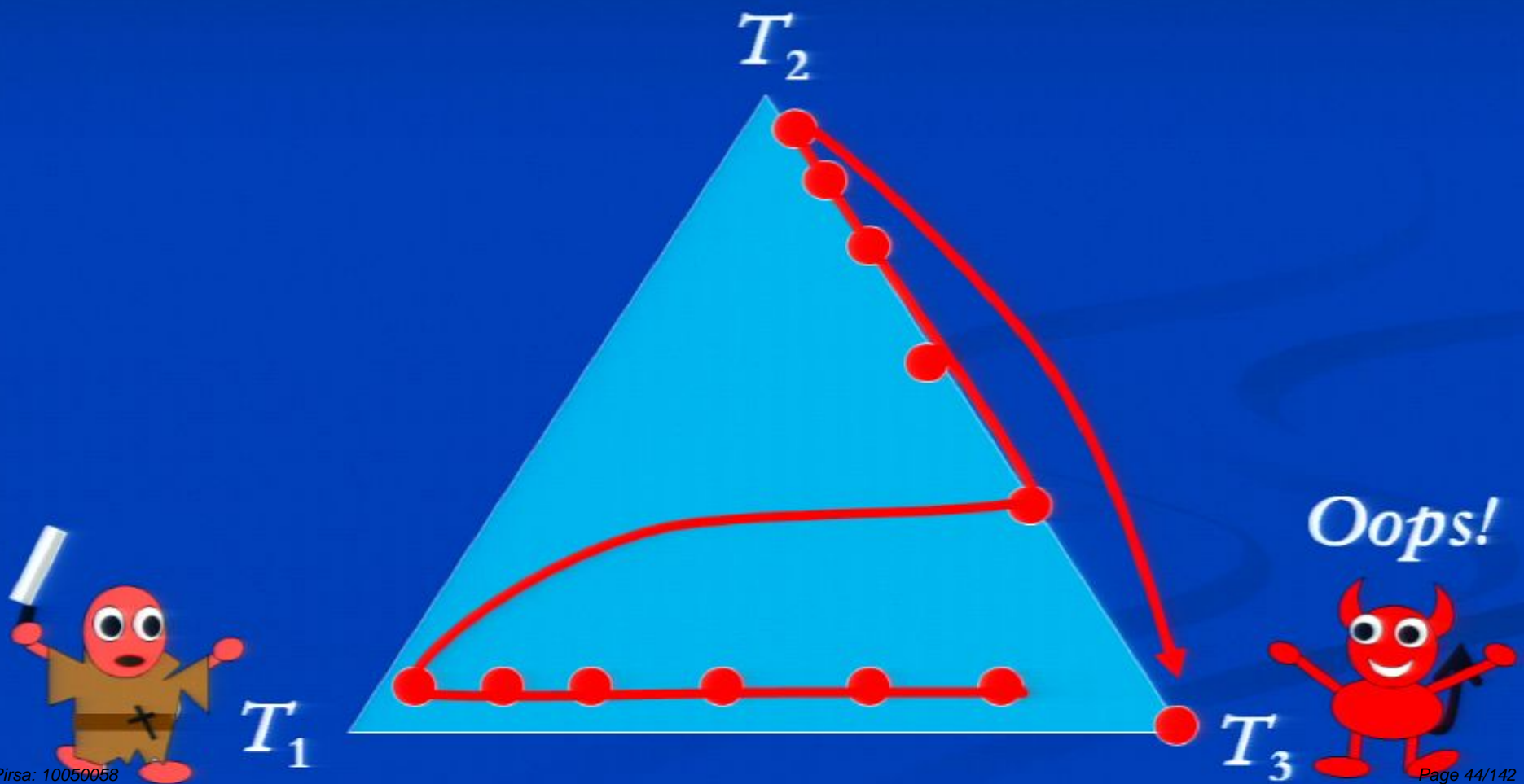
Ockham Bayesians Converge



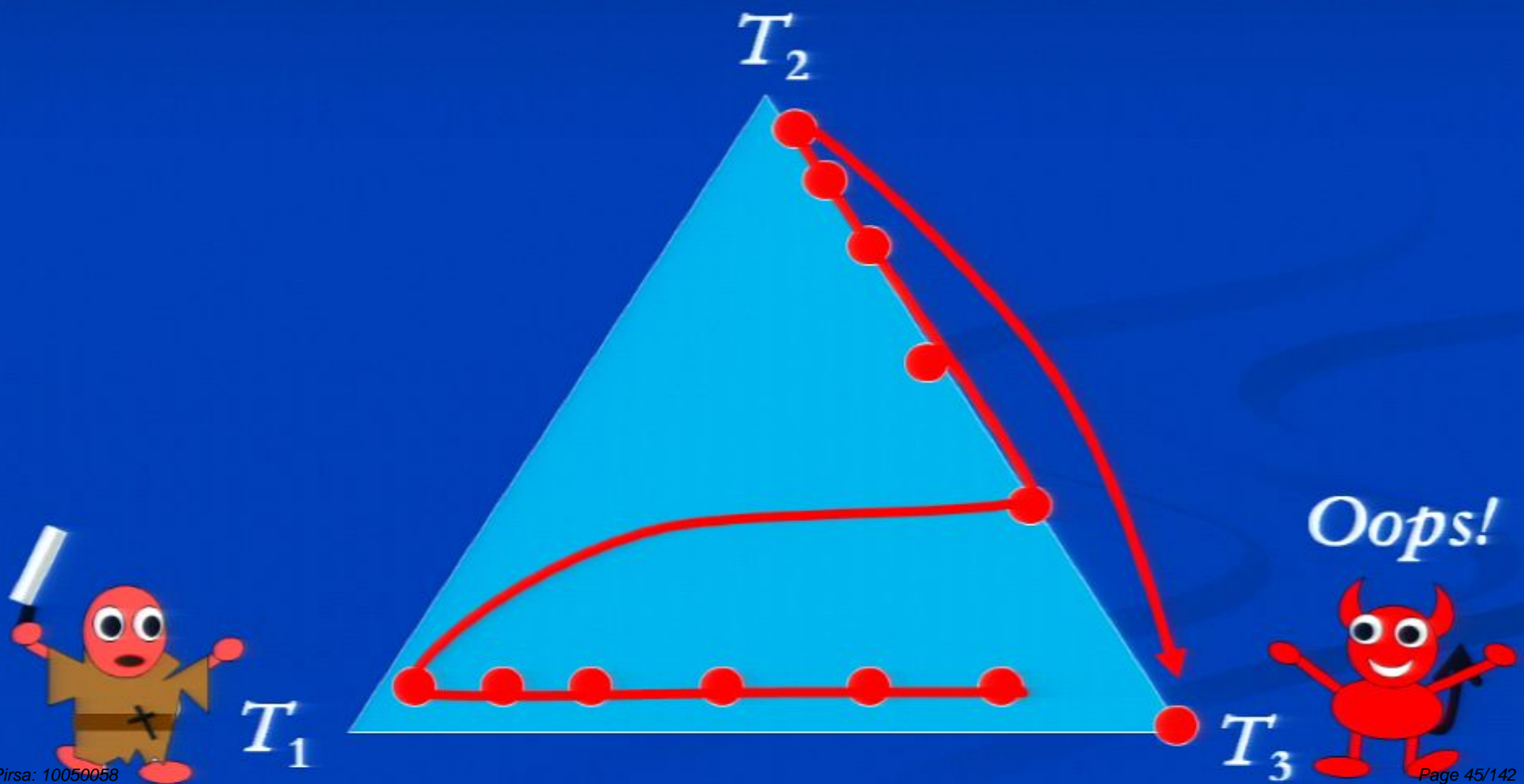
Ockham Bayesians Converge



But So Would Other Methods



But So Would Other Methods



Summary of Bayesian Approach

- **Prior-based** explanations of Ockham's razor are circular.
- **Convergence-based** explanations of Ockham's razor fail to single out Ockham's razor.

2. Risk Minimization

- Ockham's razor minimizes expected distance of empirical estimates from the true value.



Unconstrained Estimates

- are **Centered** on truth but **spread** around it.



Constrained Estimates

- Off-center but less spread.



Constrained Estimates

- Off-center but less spread
- Overall improvement in expected distance from truth...



Constrained Estimates

- Off-center but less spread.



Clamped aim



Unconstrained Estimates

- are **Centered** on truth but **spread** around it.



Constrained Estimates

- Off-center but less spread.



Clamped aim



Constrained Estimates

- Off-center but less spread
- Overall improvement in expected distance from truth...



Doesn't Find True **Theory**

- **False** theories that aren't too false theory typically predict more accurately than the true theory.



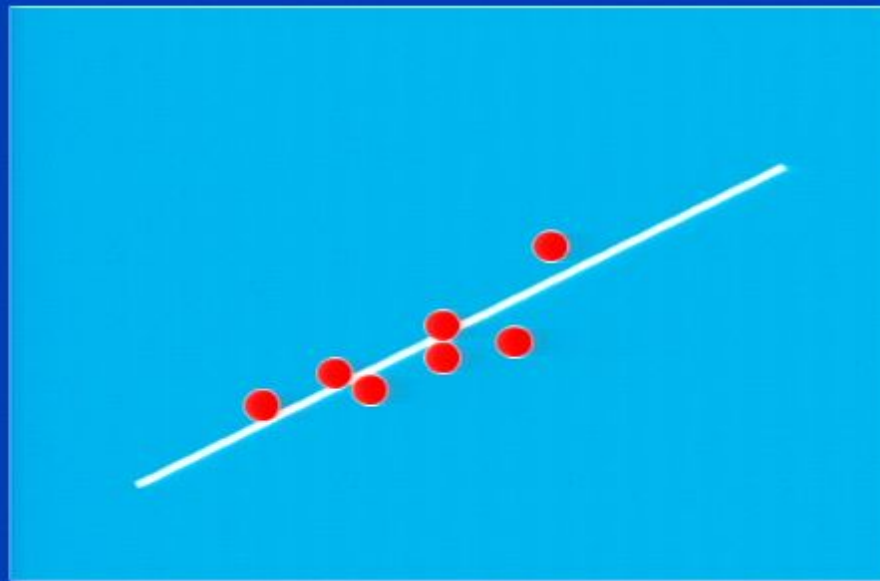
Clamped aim



Four eyes!

When Theory Doesn't Matter

- Predicting lung cancer from ash trays.



Doesn't Find True **Theory**

- **False** theories that aren't too false theory typically predict more accurately than the true theory.



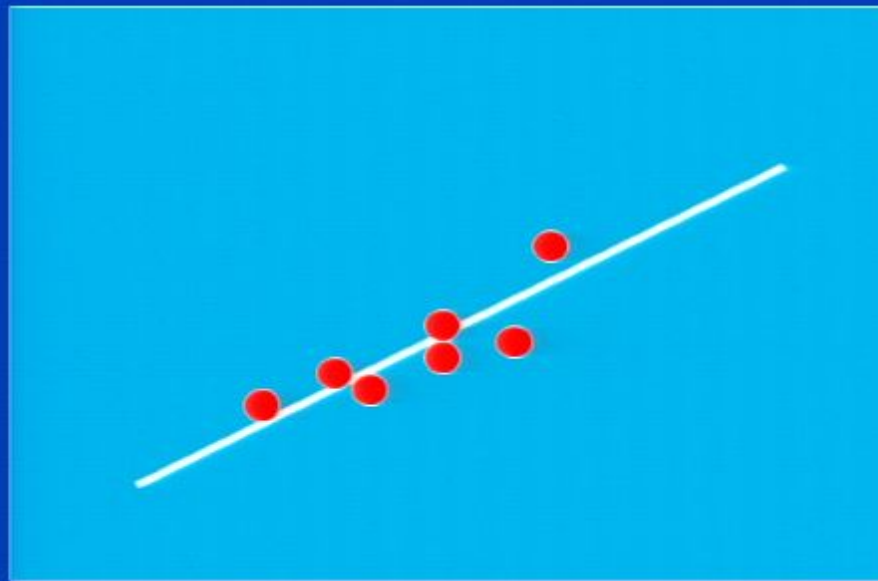
Clamped aim



Four eyes!

When Theory Doesn't Matter

- Predicting lung cancer from ash trays.



Doesn't Find True **Theory**

- **False** theories that aren't too false theory typically predict more accurately than the true theory.



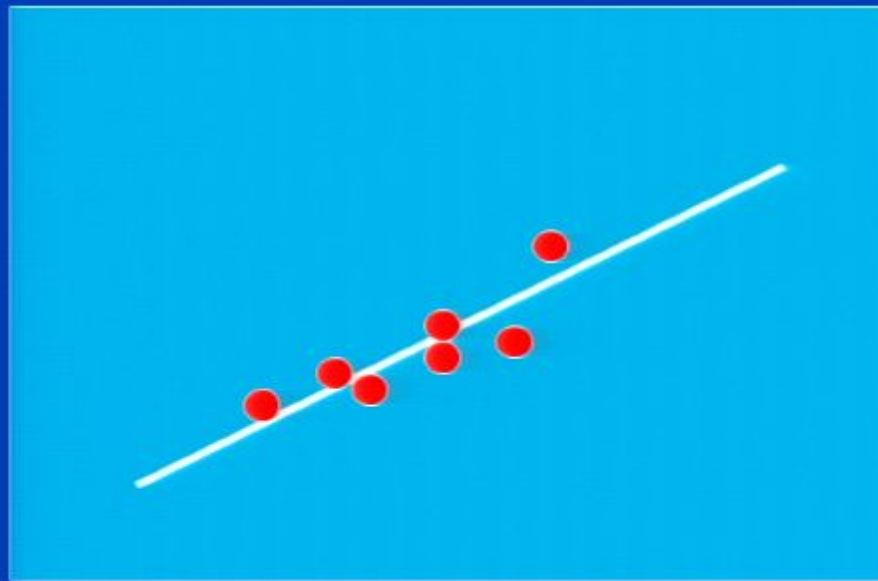
Clamped aim



Four eyes!

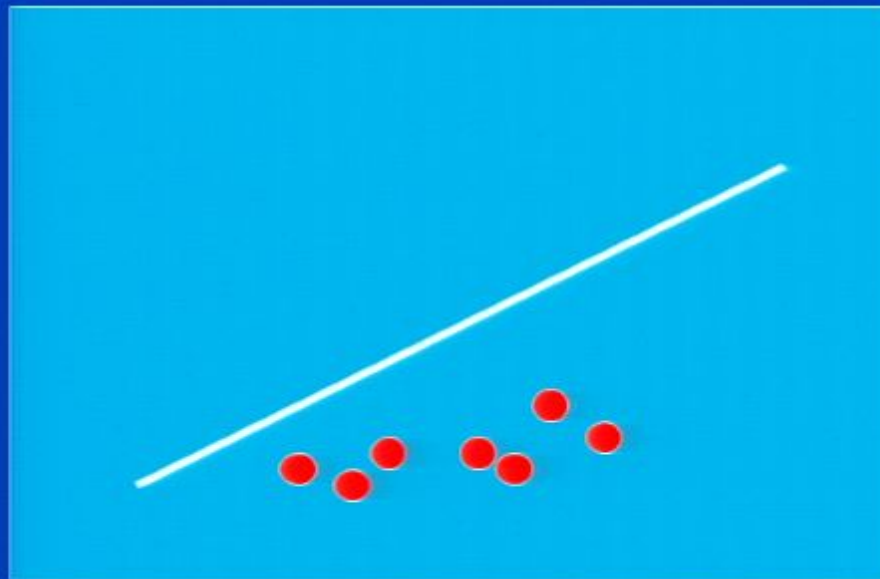
When Theory Doesn't Matter

- Predicting lung cancer from ash trays.



When Theory Does Matter

- Predicting effectiveness of a ban on ash trays.



Great News

Correlation **does** imply causation if there are multiple variables, some of which are common effects. [Pearl, Spirtes, Glymour and Scheines, etc.]



Core assumptions

Joint distribution p is **causally compatible** with causal network G iff:

Causal Markov Condition: each variable X is independent of its non-effects given its immediate causes.

Faithfulness Condition: no other conditional independence relations hold in p .

Tell-tale Correlations



*Given C ,
 $F1$ gives no further
info about $F2$*
(Markov)

Pirsa: 10050058



*Given F ,
 H gives
some info about C*
(Faithfulness)

Core assumptions

Joint distribution p is **causally compatible** with causal network G iff:

Causal Markov Condition: each variable X is independent of its non-effects given its immediate causes.

Faithfulness Condition: no other conditional independence relations hold in p .

Tell-tale Correlations



*Given C ,
 $F1$ gives no further
info about $F2$*
(Markov)

Pirsa: 10050058



*Given F ,
 H gives
some info about C*
(Faithfulness)

Core assumptions

Joint distribution p is **causally compatible** with causal network G iff:

Causal Markov Condition: each variable X is independent of its non-effects given its immediate causes.

Faithfulness Condition: no other conditional independence relations hold in p .

Tell-tale Correlations



*Given C ,
 $F1$ gives no further
info about $F2$*
(Markov)

Pirsa: 10050058



*Given F ,
 H gives
some info about C*
(Faithfulness)

Standard Applications

- **Linear Causal Case:** each variable X is a linear function of its parents and a normally distributed hidden variable called an “error term”. The error terms are mutually independent.
- **Discrete Multinomial Case:** each variable X takes on a finite range of values.

The Catch



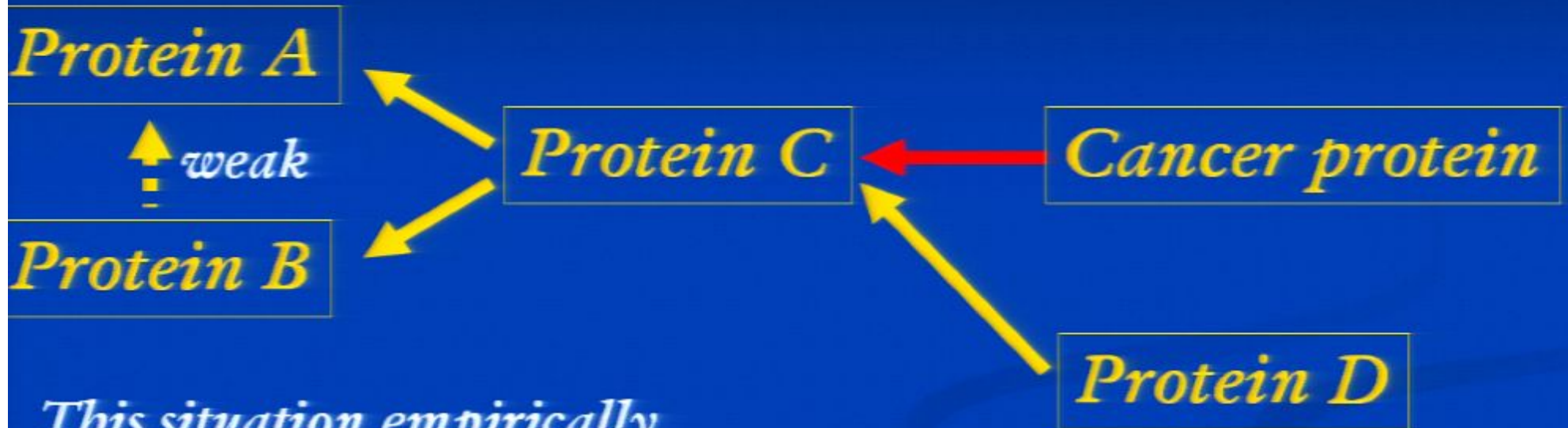
English Breakfast?



It's full of protein C!



As the Sample Increases...



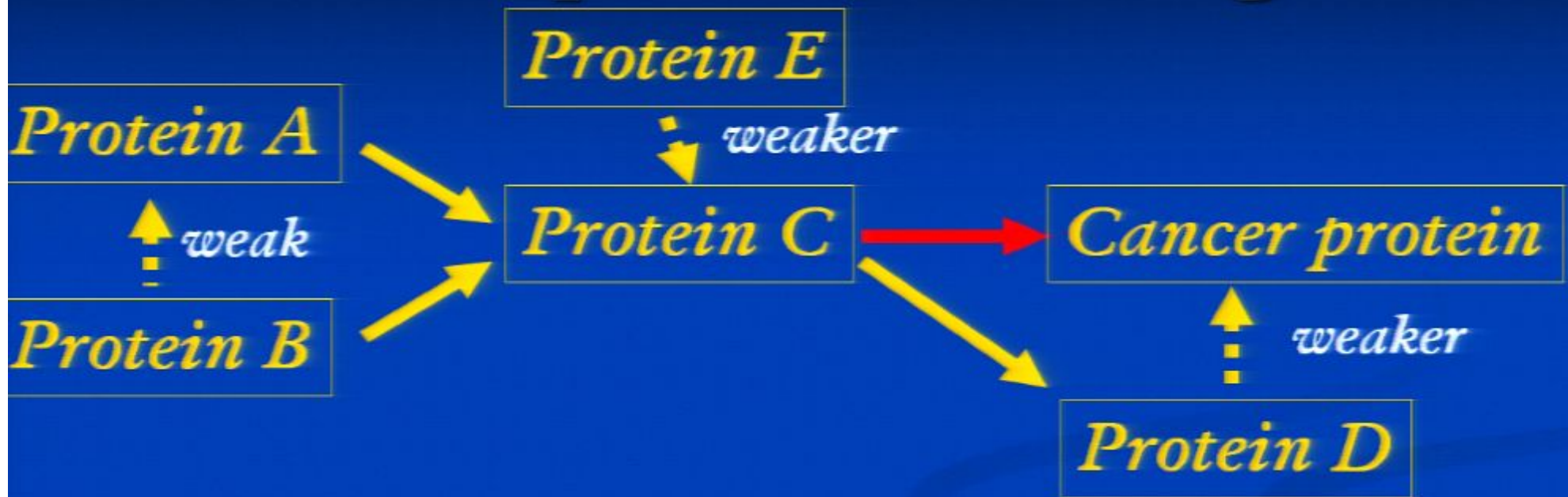
*This situation empirically
approximates the last one.
So who cares?*



I do! Bon appetit!



As the Sample Increases Again...



Wasn't that last approximation to the truth good enough?



Causal Flipping Theorem

For each convergent M and standard (G, p) , there exists standard (G', p') such that:

- p' is indistinguishable from p at the current sample size and
- consistent M flips the orientation of causal arrow $X \rightarrow Y$ in (G', p') at least n times as sample size increases.



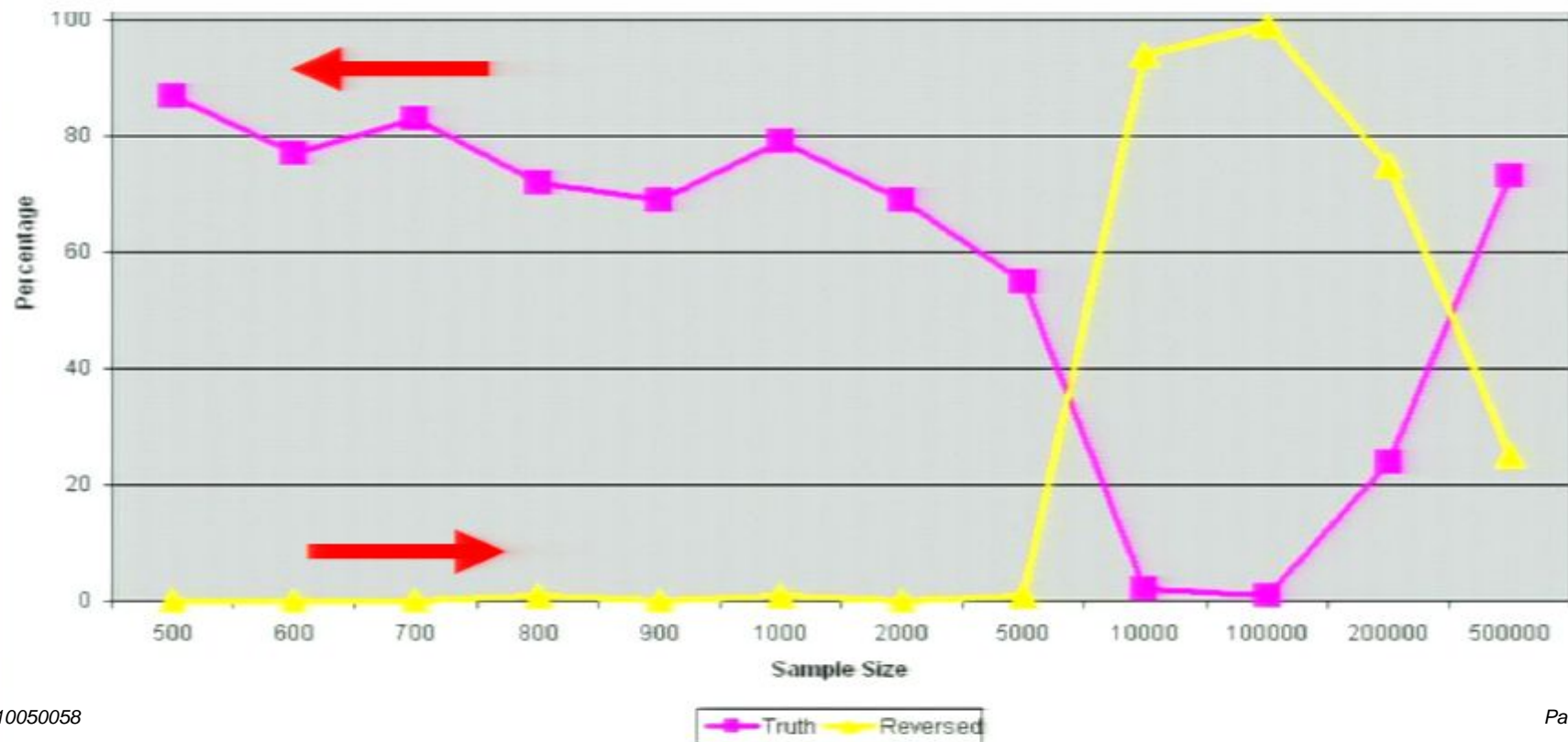
oops
I meant ← oops
I meant → oops
I meant ←



Simulation Using CPC Algorithm

CPC Algorithm

Proportion of outputs out of 1000 trials at each sample size.



Causal Flipping Theorem

For each convergent M and standard (G, p) , there exists standard (G', p') such that:

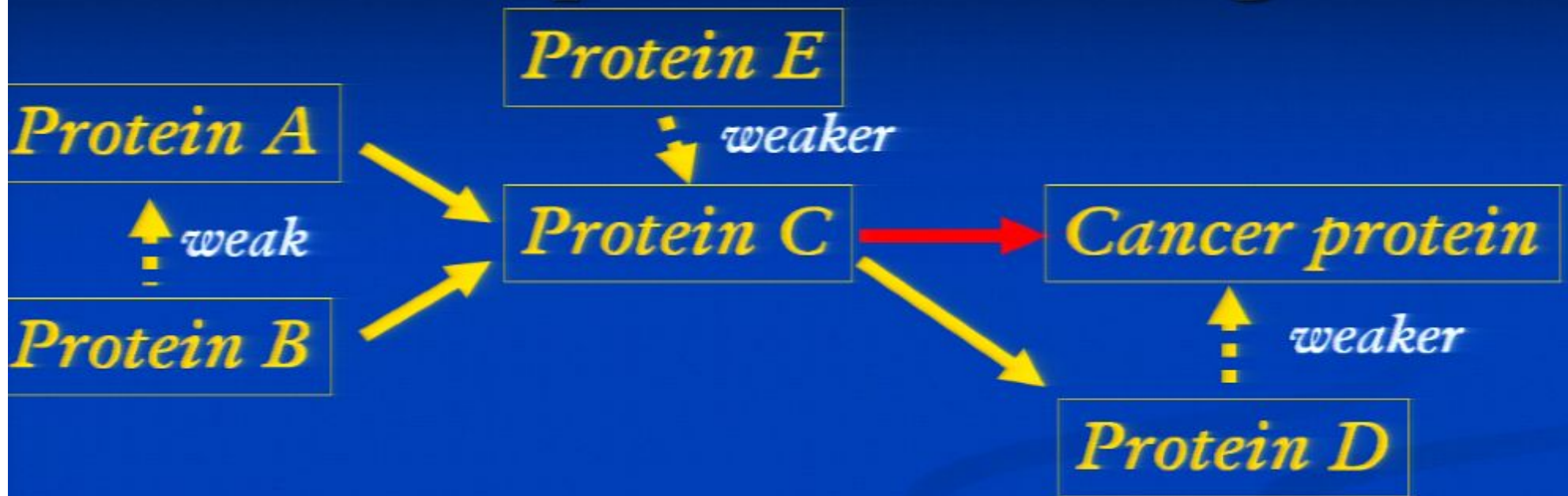
- p' is indistinguishable from p at the current sample size and
- consistent M flips the orientation of causal arrow $X \rightarrow Y$ in (G', p') at least n times as sample size increases.



I meant  *oops*
I meant  *oops*
I meant  *oops*
I meant 



As the Sample Increases Again...



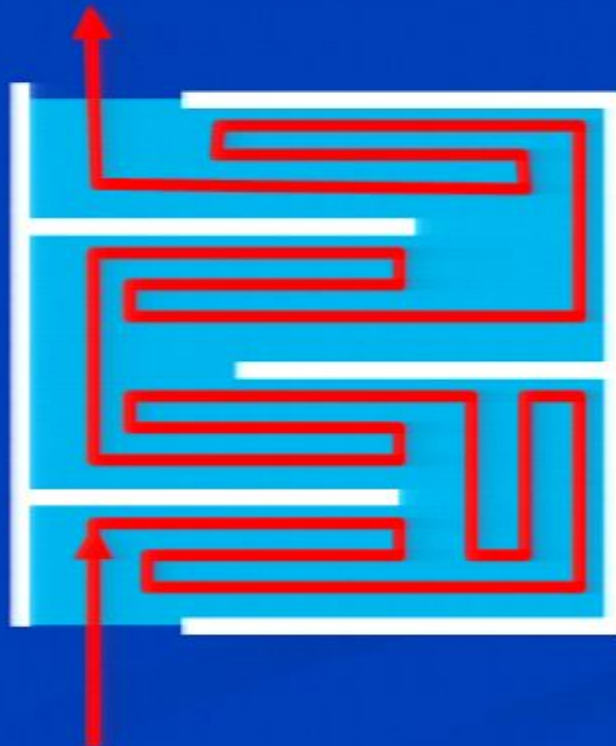
Wasn't that last approximation to the truth good enough?



Pursuit of Truth

■ Long-run Convergence

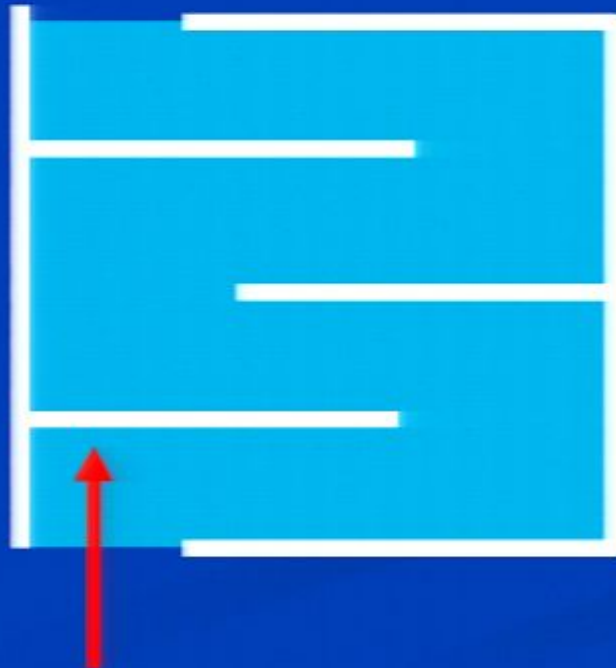
- *Too weak to single out Ockham's razor*



Pursuit of Truth

■ Short-run Tracking

- *Too strong to be feasible when theory matters.*



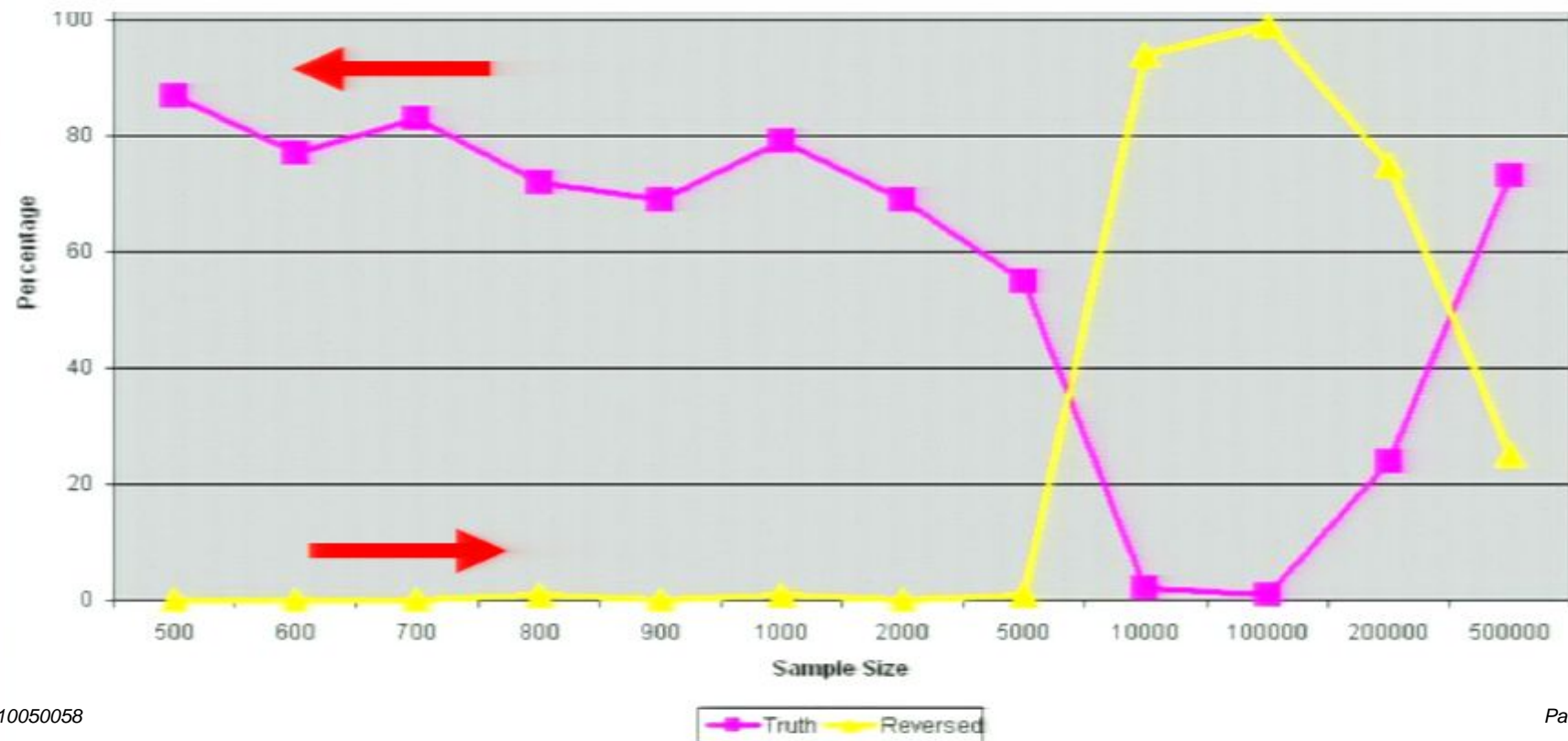
Zen Response

- Many explanations have been offered to make sense of the here-today-gone-tomorrow nature of medical wisdom — what we are advised **with confidence** one year is **reversed** the next — but the simplest one is that it is **the natural rhythm of science**.
- (*Do We Really Know What Makes us Healthy*, NY Times Magazine, Sept. 16, 2007).

Simulation Using CPC Algorithm

CPC Algorithm

Proportion of outputs out of 1000 trials at each sample size.



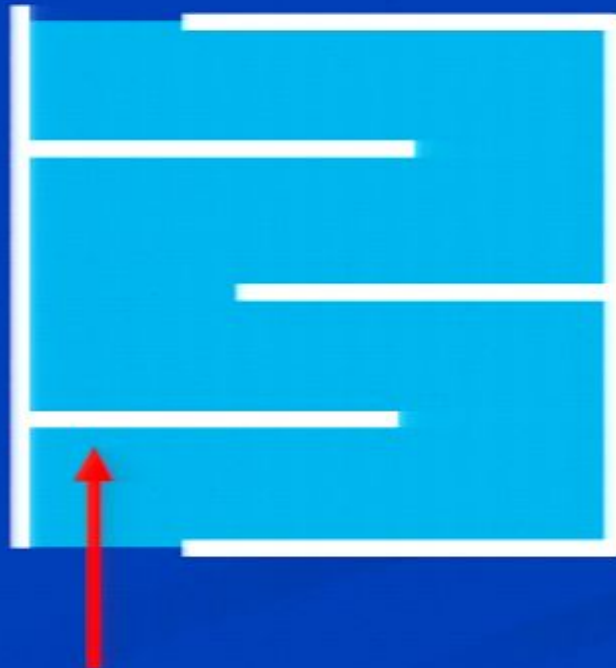
Zen Response

- Many explanations have been offered to make sense of the here-today-gone-tomorrow nature of medical wisdom — what we are advised **with confidence** one year is **reversed** the next — but the simplest one is that it is **the natural rhythm of science**.
- (*Do We Really Know What Makes us Healthy*, NY Times Magazine, Sept. 16, 2007).

Pursuit of Truth

■ Short-run Tracking

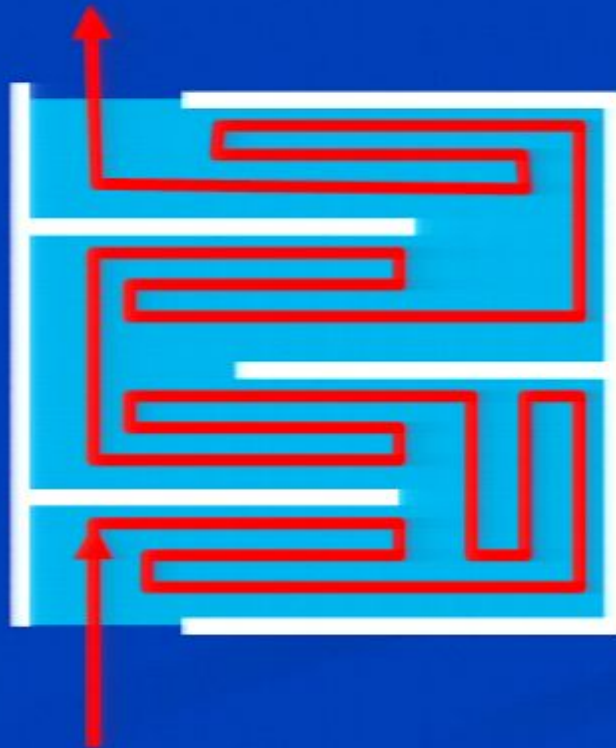
- *Too strong to be feasible when theory matters.*



Pursuit of Truth

■ Long-run Convergence

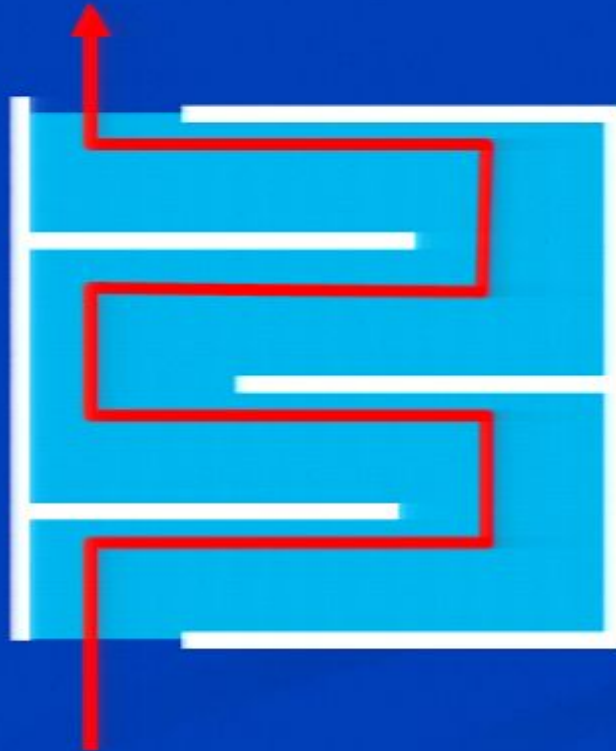
- *Too weak to single out Ockham's razor*



The Middle Way

■ Straightest Convergence

■ *Just right?*



Ancient Roots

"Living in the midst of ignorance and considering themselves intelligent and enlightened, the senseless people go round and round, following crooked courses, just like the blind led by the blind." *Katha Upanishad*, I. ii. 5, c. 600 BCE.

II. Navigation by Broken Compass

simple



Asking for Directions

Where's ...



Asking for Directions

Turn around. The freeway ramp is on the left.



Asking for Directions



Best Route



Best Route to Any Goal



Disregarding Advice is Bad



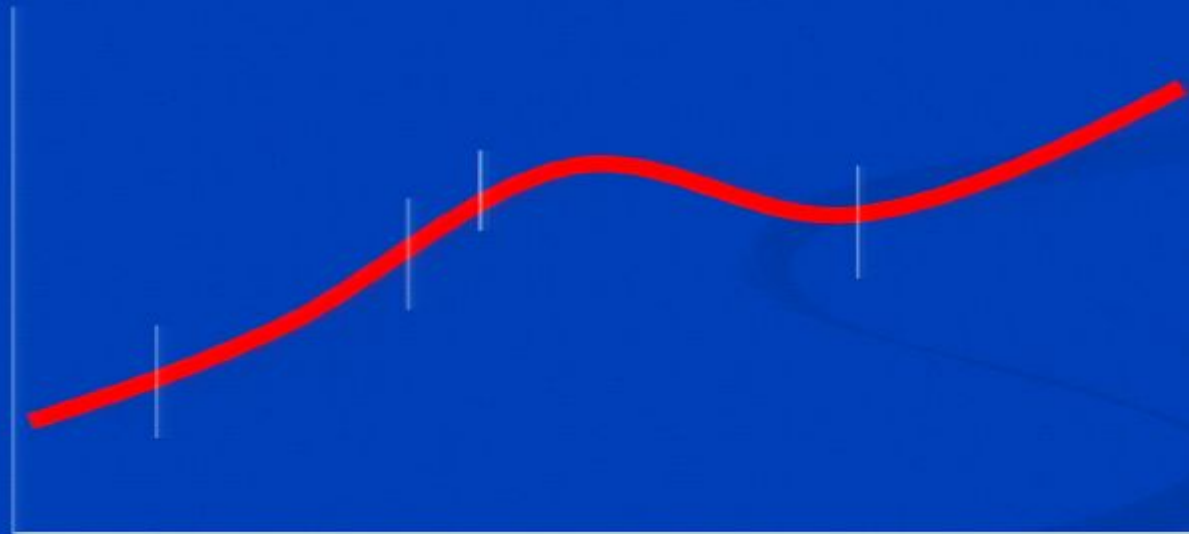
Helpful A Priori Advice



...so **fixed** advice can help you reach a **hidden** goal
without circles, evasions, or magic.

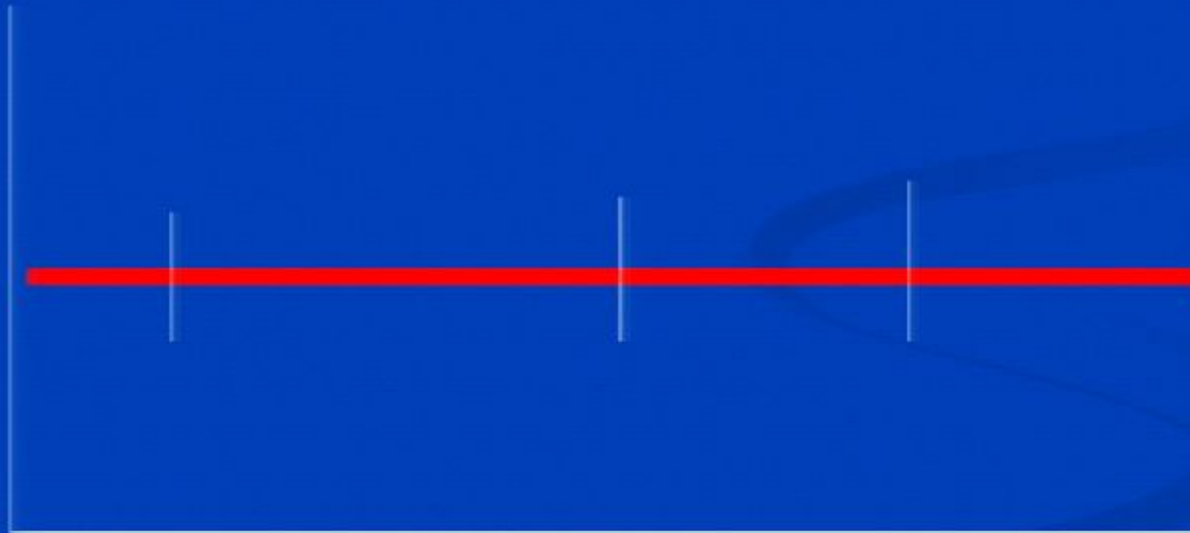
Polynomial Laws

- Data = open intervals around Y at rational values of X .



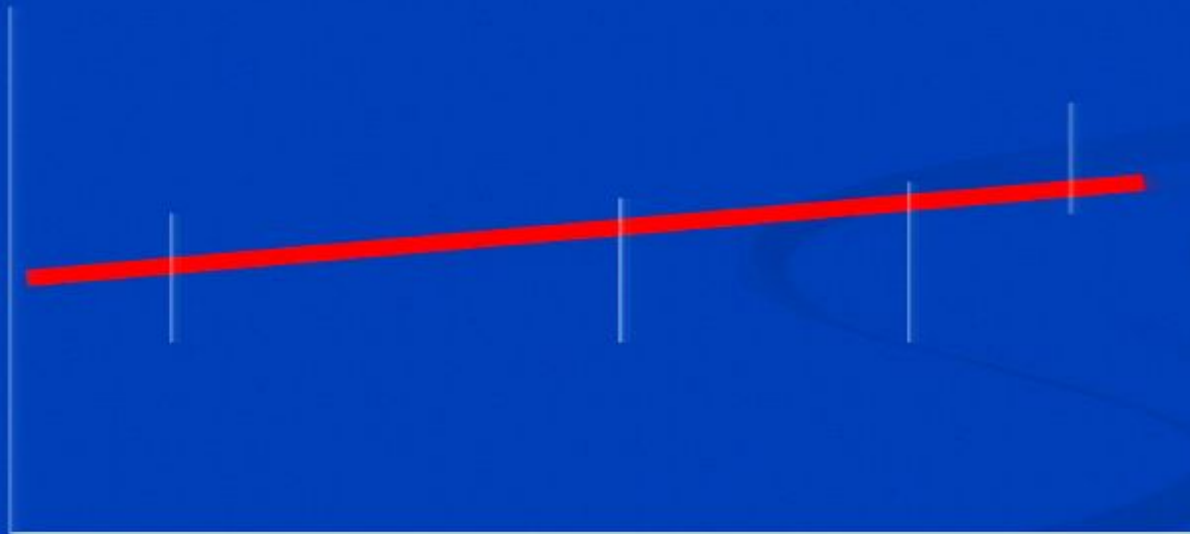
Polynomial Laws

■ No effects:



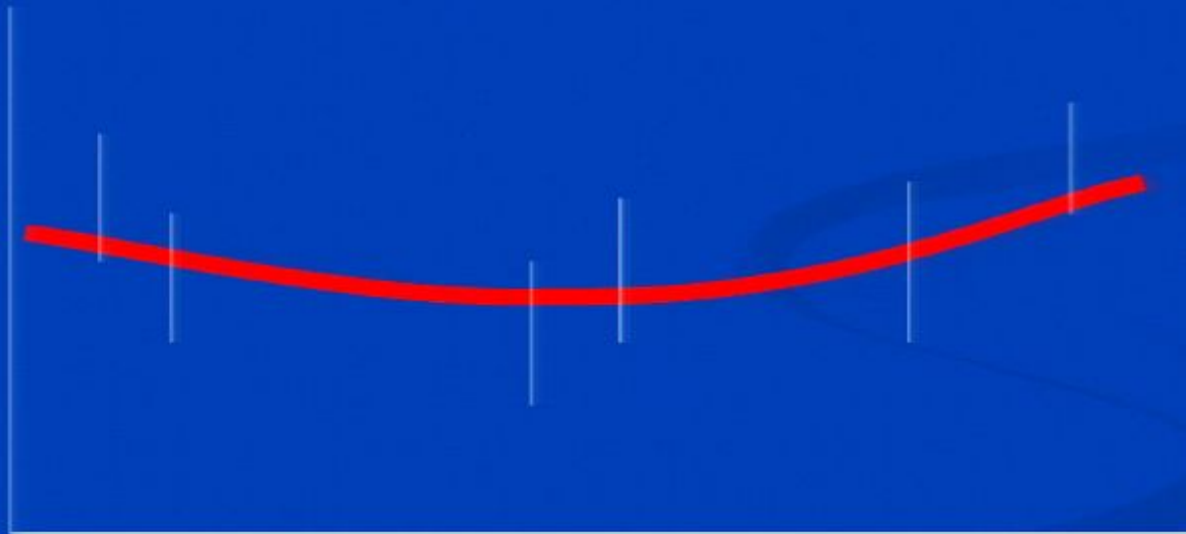
Polynomial Laws

■ First-order effect:



Polynomial Laws

■ Second-order effect:



In Step with the Demon

There yet?

Maybe.



Constant



Linear



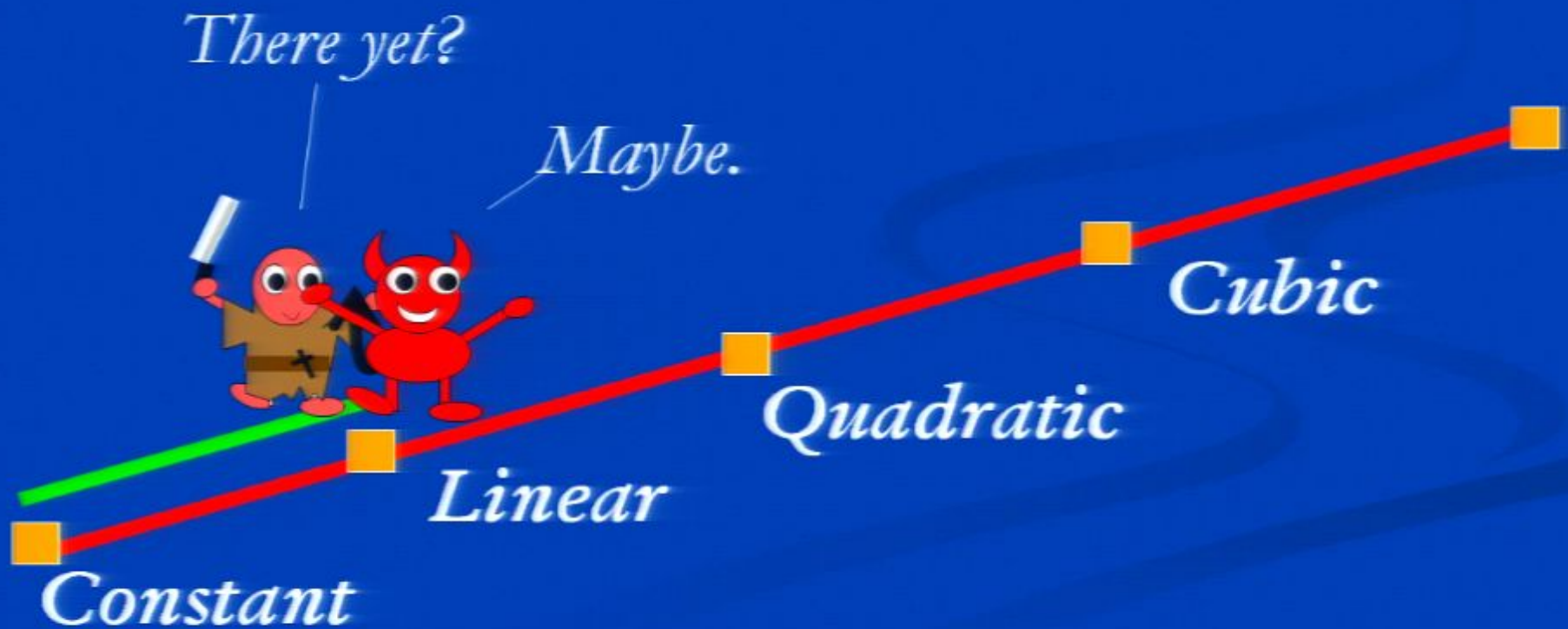
Quadratic



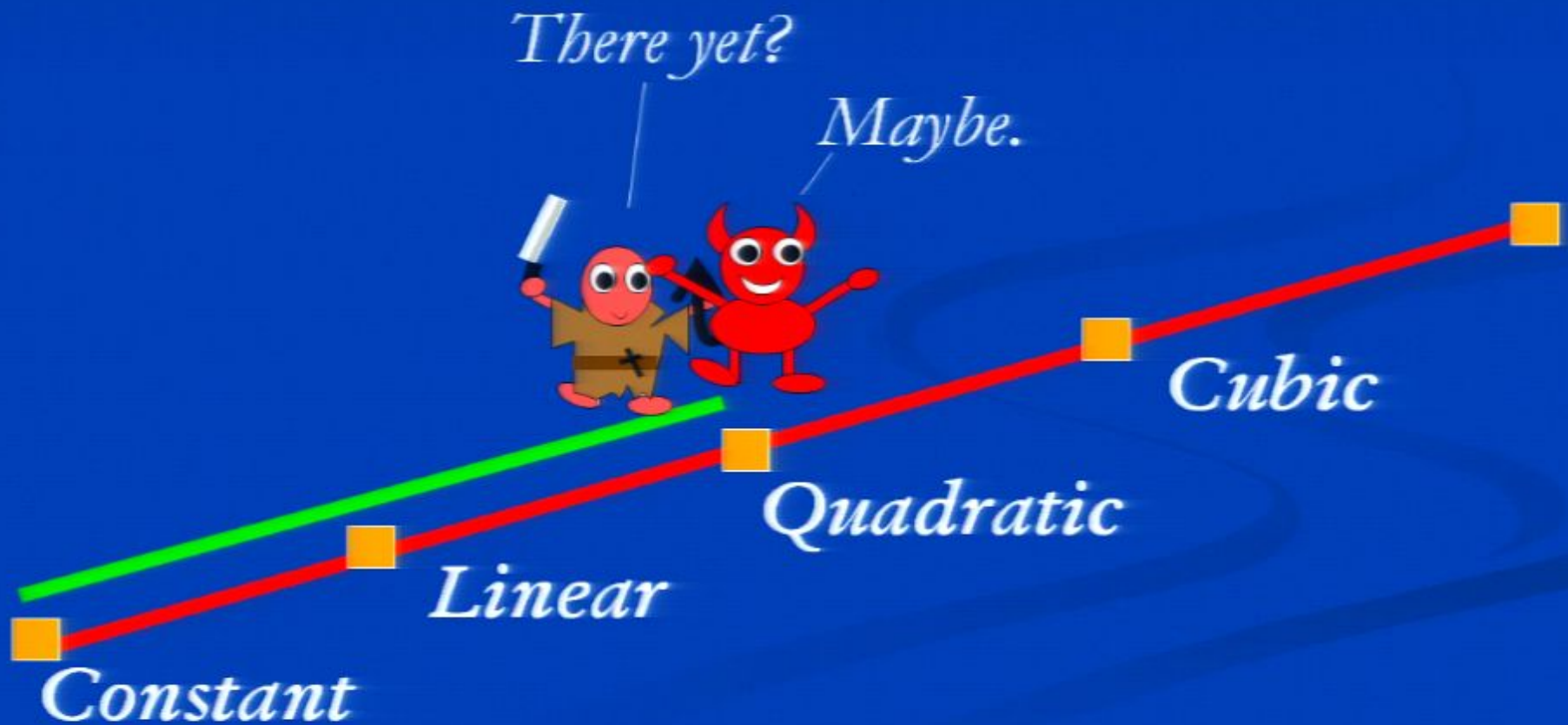
Cubic



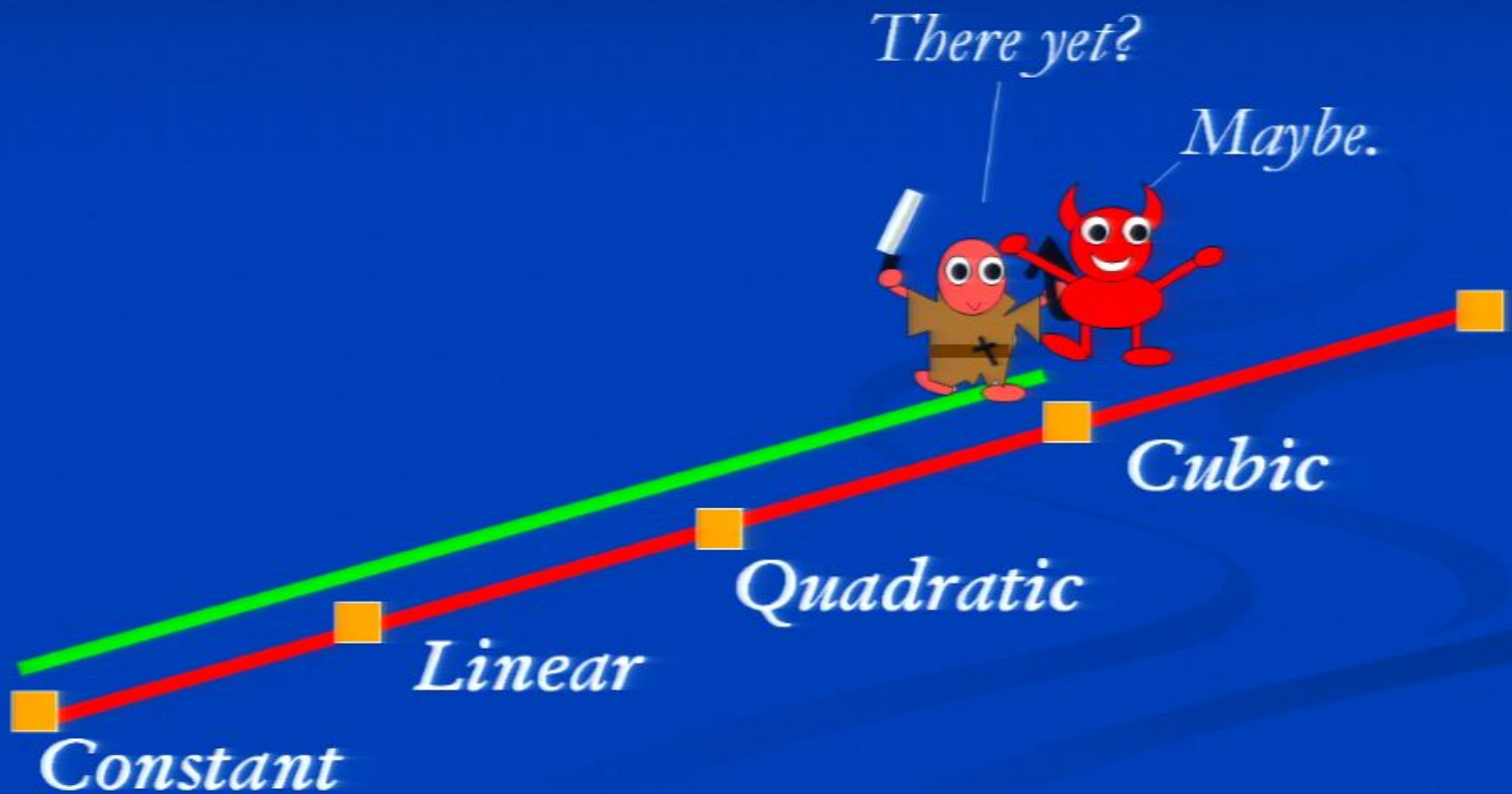
In Step with the Demon



In Step with the Demon



In Step with the Demon

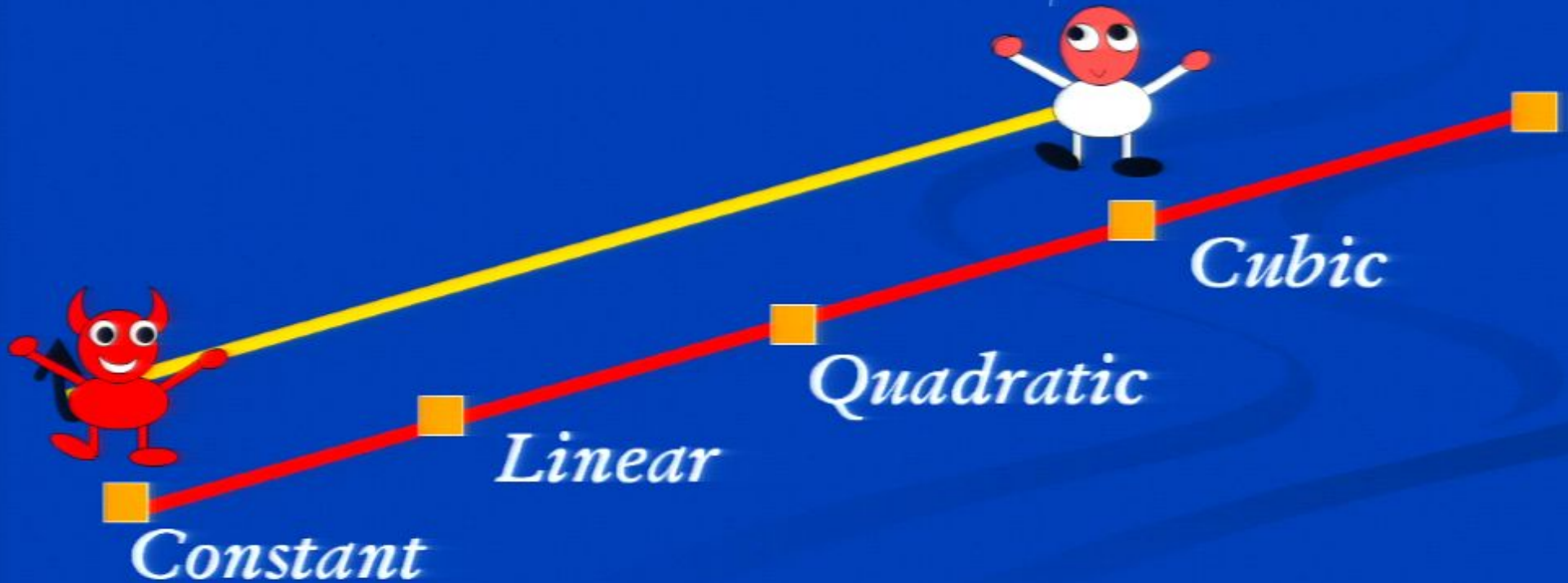


Passing the Demon



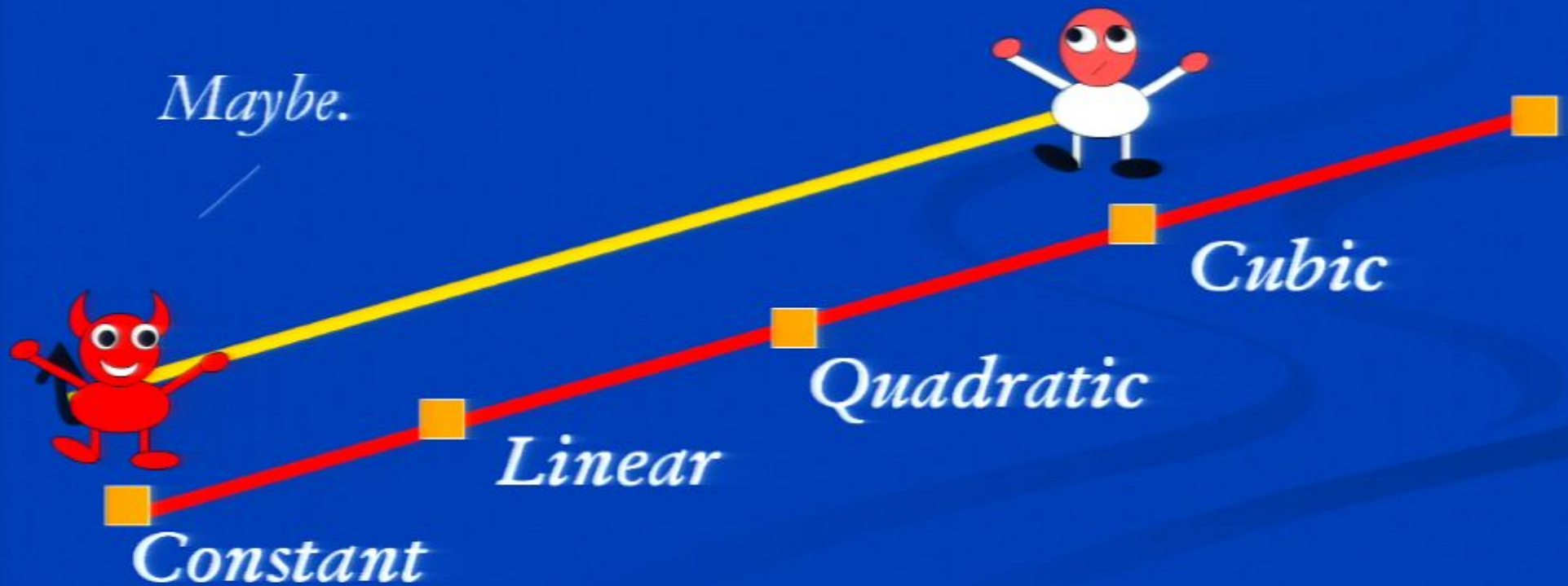
Passing the Demon

I know you're coming!



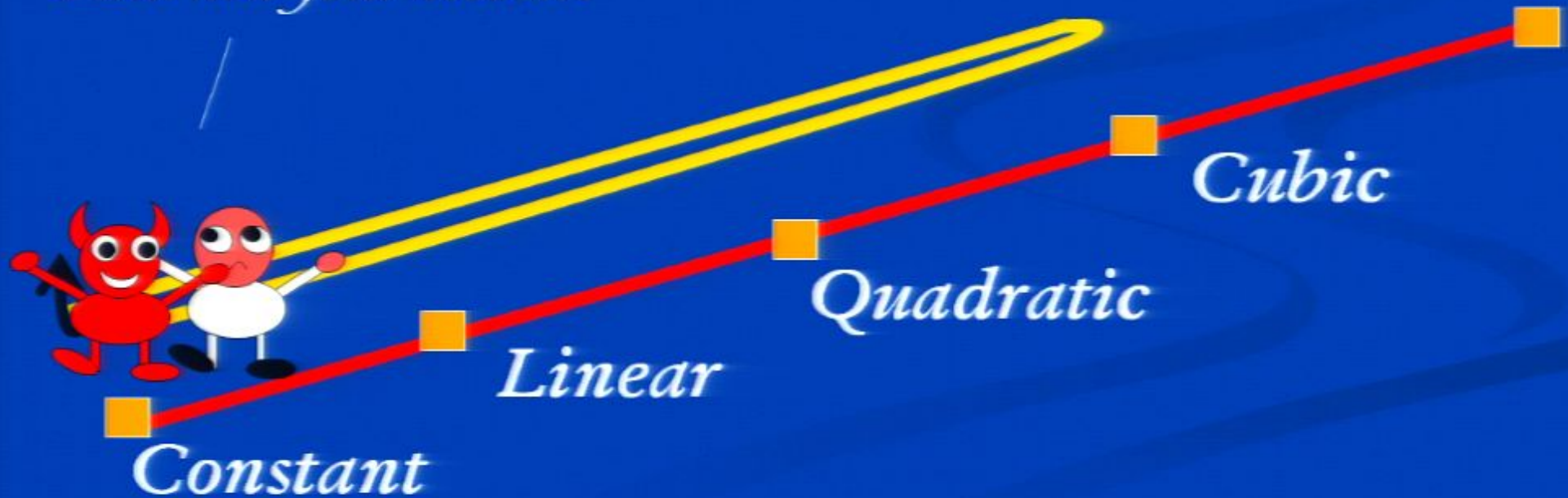
Passing the Demon

Maybe.



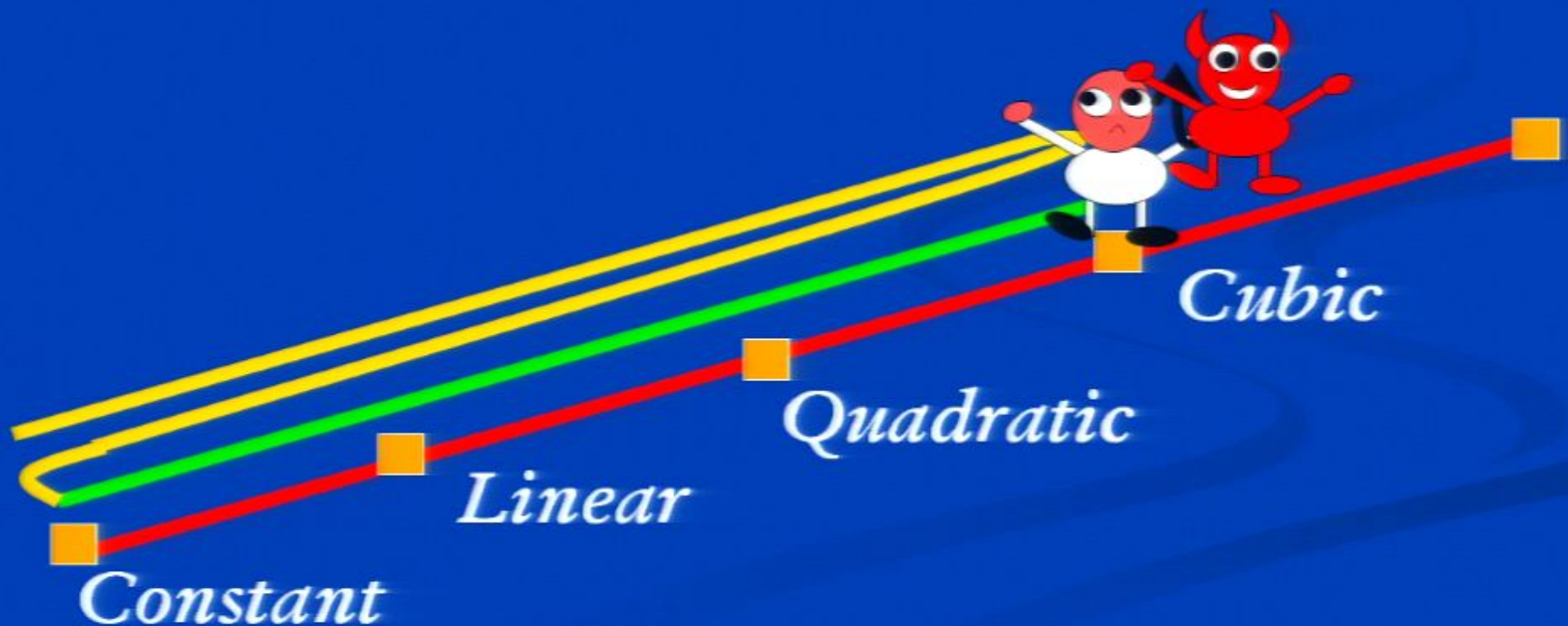
Passing the Demon

You're back!
Learned your lesson?

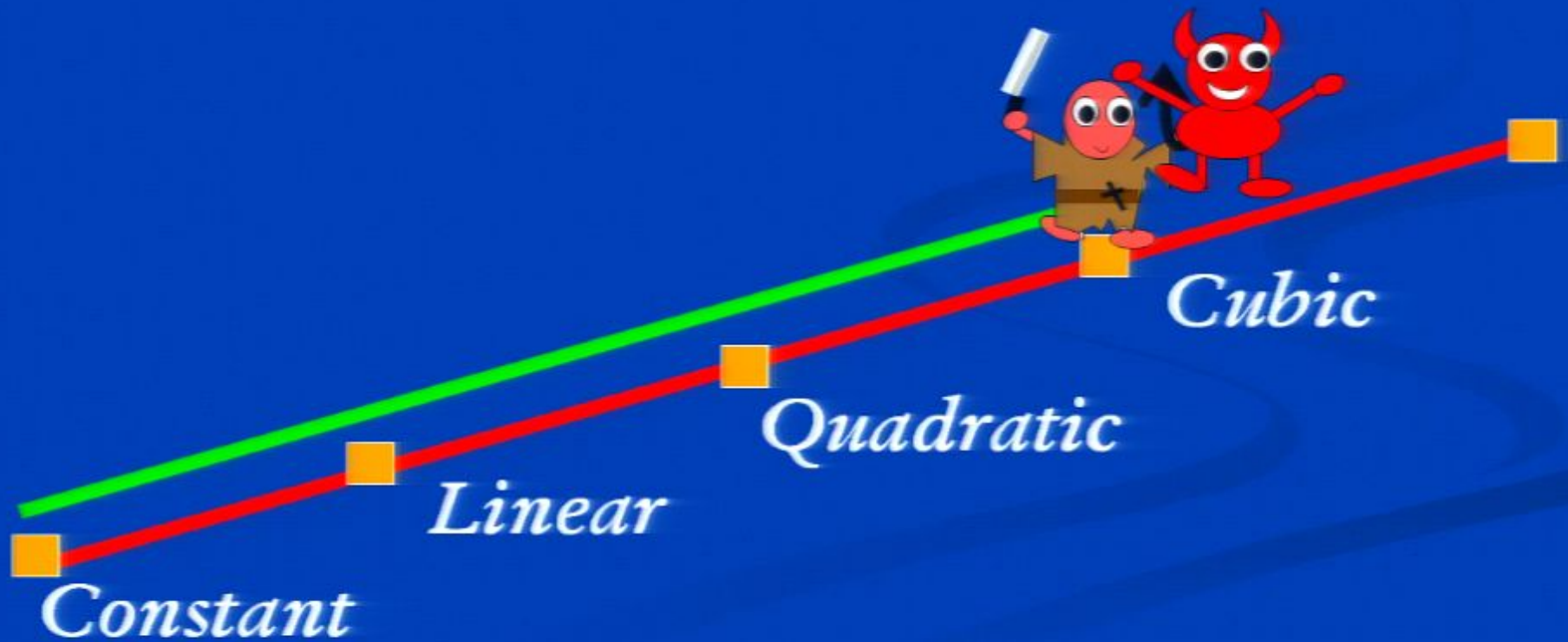


Ockham Violator's Path

*See, you shouldn't run ahead
Even if you are right!*



Ockham Path



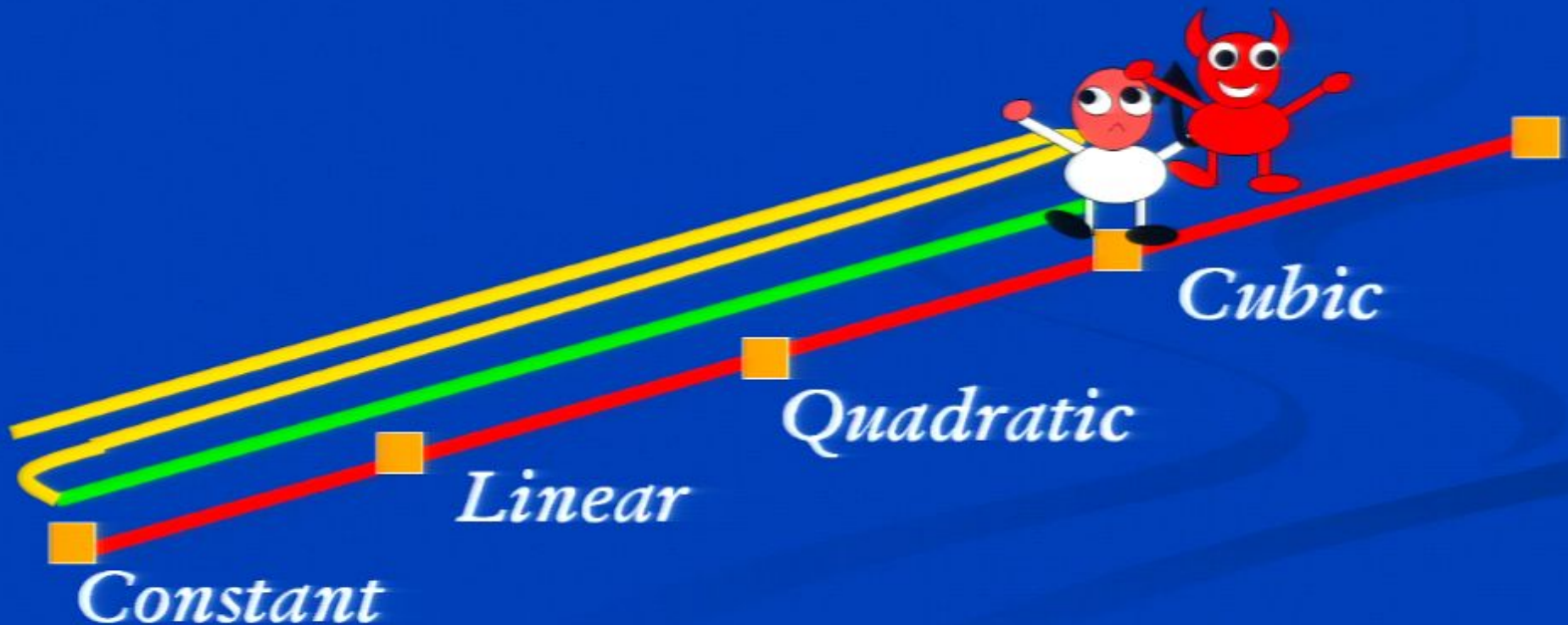
Empirical Problems

- Set K of infinite **input** sequences.
- Partition of K into alternative **theories**.

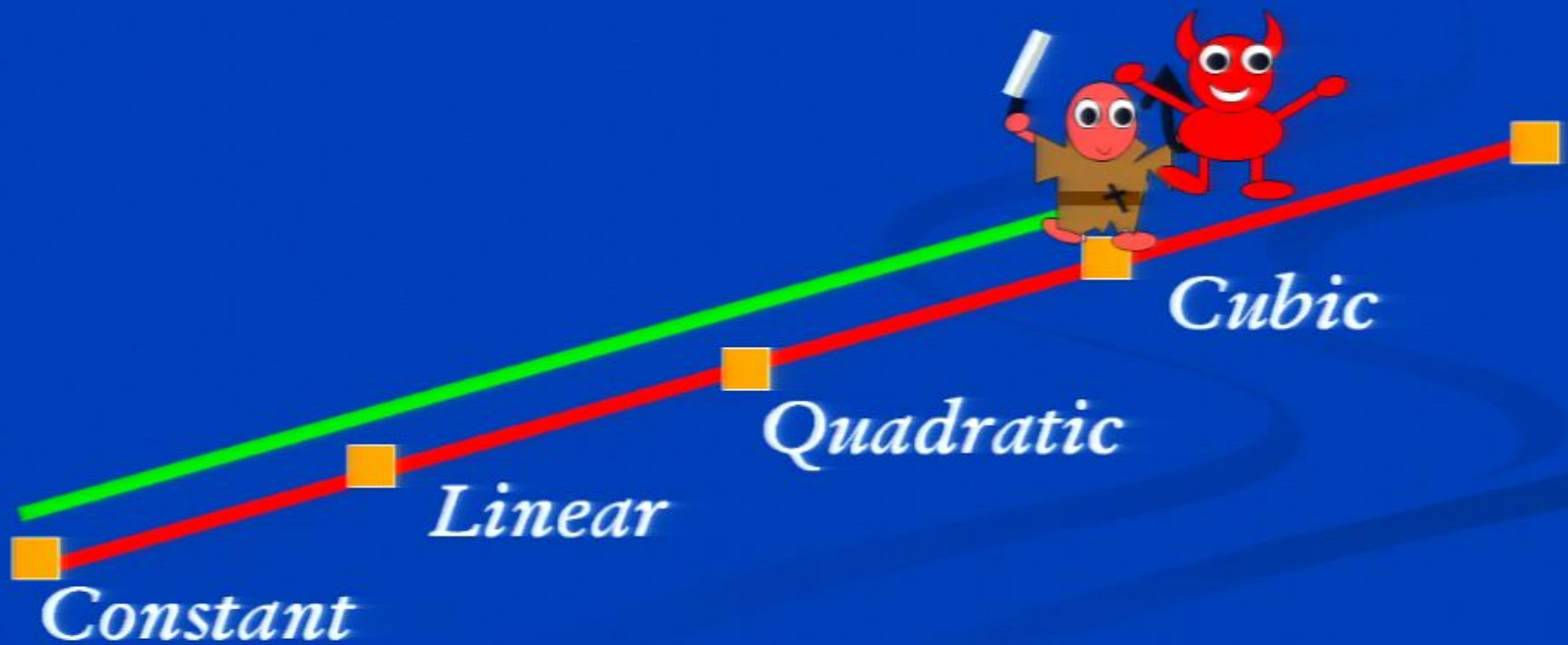


Ockham Violator's Path

*See, you shouldn't run ahead
Even if you are right!*



Ockham Path



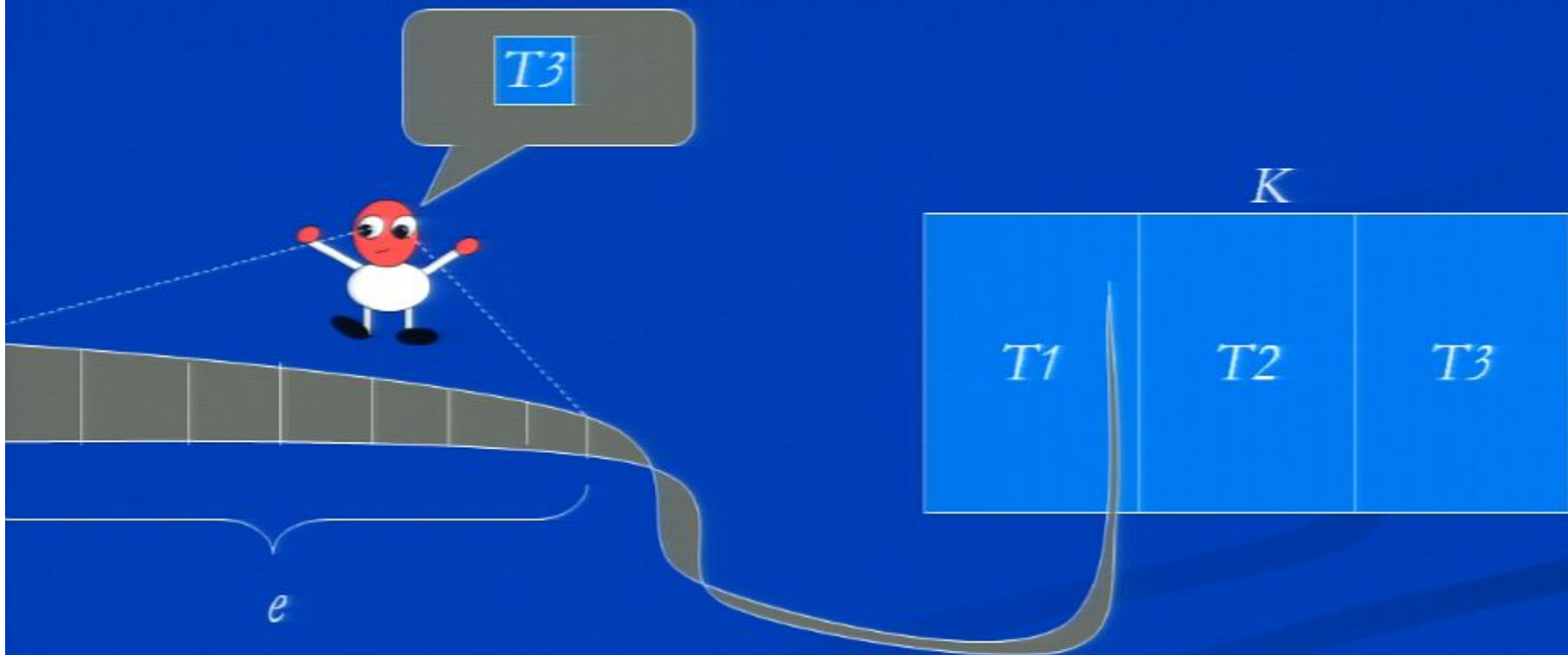
Empirical Problems

- Set K of infinite **input** sequences.
- Partition of K into alternative **theories**.



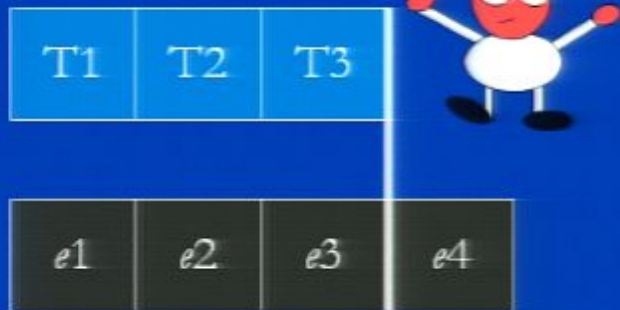
Empirical Methods

- Map finite input sequences to theories or to “?”.



Method Choice

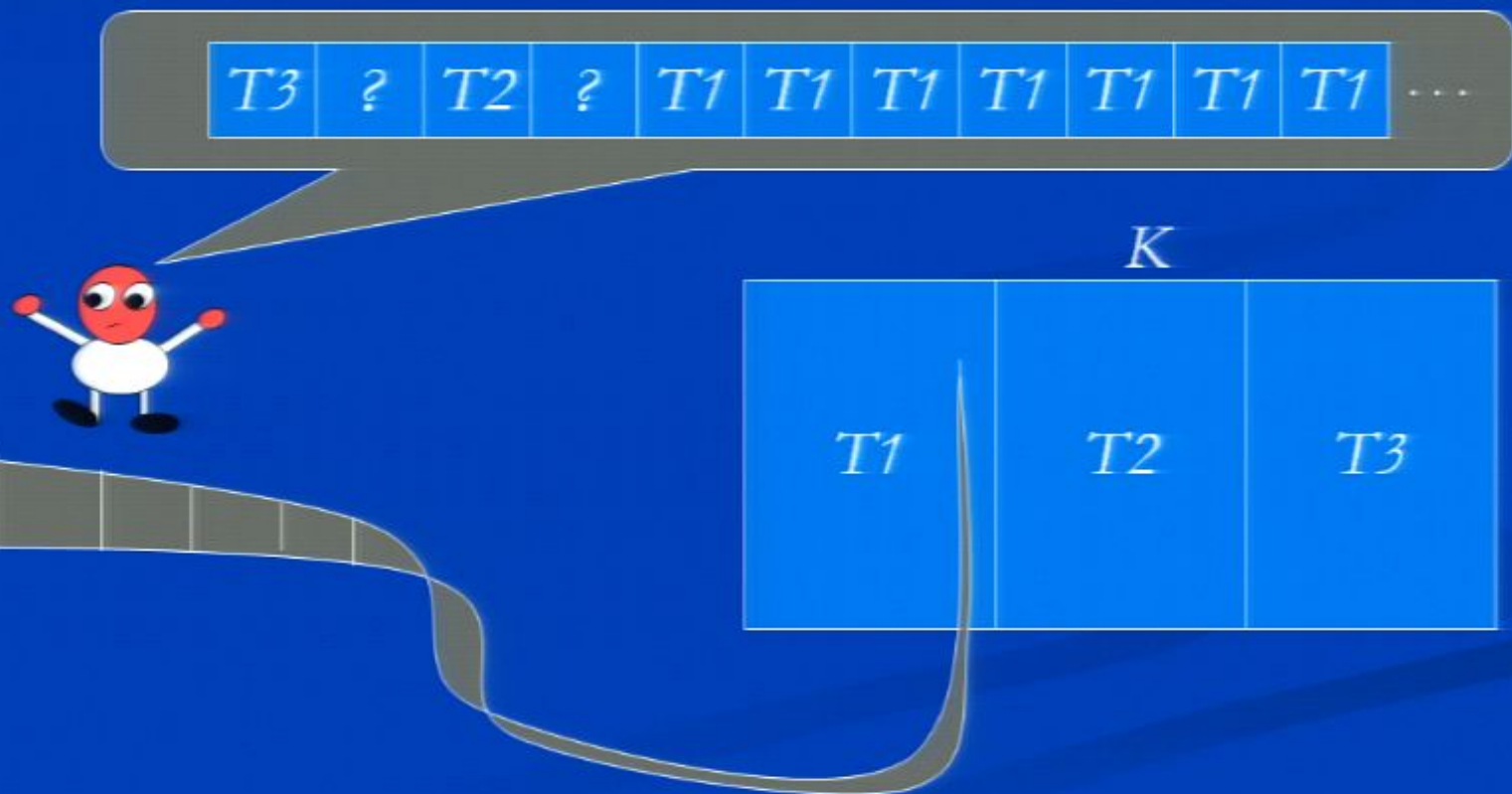
Output history



Input history

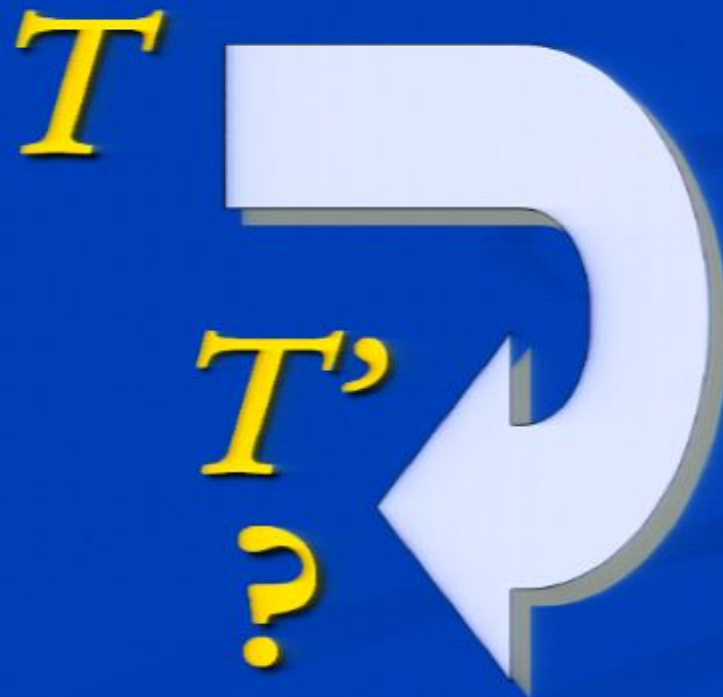
At each stage, scientist can choose a new method (agreeing with past theory choices).

Aim: Converge to the Truth

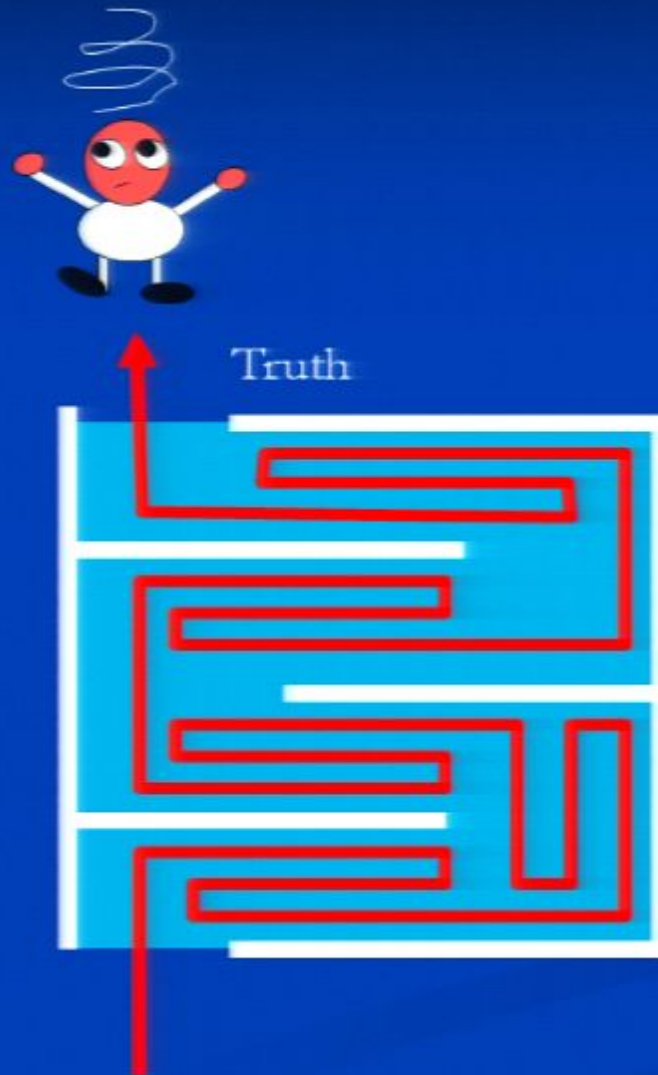


Retraction

- Choosing T and then not choosing T next



Aim: Eliminate **Needless Retractions**



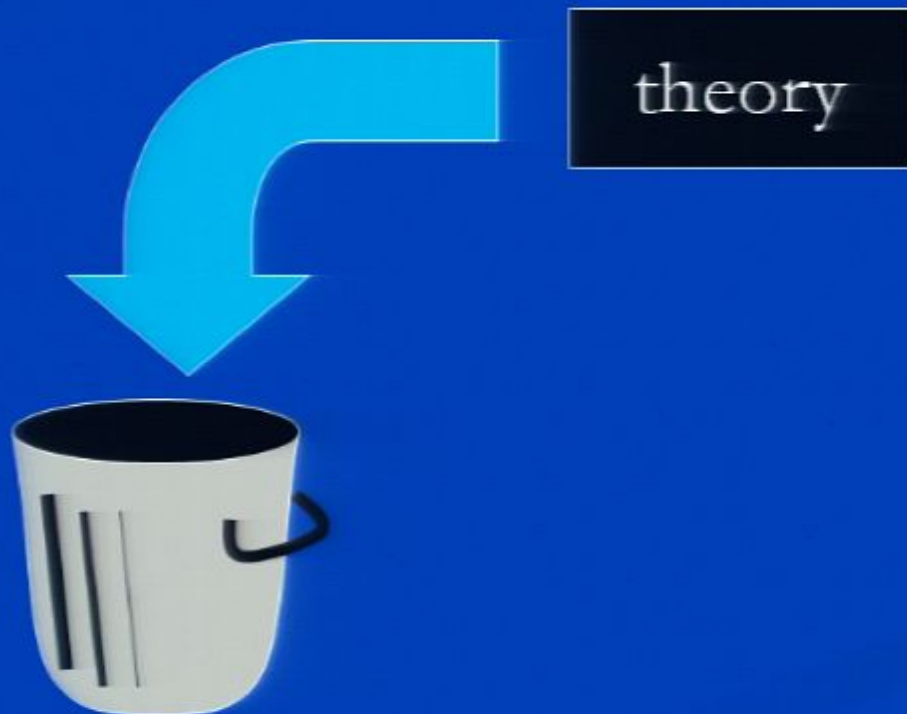
Aim: Eliminate Needless Retractions



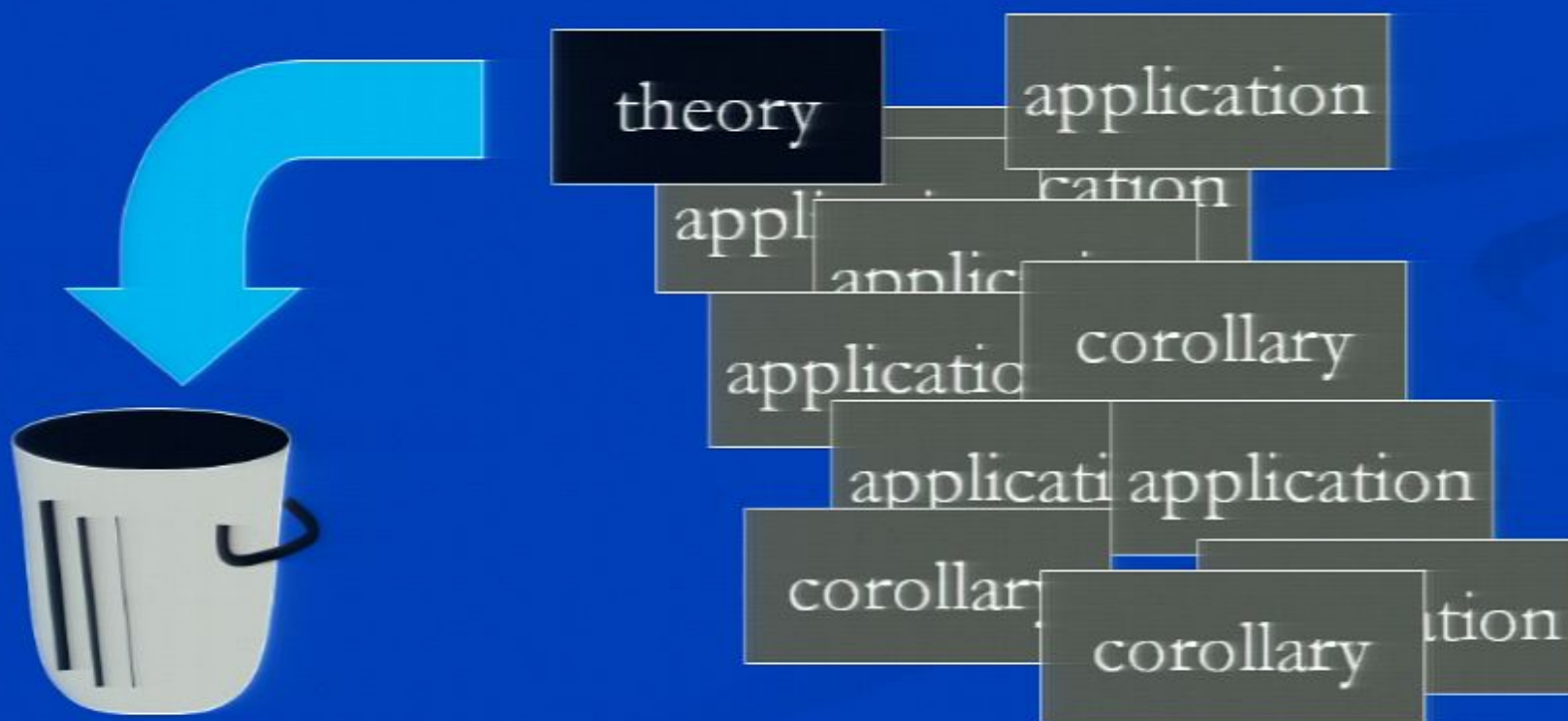
Truth



Aim: Eliminate **Needless Delays** to Retractions



Aim: Eliminate **Needless Delays** to Retractions



An Epistemic Motive

1. Future retraction is Gettier situation even if current belief is true.
2. If Gettier situations are bad, more of them are worse.

Easy Retraction Time Comparisons

Method 1



at least as **many**
at least as **late**

Method 2



Worst-case Retraction Time Bounds

$(1, 2, \dots, \infty)$
 $\downarrow \quad \downarrow \quad \downarrow$
 $\vdots \quad \vdots \quad \vdots$

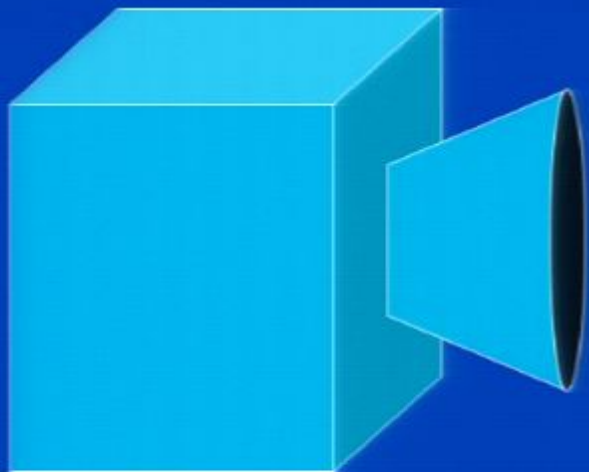


| | | | | | | | |
|----|----|----|----|----|----|----|-----|
| T1 | T2 | T3 | T3 | T3 | T3 | T4 | ... |
| T1 | T2 | T3 | T3 | T3 | T4 | T4 | ... |
| T1 | T2 | T3 | T3 | T4 | T4 | T4 | ... |
| T1 | T2 | T3 | T4 | T4 | T4 | T4 | ... |

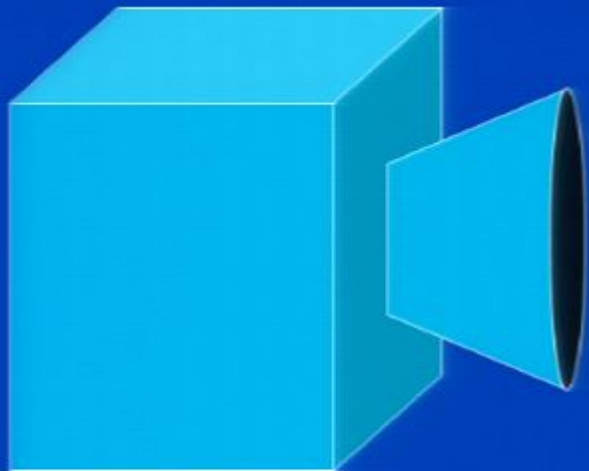
Output sequences

IV. Ockham Without Circles, Evasions, or Magic

Empirical Effects



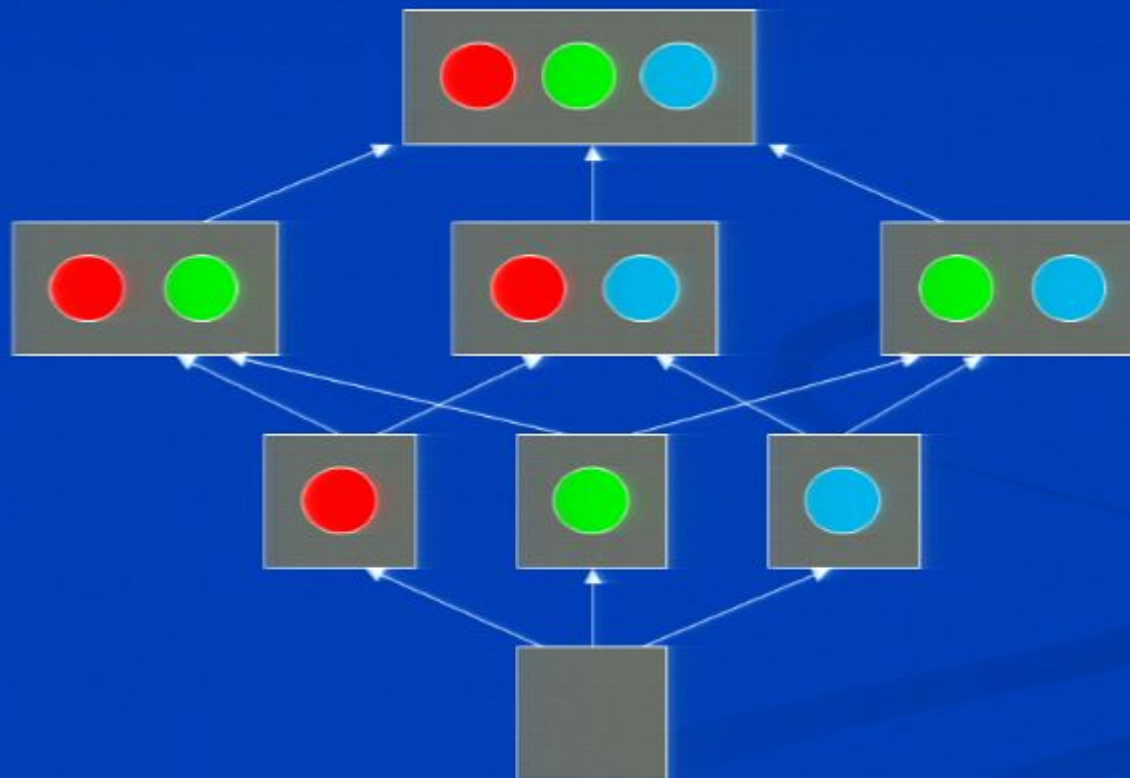
Empirical Effects



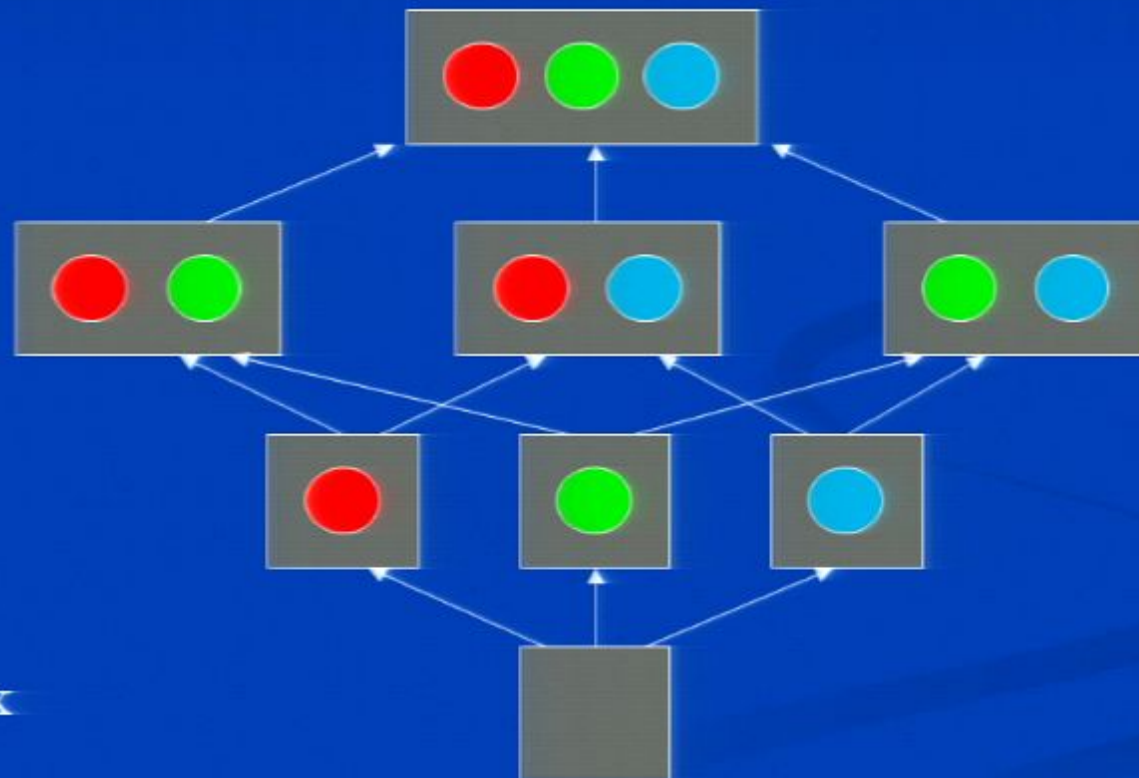
May take arbitrarily long to discover

Empirical Theories

- True theory determined by which effects appear.



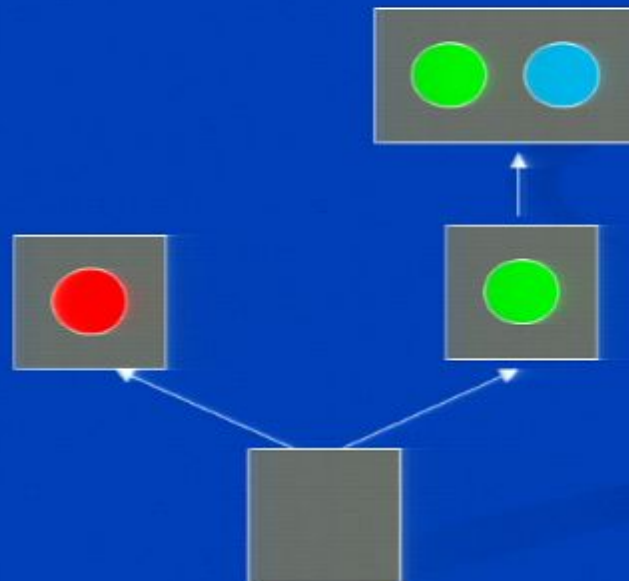
Empirical Complexity



More complex

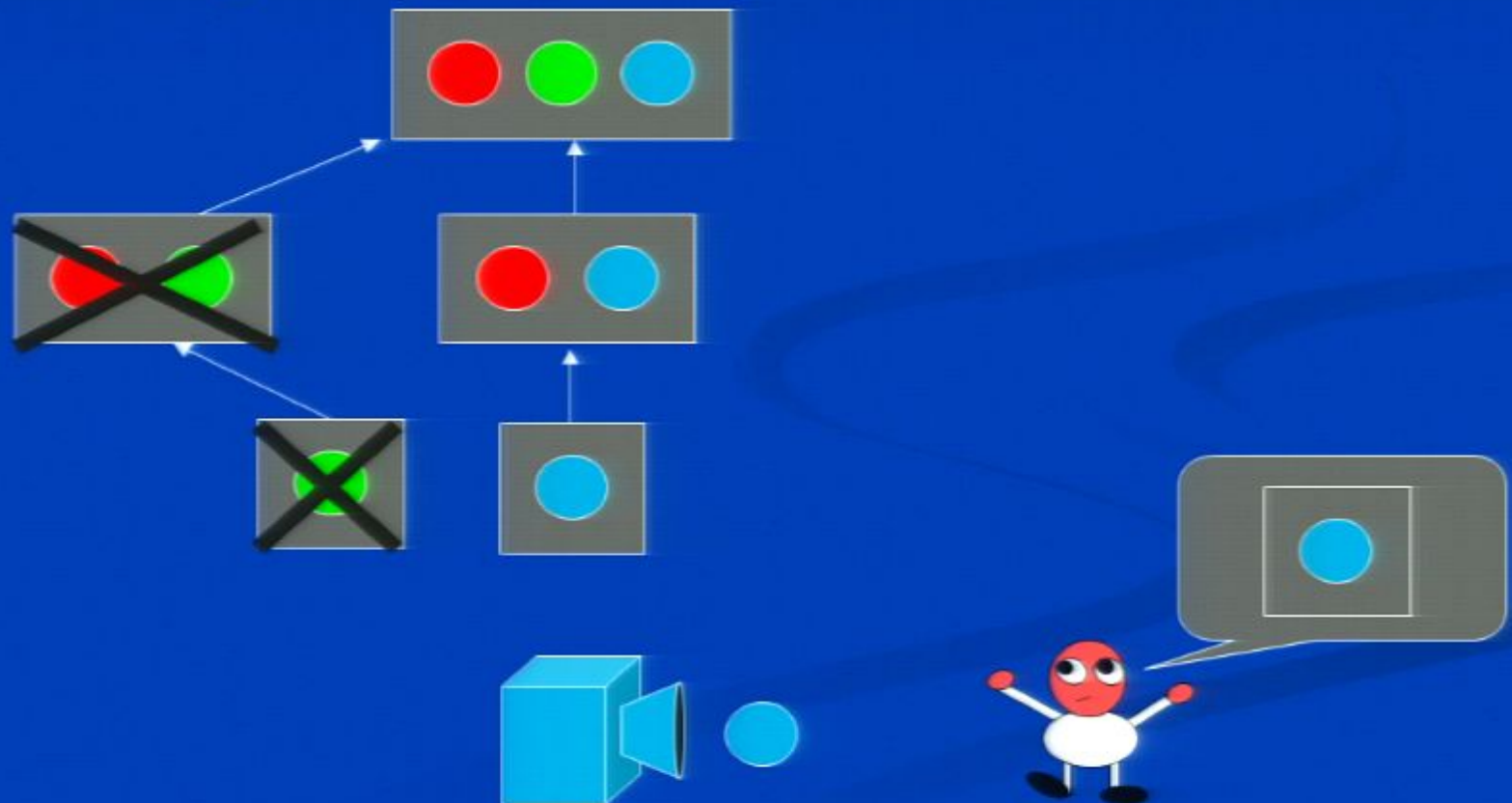
Assume **No Short Paths**

Weaker results if some path is shorter than some other path.



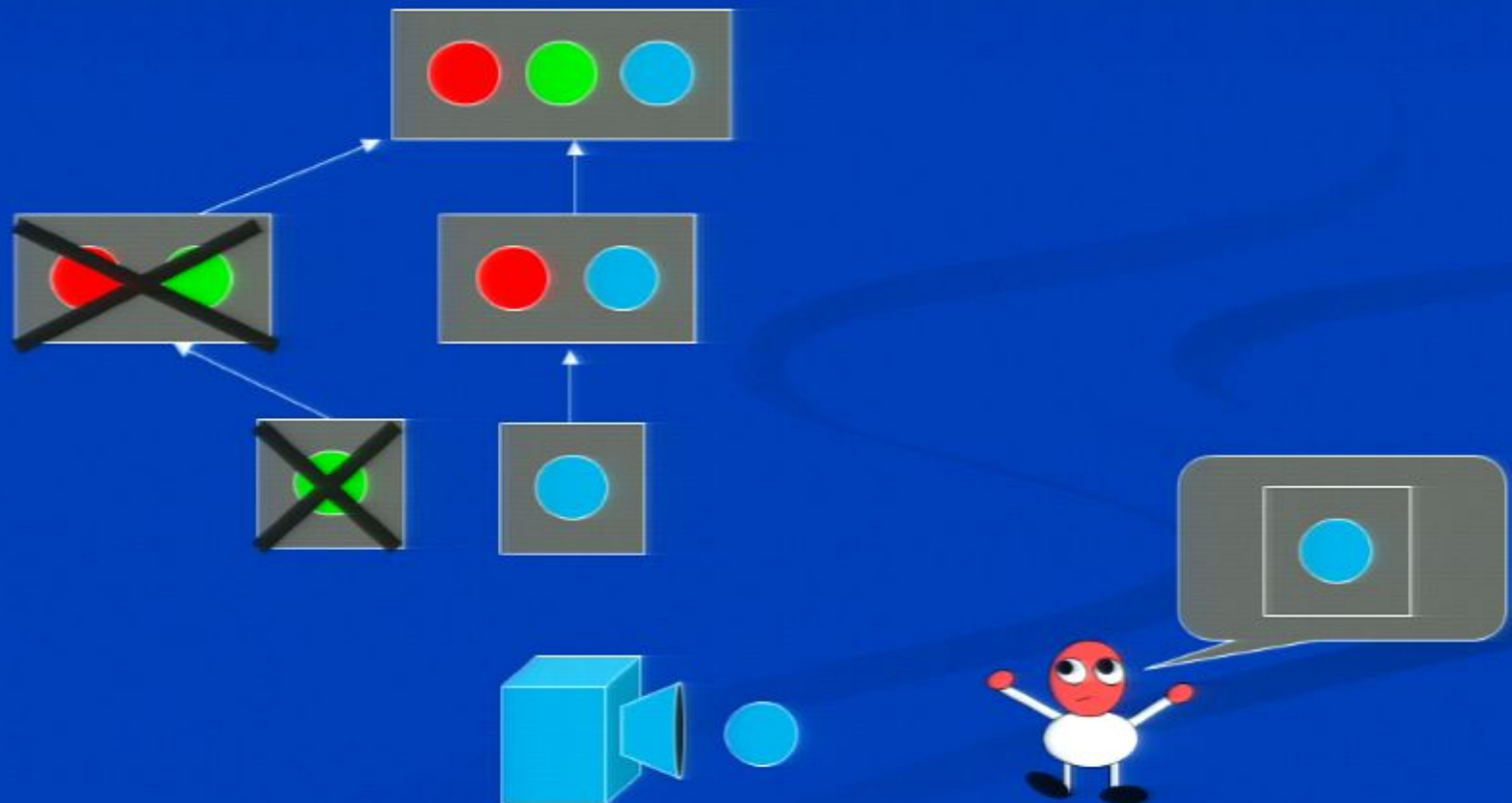
Ockham's Razor

- Don't select a theory unless it is uniquely simplest in light of experience.



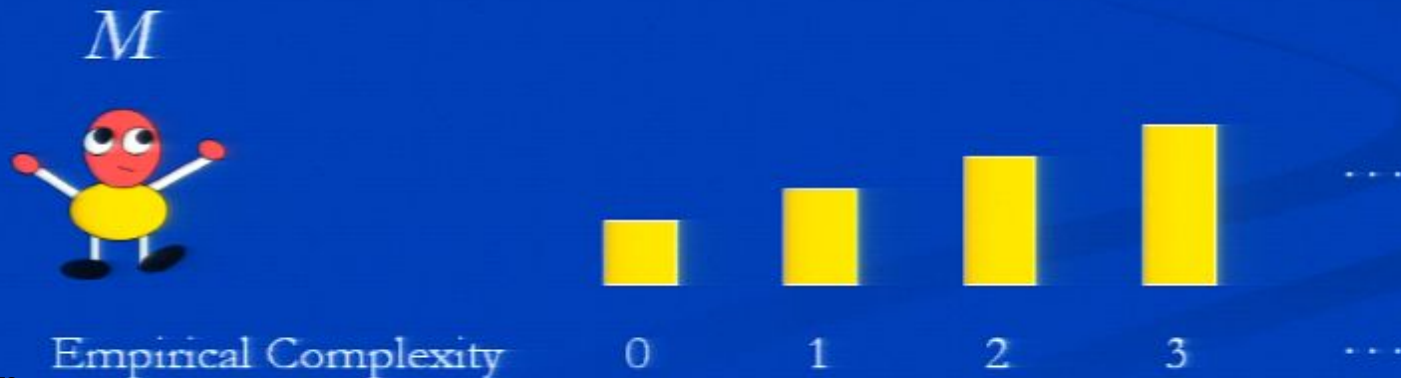
Stalwartness

- Don't retract your answer while it is uniquely simplest.



Timed Retraction Bounds

$r(M, e, n)$ = the least timed retraction bound covering the total timed retractions of M along input streams of complexity n that extend e

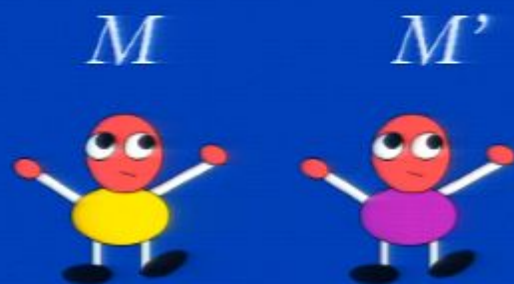


Efficiency of Method M at e

M converges to the truth no matter what;

For each convergent M' that agrees with M up to the end of e , and for each n :

$$\blacksquare r(M, e, n) \leq r(M', e, n)$$



Empirical Complexity



M is **Beaten** at e

There exists convergent M' that agrees with M up to the end of e , such that

- $r(M, e, n) > r(M', e, n)$, for each n .



Basic Idea

- **Ockham efficiency:** Nature can force arbitrary, convergent M to produce the successive answers down an effect path arbitrarily late, so stalwart, Ockham solutions are efficient.

Basic Idea

- **Unique Ockham efficiency:** A violator of Ockham's razor or stalwartness can be forced into an extra retraction or a late retraction at the time of the violation, so the violator is beaten by each stalwart, Ockham solution.

Ockham Efficiency Theorem

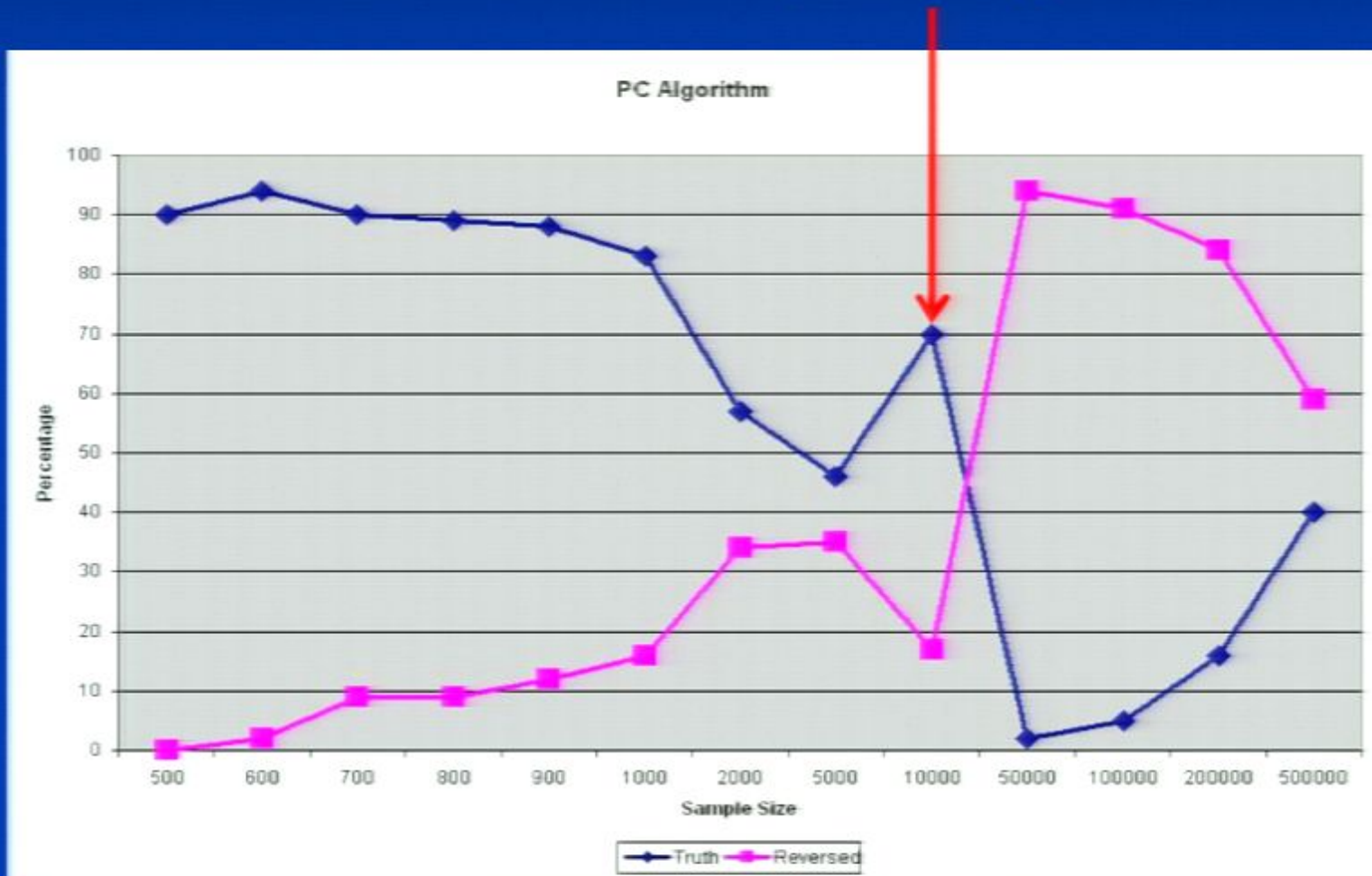
- Let M be a solution.
- The following are equivalent:
 - M is henceforth Ockham and stalwart;
 - M is henceforth efficient;
 - M is henceforth never beaten.

Some Applications

- Polynomial laws
- Conservation laws
- Causal networks

Concrete Recommendations

Extra retraction by PC algorithm

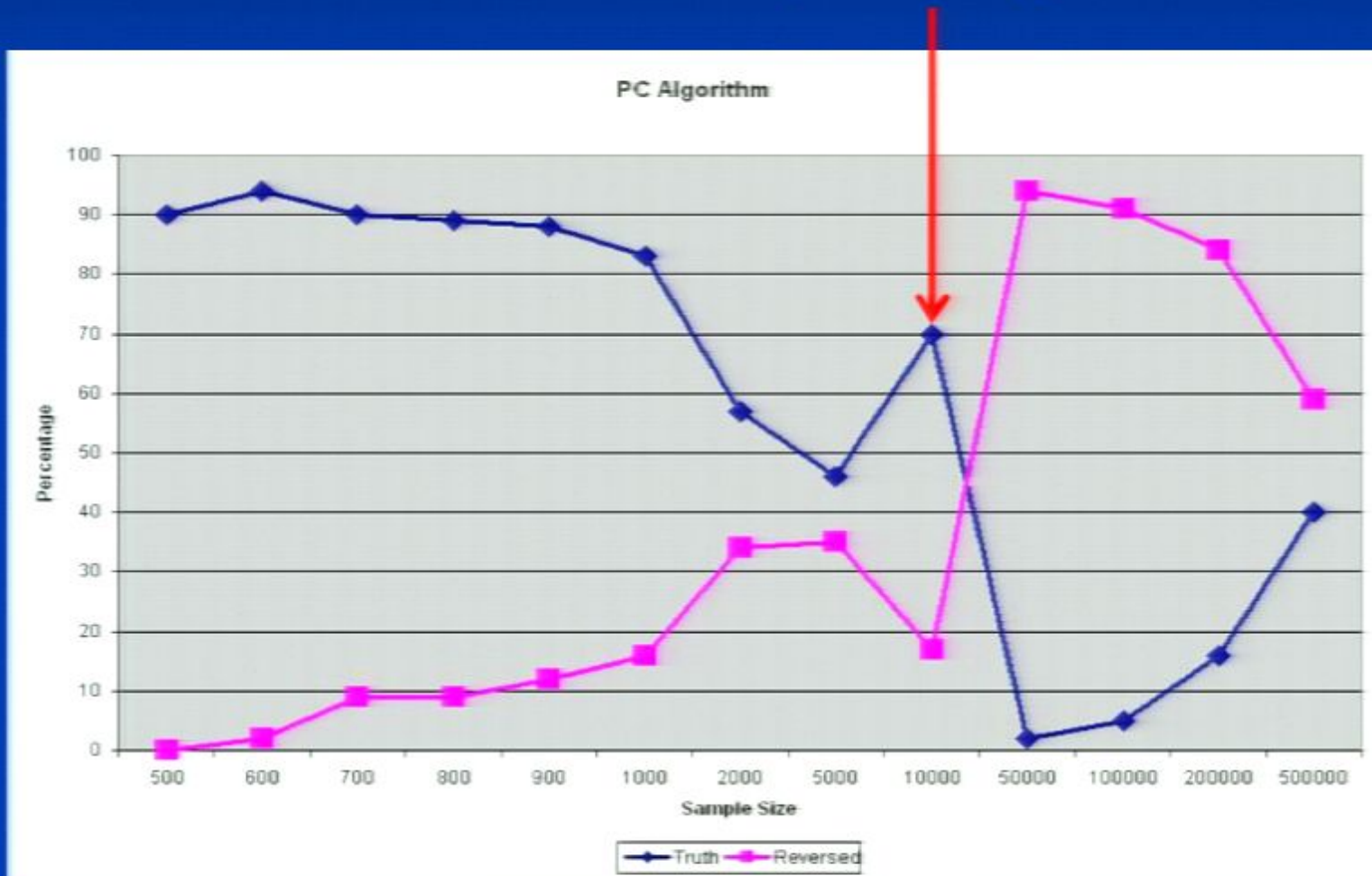


Generalizations

- Generalized Simplicity
- Bayesian retractions
- Mixed Strategies
 - Times of retractions in chance
 - Expected retraction times
- Drop the no short paths assumption...
- Retractions weighted by content...
- Statistical inference...

Concrete Recommendations

Extra retraction by PC algorithm



Generalizations

- Generalized Simplicity
- Bayesian retractions
- Mixed Strategies
 - Times of retractions in chance
 - Expected retraction times
- Drop the no short paths assumption...
- Retractions weighted by content...
- Statistical inference...

Conclusions

- Ockham's razor is necessary for staying on the straightest path to the true theory but does not point at the true theory.
- A priori justification---no evasions or circles are required.
- Analogous to computational complexity theory.

