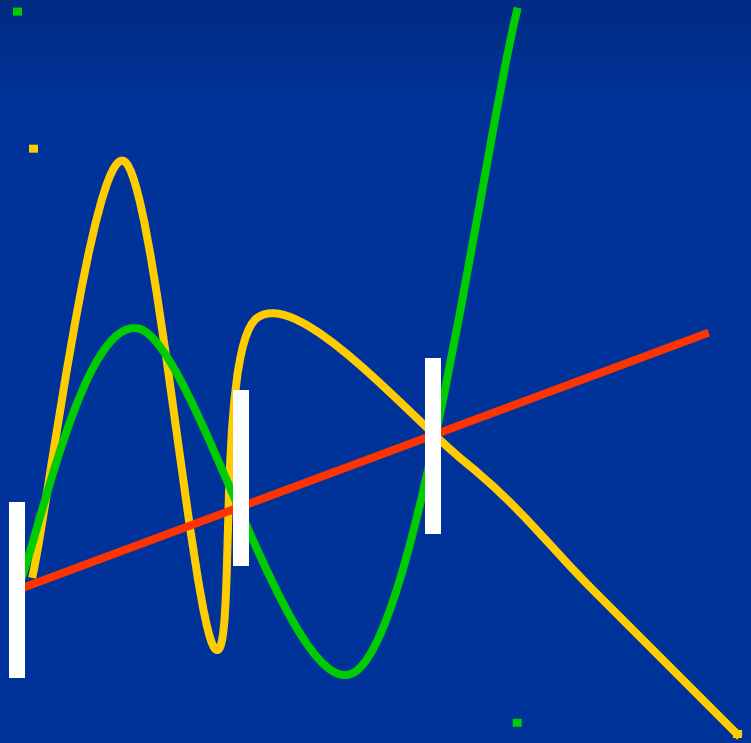
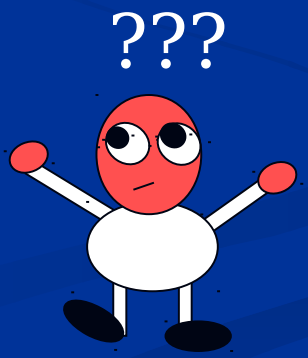
Ockham's Razor without Circles, Evasions, or Magic

Kevin T. Kelly
Department of Philosophy
Carnegie Mellon University
www.cmu.edu

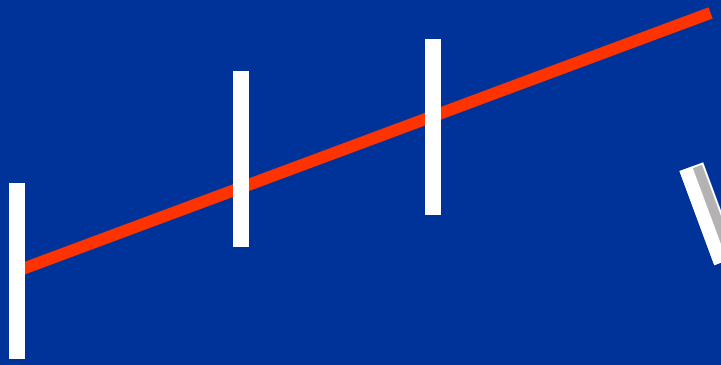
I. The Simplicity Puzzle

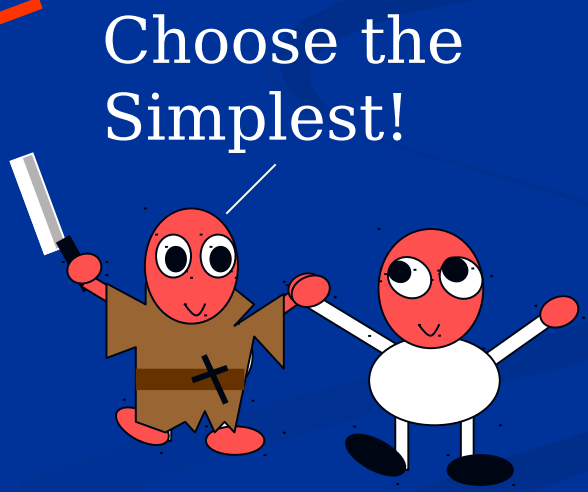
Which Explanation is Right?



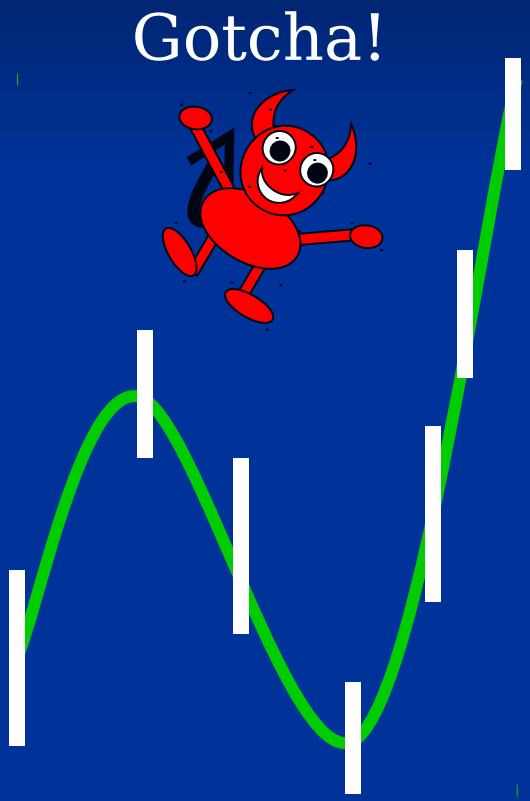


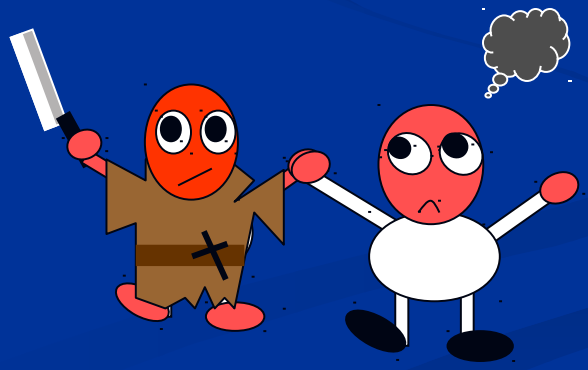
Ockham Says:





But Maybe...





The Simplicity Puzzle

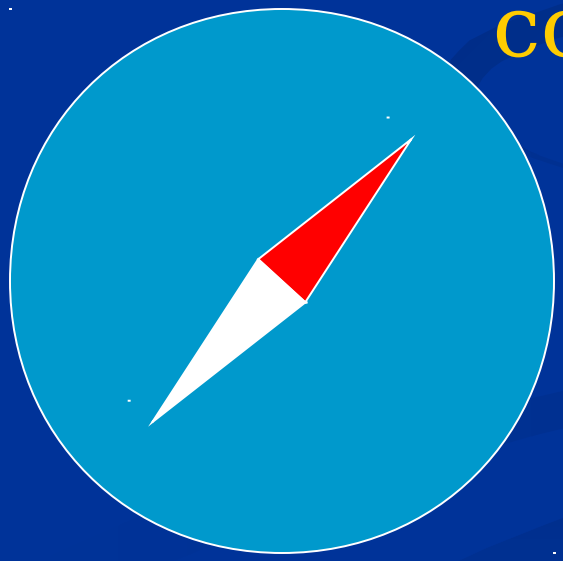
- An indicator must be sensitive to what it indicates.

simple

The Simplicity Puzzle

- An indicator must be sensitive to what it indicates.

complex



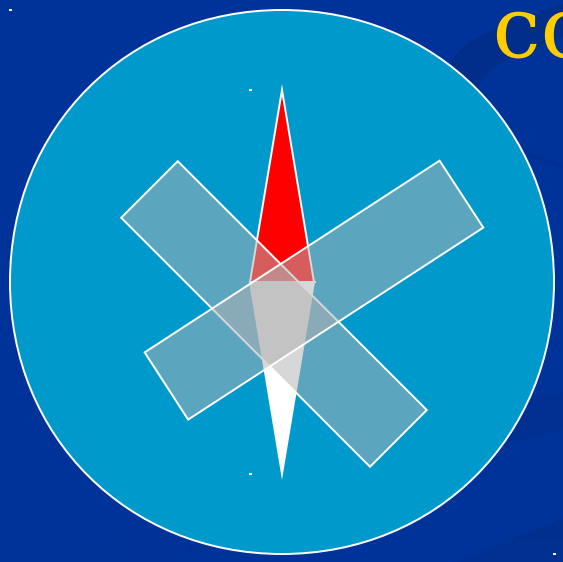
The Simplicity Puzzle

- But Ockham's razor always points at simplicity.

simple

The Simplicity Puzzle

- But Ockham's razor always points at simplicity.



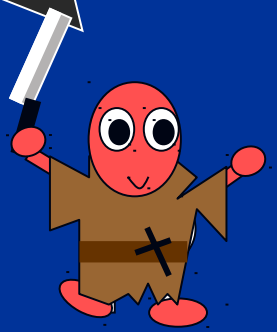
complex

A. Circular Accounts

Prior Probability

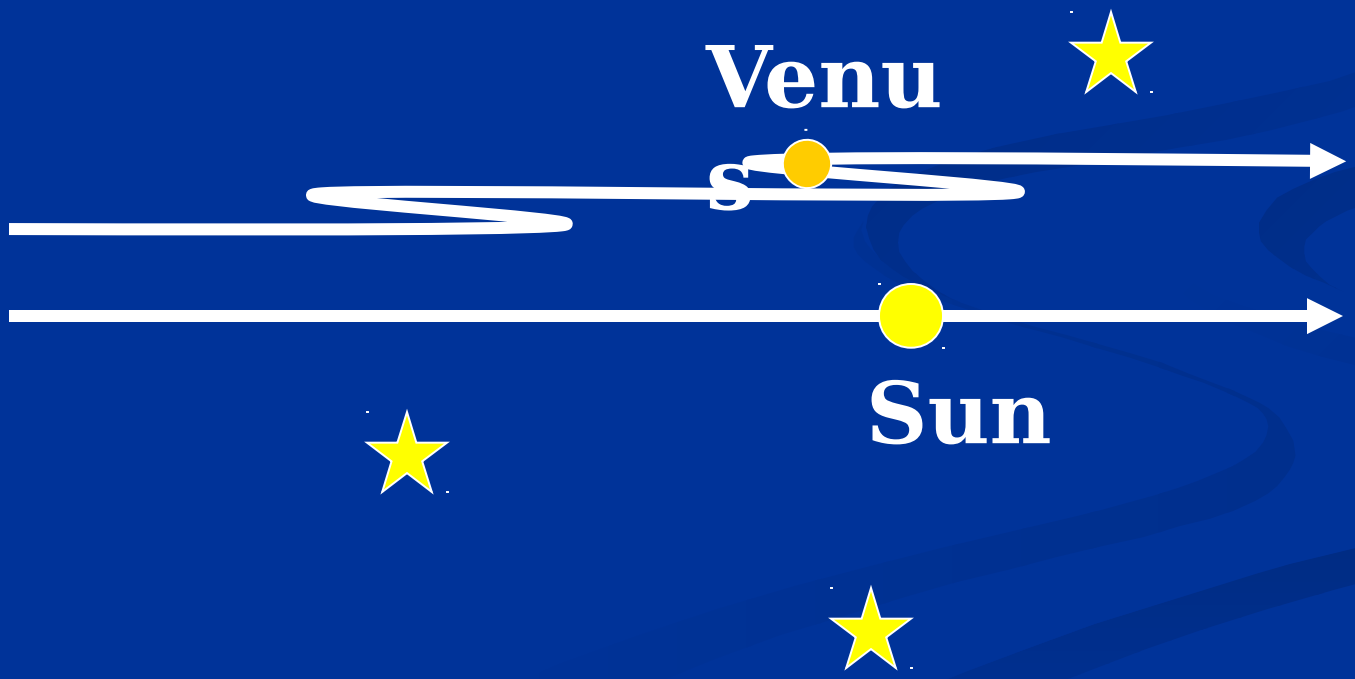
- Assign high prior probability to simple theories.

On the presumption of simplicity,
presume simplicity.



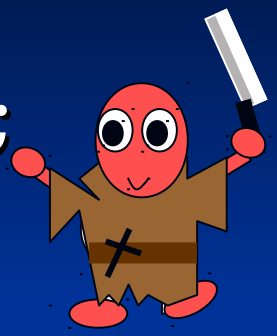
Miracle Argument

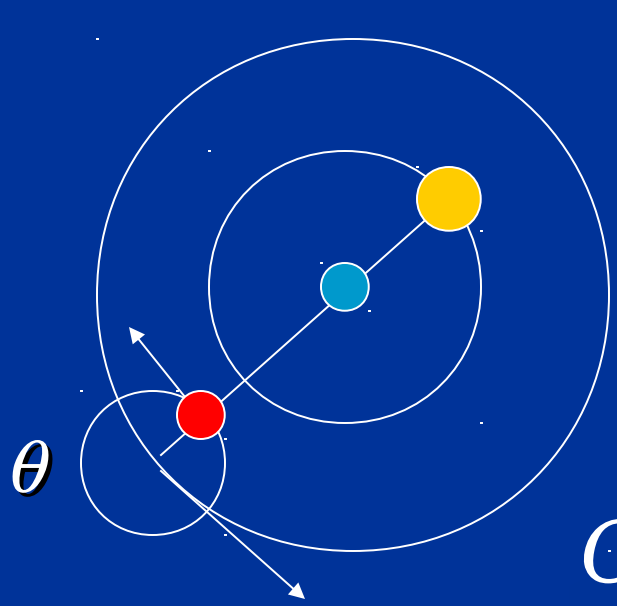
- Phenomenon *e*: Venus appears to bob back and forth about the sun against the fixed stars.



Miracle Argument

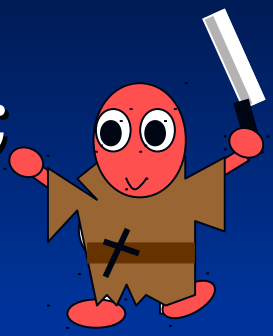
- e would not be a miracle given S;
- e would be a miracle given C.

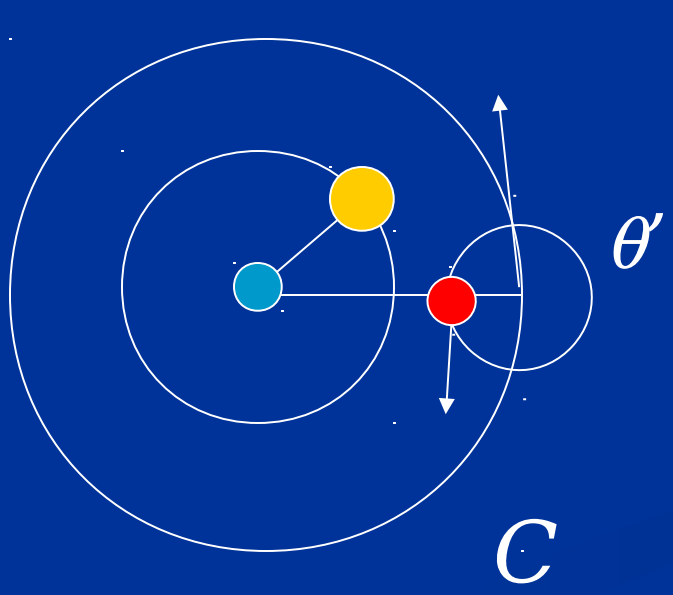




Miracle Argument

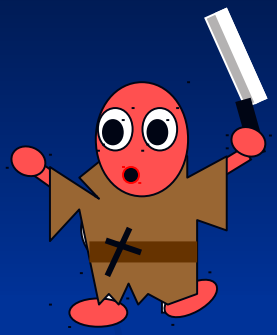
- e would not be a miracle given S;
- e would be a miracle given C.



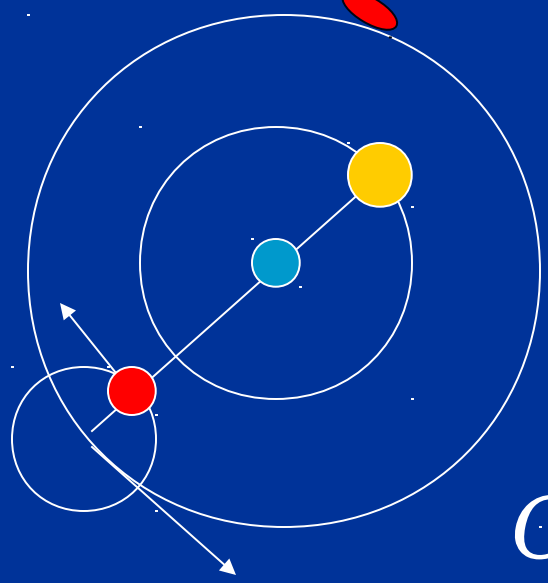


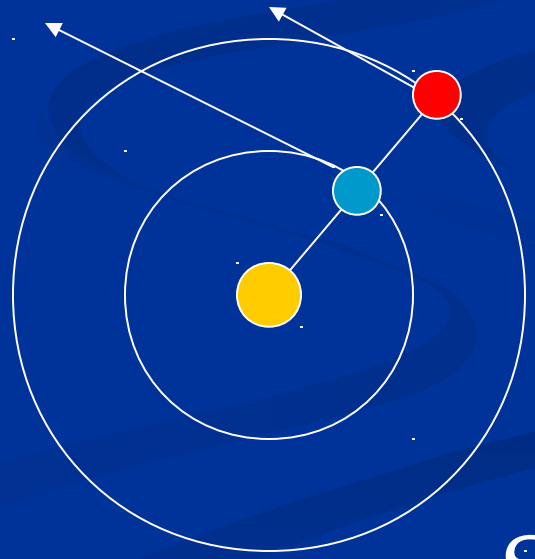
However…

- e would not be a miracle given $C(\theta)$;



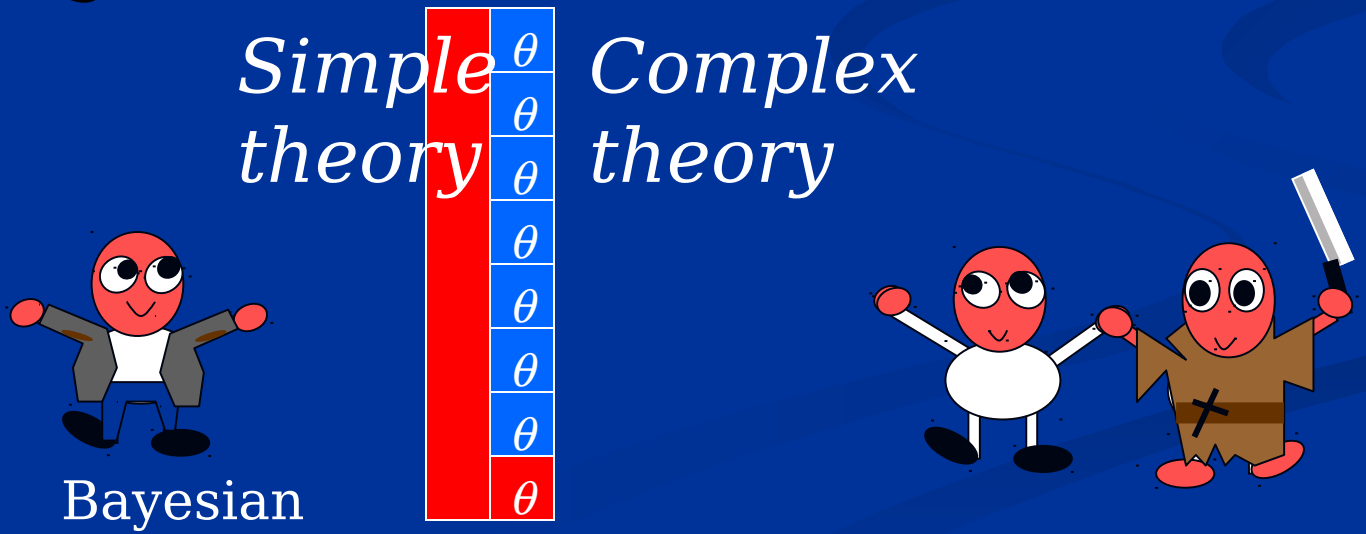
Why not this?





Bayesian Explanation

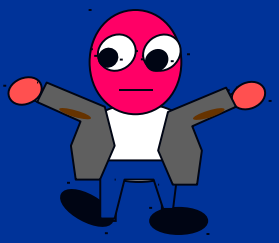
- Ignorance (over theories)
- + Ignorance (over parameter settings in each theory)
- = Knowledge (against complex parameter settings).



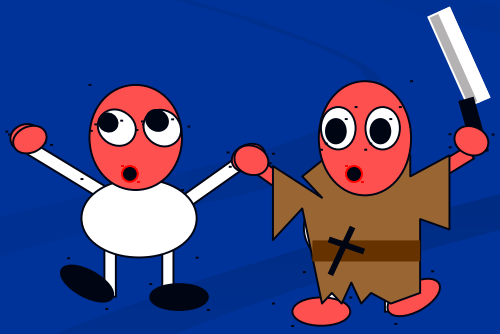
= The Old Paradox of Indifference

- Ignorance (over blue, non-blue)
- \+ Ignorance (over ways of being non-blue)
- = Knowledge (that the truth is blue as opposed to any other color)

Blue *Not-Blue*

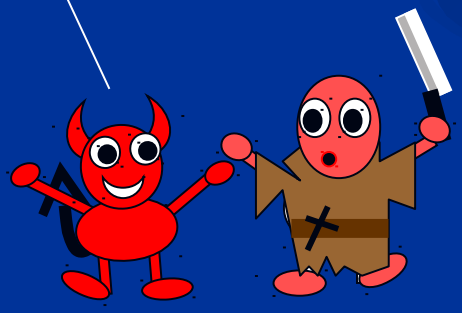


Bayesian



In Any Event…

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



B. Evasive Accounts

Theoretical Virtues

- Simple theories have virtues:
  - Testable
  - Unified
  - Explanatory
  - Symmetrical
  - Bold
  - Compress data

Theoretical Virtues

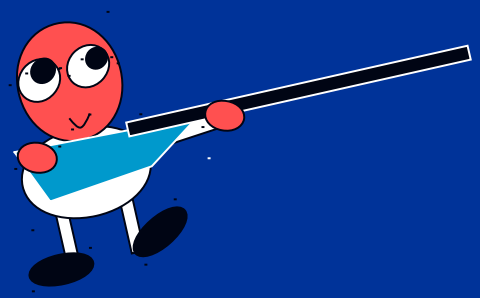
- Simple theories have virtues:
  - Testable
  - Unified
  - Explanatory
  - Symmetrical
  - Bold
  - Compress data

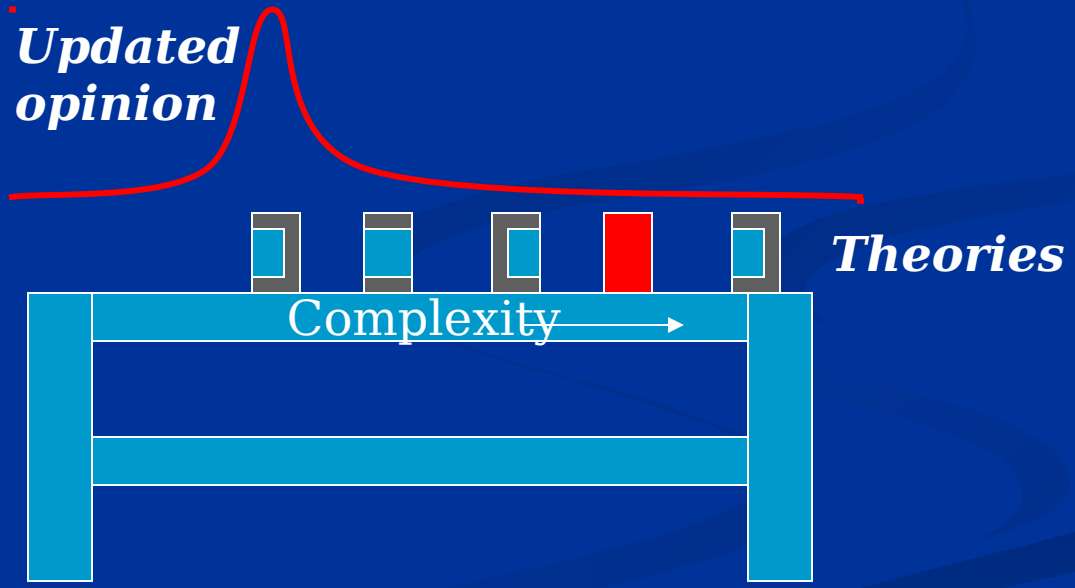
Kelly Maximus

- But to conclude that the virtuous theory is true is wishful thinking.

Bayesian Convergence

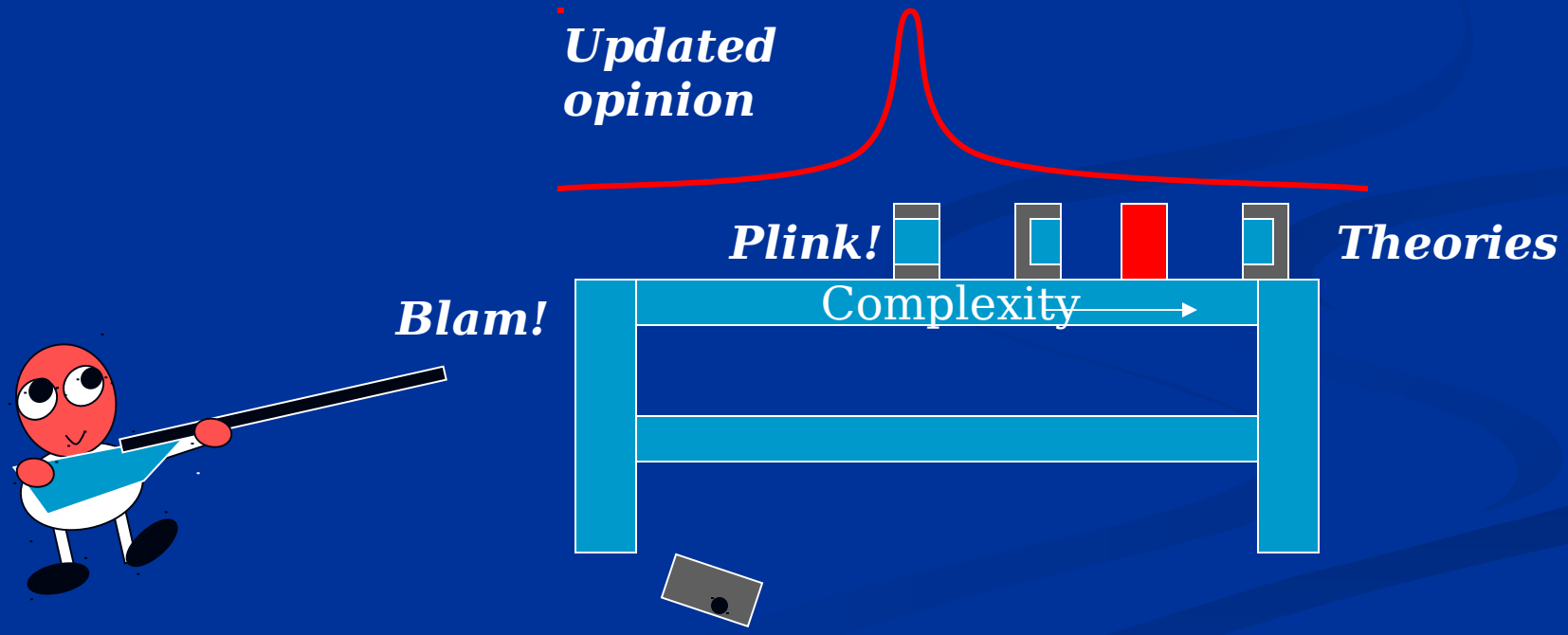
- Too-simple theories get shot down…





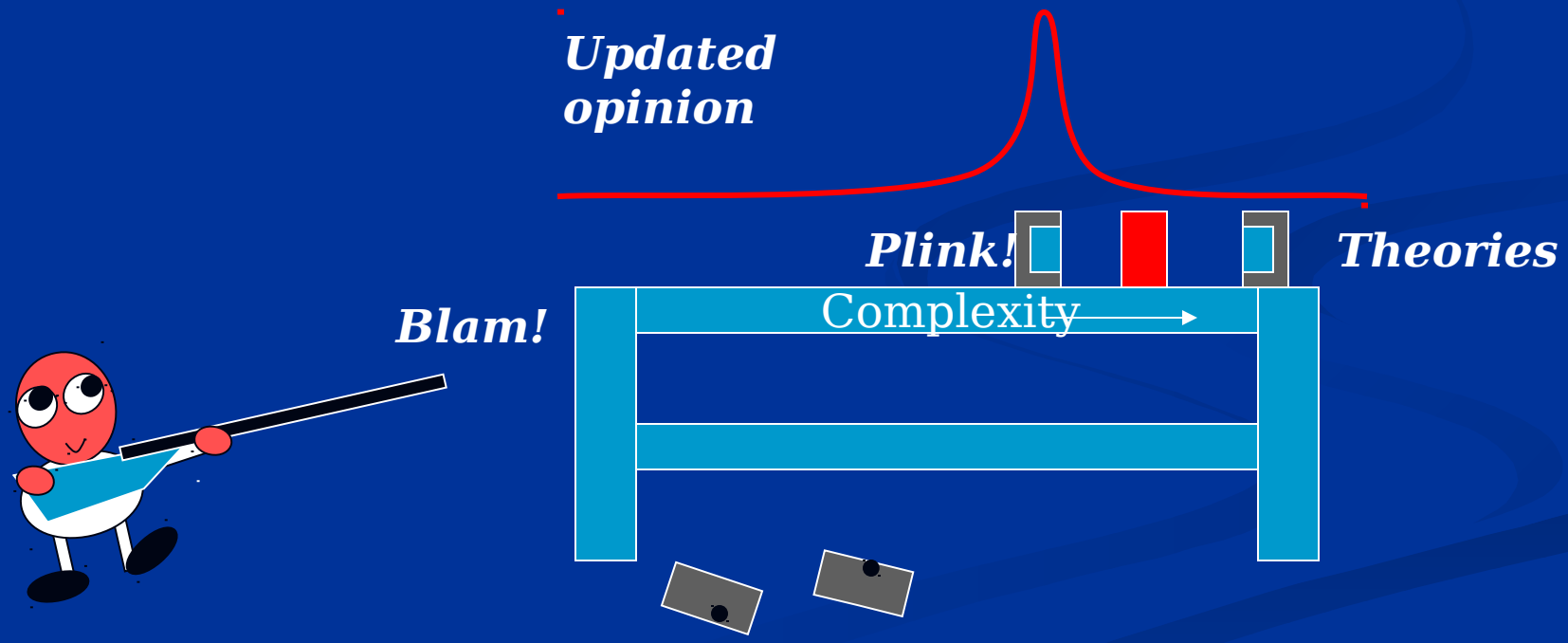
Bayesian Convergence

- Plausibility is transferred to the next-simplest theory…



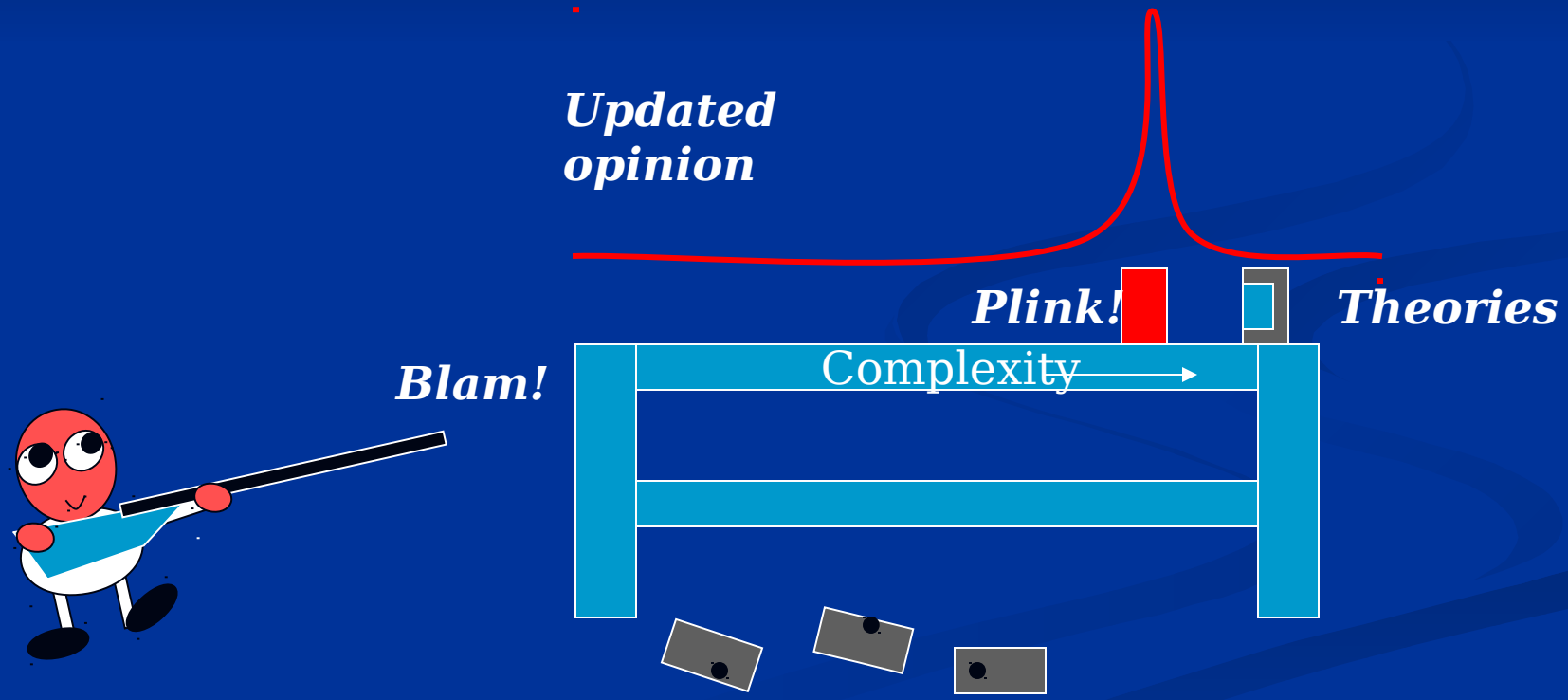
Bayesian Convergence

- Plausibility is transferred to the next-simplest theory…



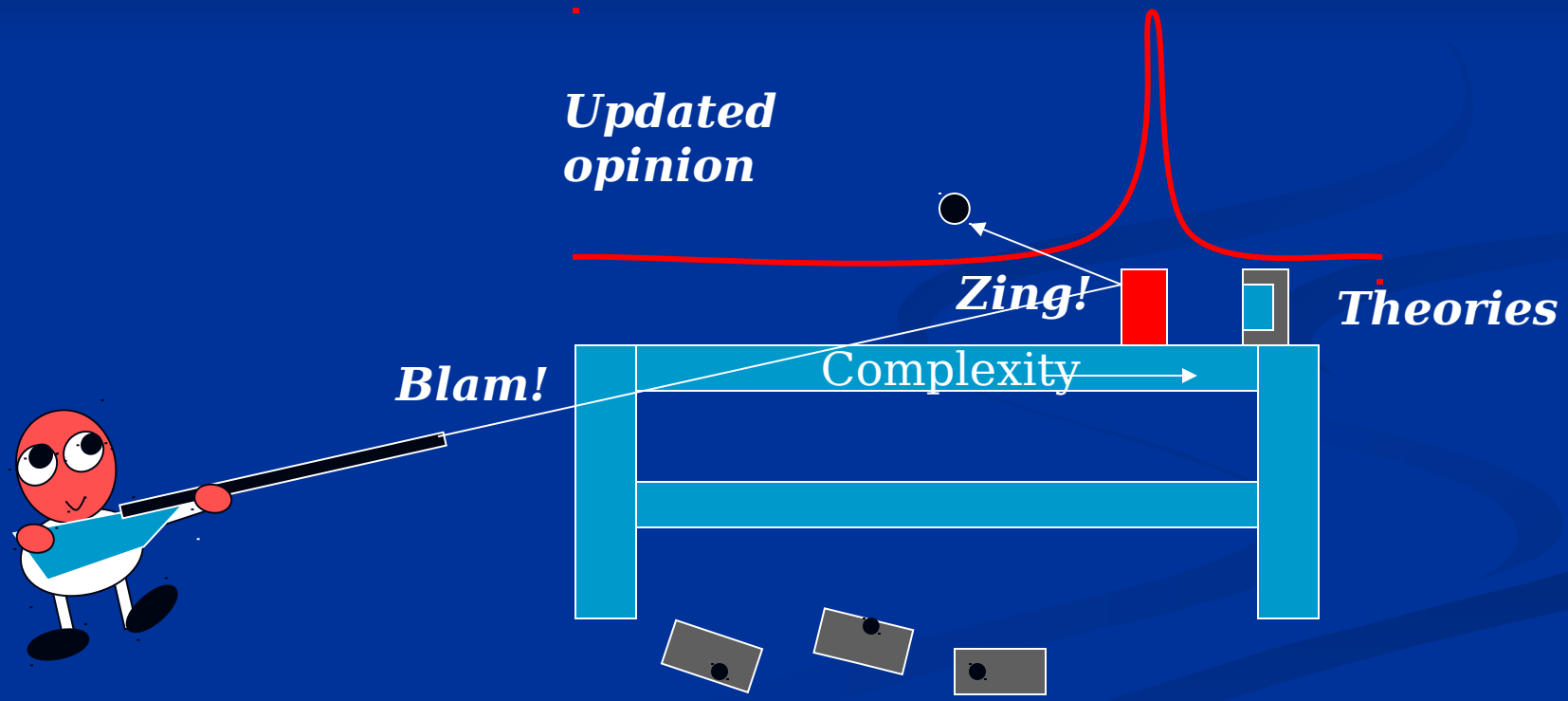
Bayesian Convergence

- Plausibility is transferred to the next-simplest theory…



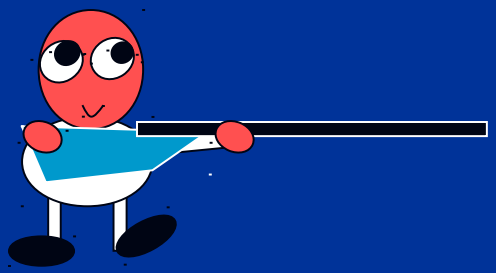
Bayesian Convergence

- The true theory is nailed to the fence.

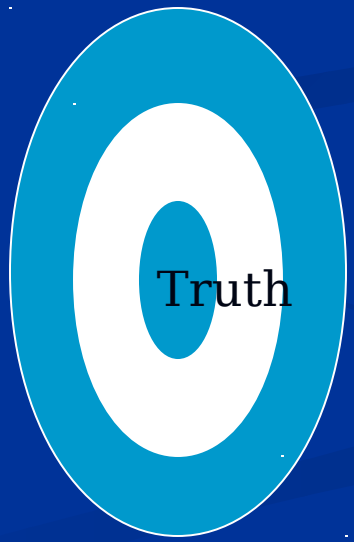


Over-fitting

- Empirical estimates based on complex models have greater expected distance from the truth.

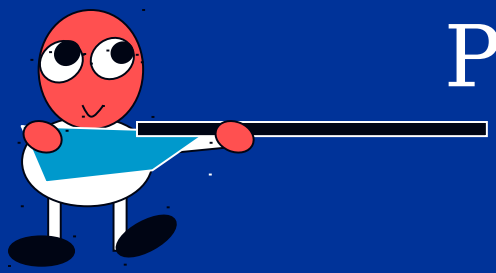


Unconstrained aim at the true value

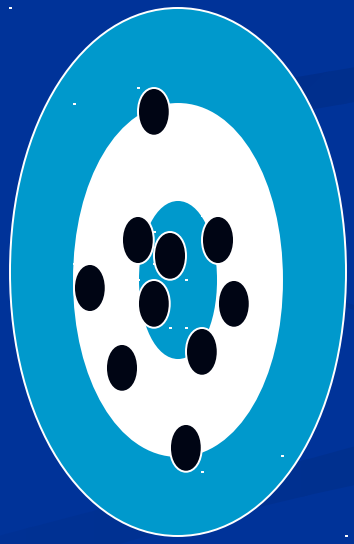


Over-fitting

- Empirical estimates based on complex models have greater expected distance from the truth.



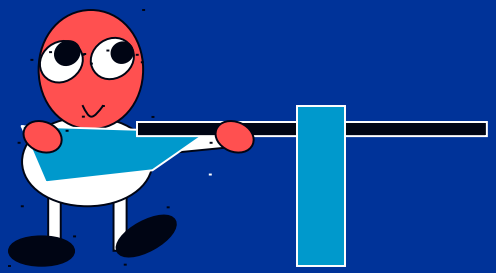
Pop!



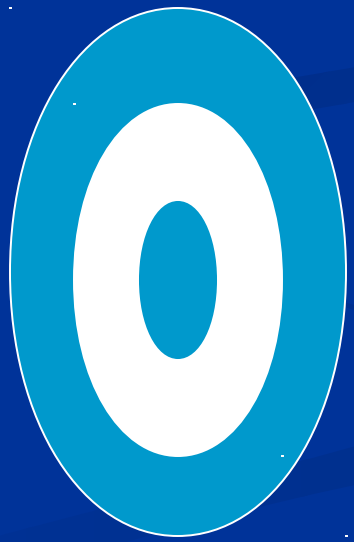
Unconstrained aim at the true value

Overfitting

- Empirical estimates based on constrained models can have lower expected distance from the truth…



Clamped aim
at the true value



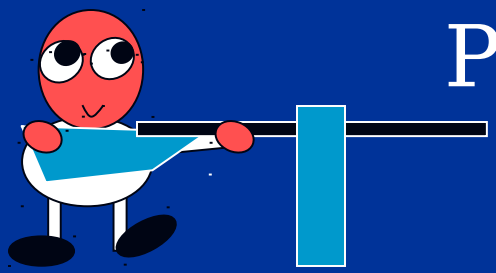
Convergence

- But alternative strategies also converge:
  - Any finite variant of a convergent strategy converges (Reichenbach, Salmon).

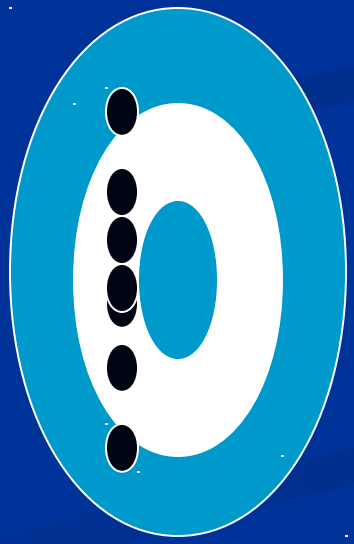
Over-fitting

- Empirical estimates based on constrained models can have lower expected distance from the truth…

(Akaike, Vapnik, Sober and Forster, etc.)



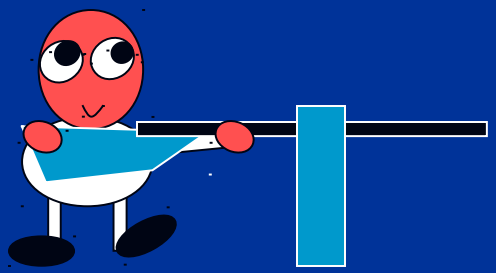
Pop!



Clamped aim
at the true value

Over-fitting

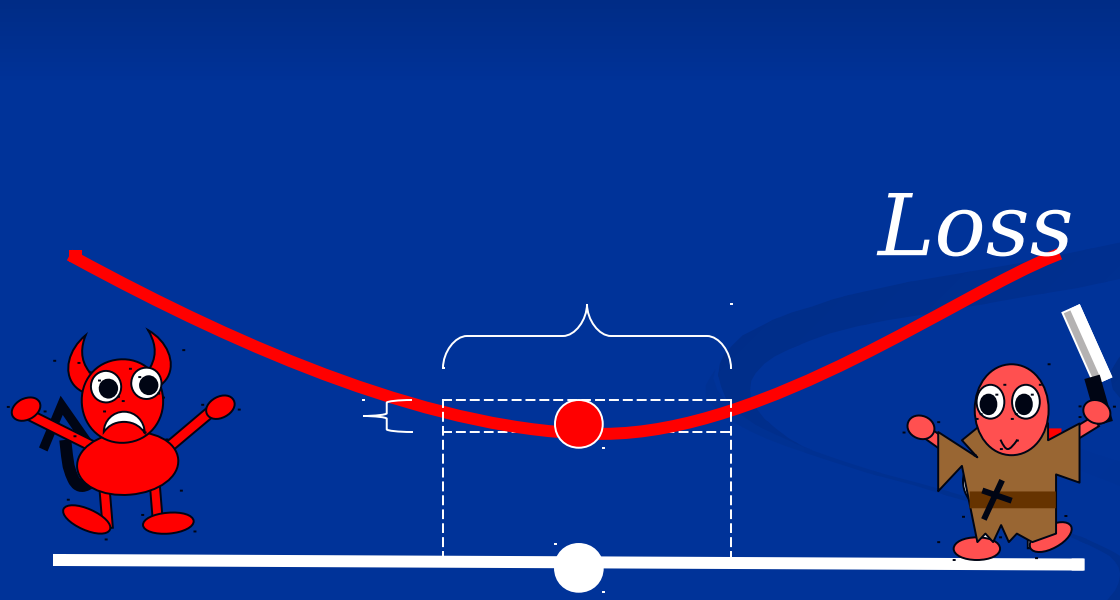
- ...even if the simple theory is known to be false...



Clamped aim
at the true value

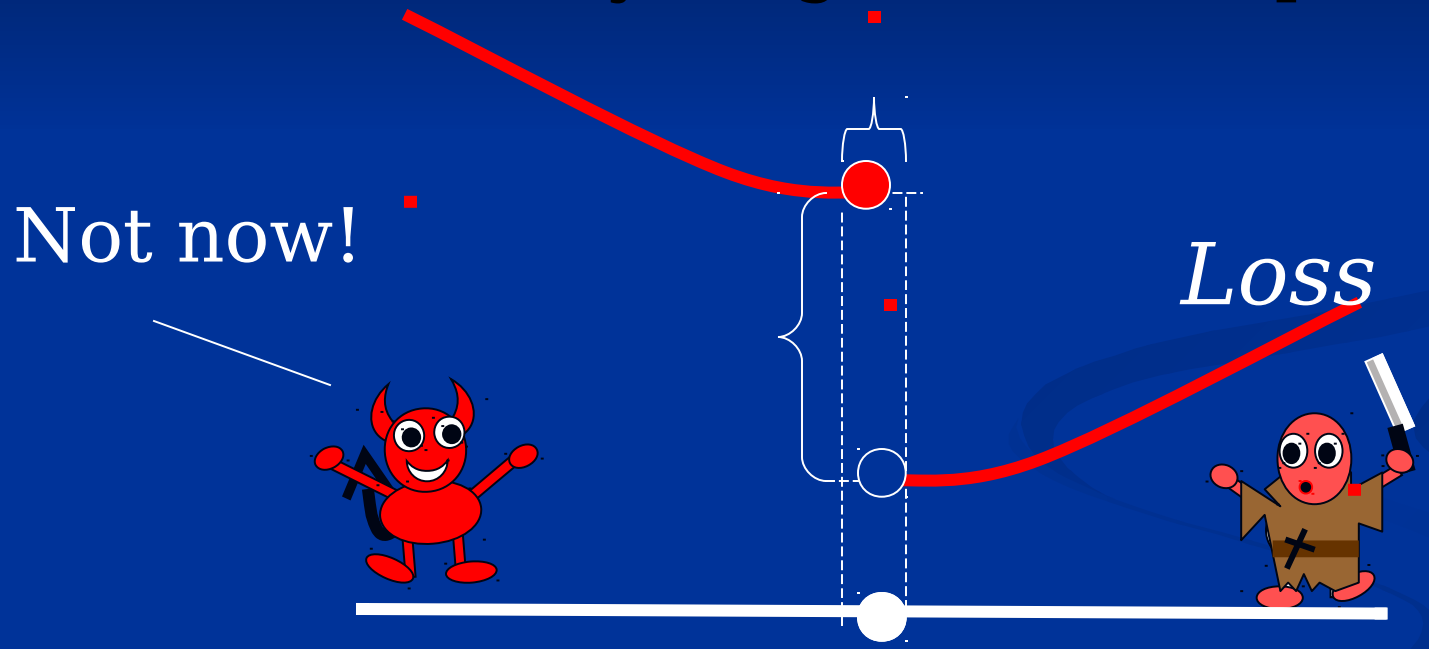
Makes Sense…

…when loss of an answer is similar in nearby distributions.



But Truth Matters…

…when loss of an answer is discontinuous with similarity---e.g., in causal prediction.

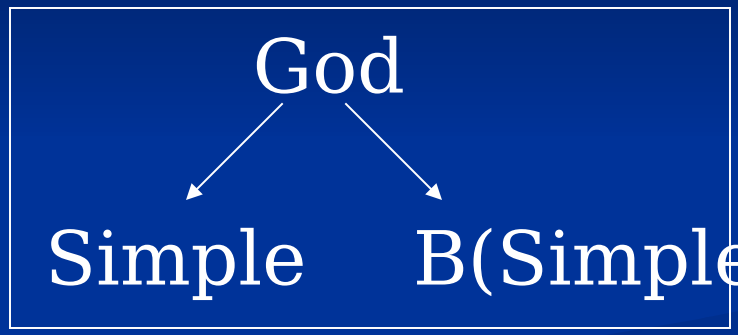


C. Magical Naturalism

Magic

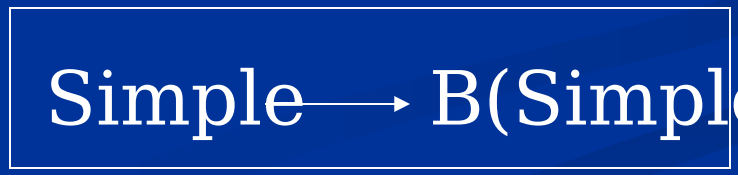
- Simplicity indicates truth via an unknown cause.

Leibniz



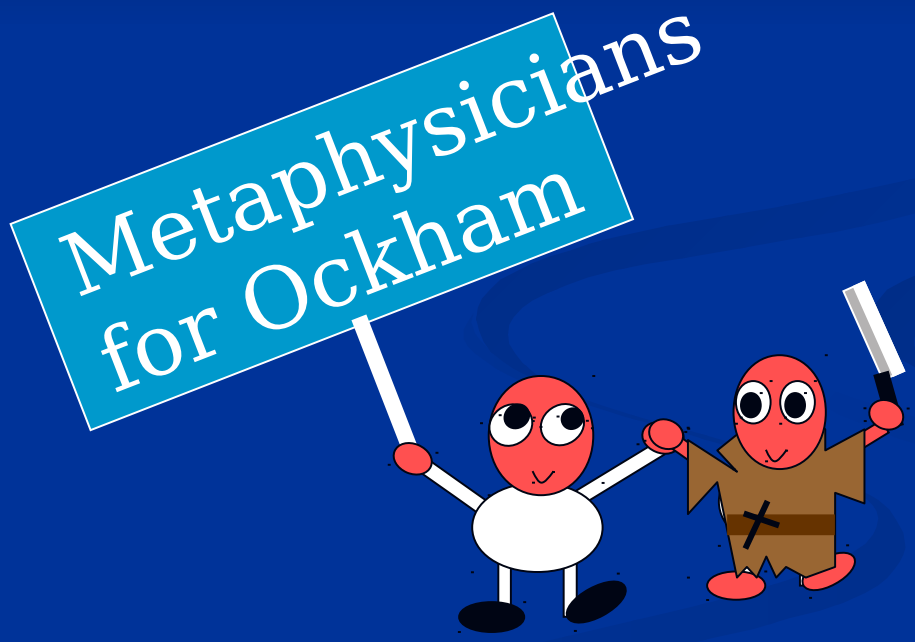
Kant

Ouija board



Magic

- Ockham says: explain Ockham's razor *without* undetected causes.



Magic

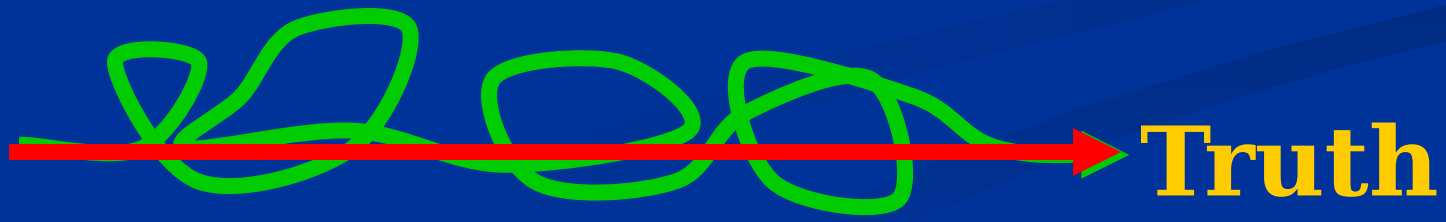
- Ockham says: explain Ockham's razor *without* undetected causes.

II. A New Direction

Ancient Hint

"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."

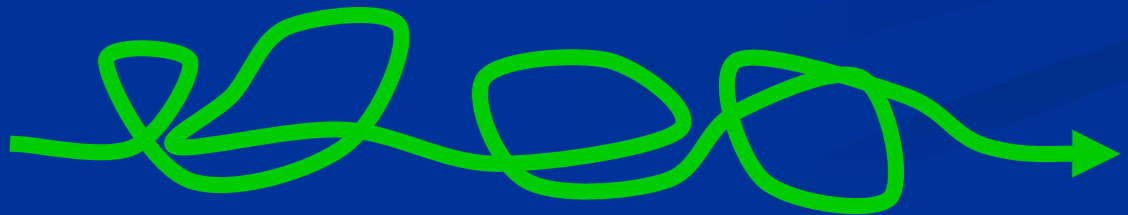
Katha Upanishad, I. ii. 5, c. 600 BCE.



Diagnosis of Standard Accounts

- **Short-run Reliability:** too strong.

- **Long-run Convergence**: too weak.



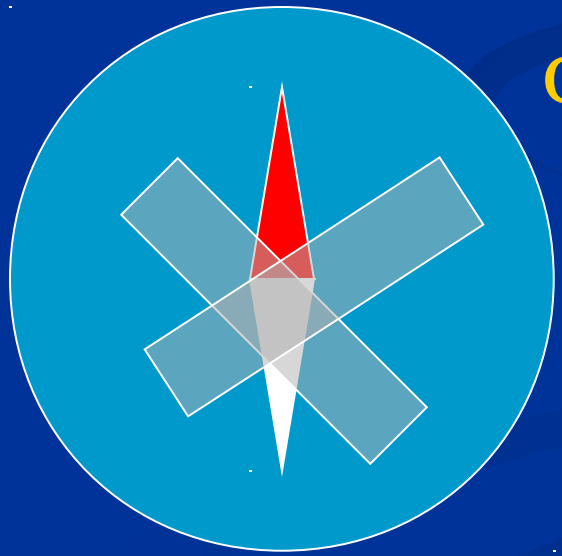
Natural Alternative

- **Short-run Reliability:** too strong.

- **Straightest Convergence:** just right?

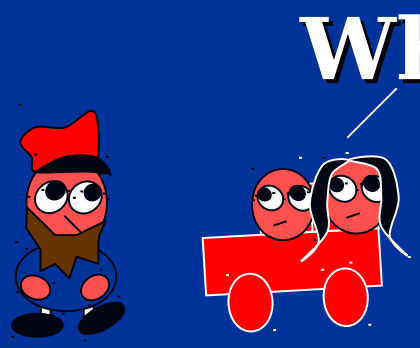
- **Long-run Convergence**: too weak.

III. Navigation by Fixed Directions



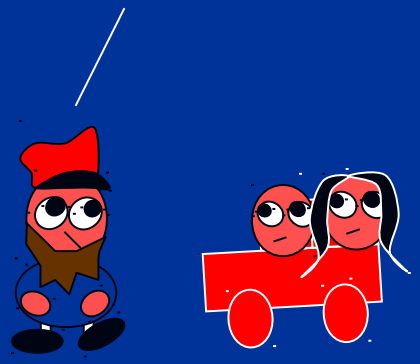
complex

Asking for Directions

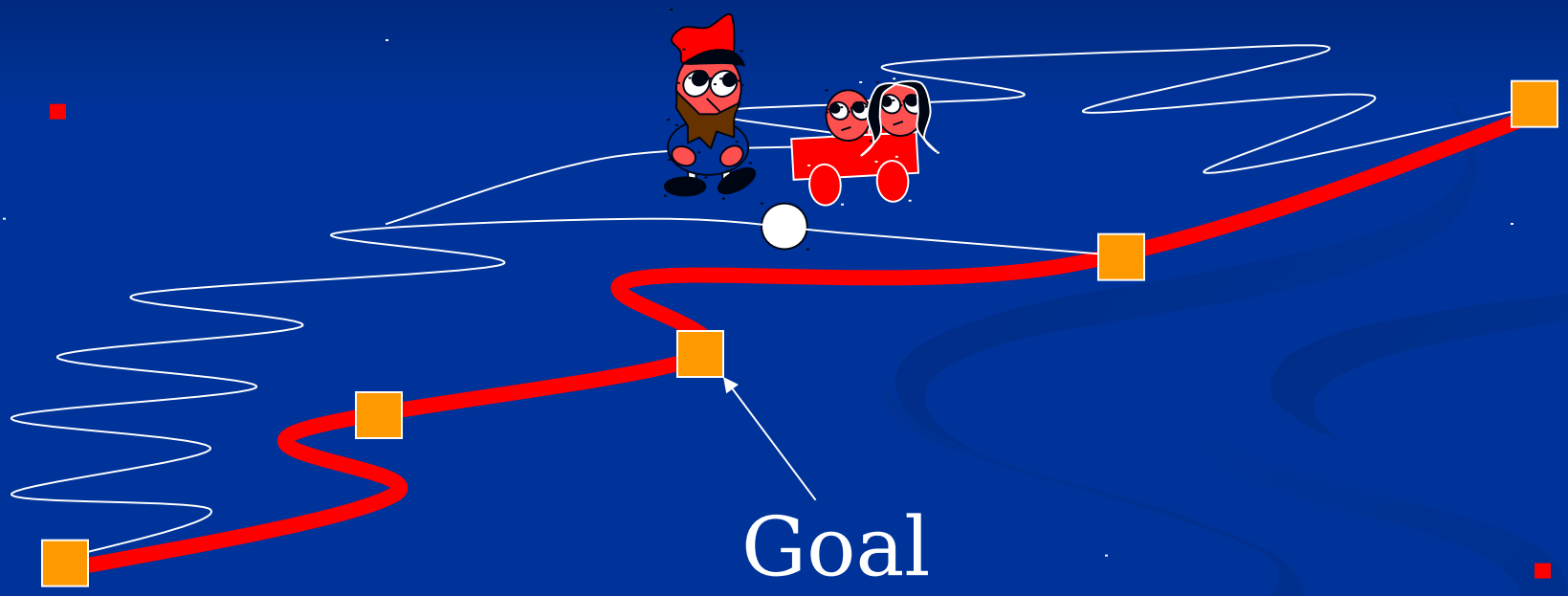


Asking for Directions

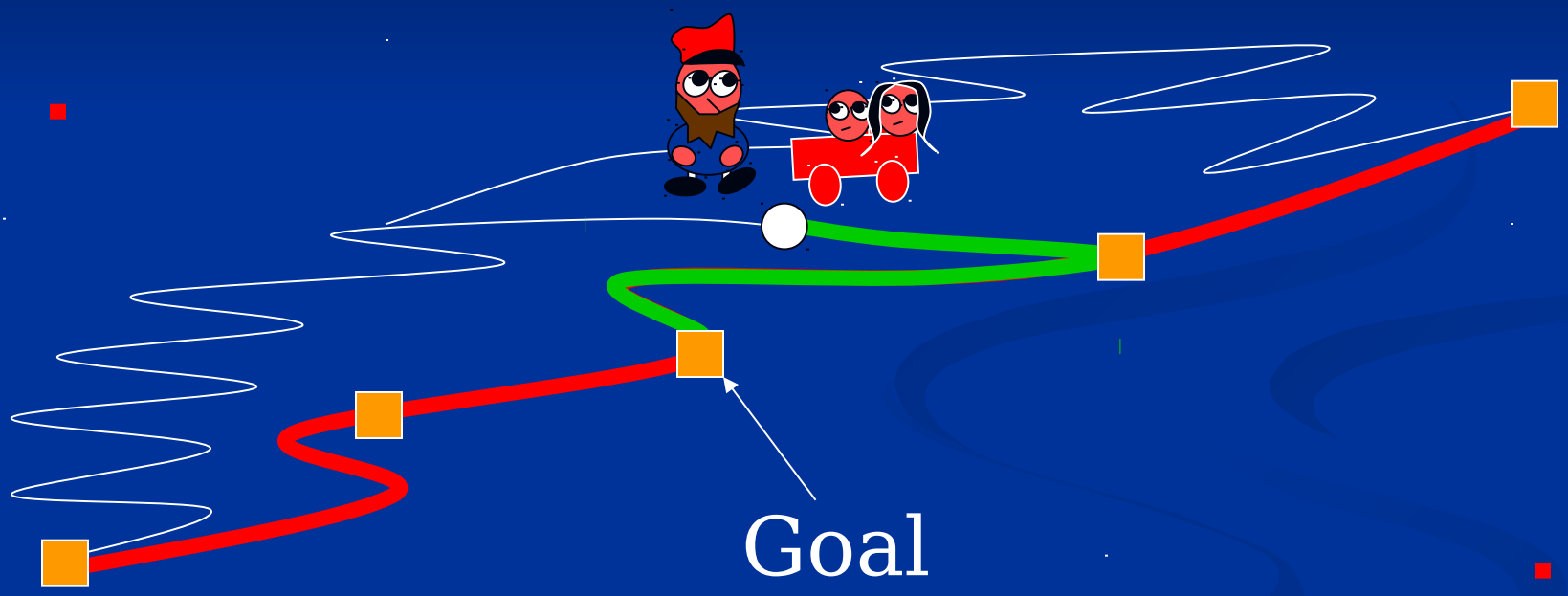
Turn around. The freeway ramp is on



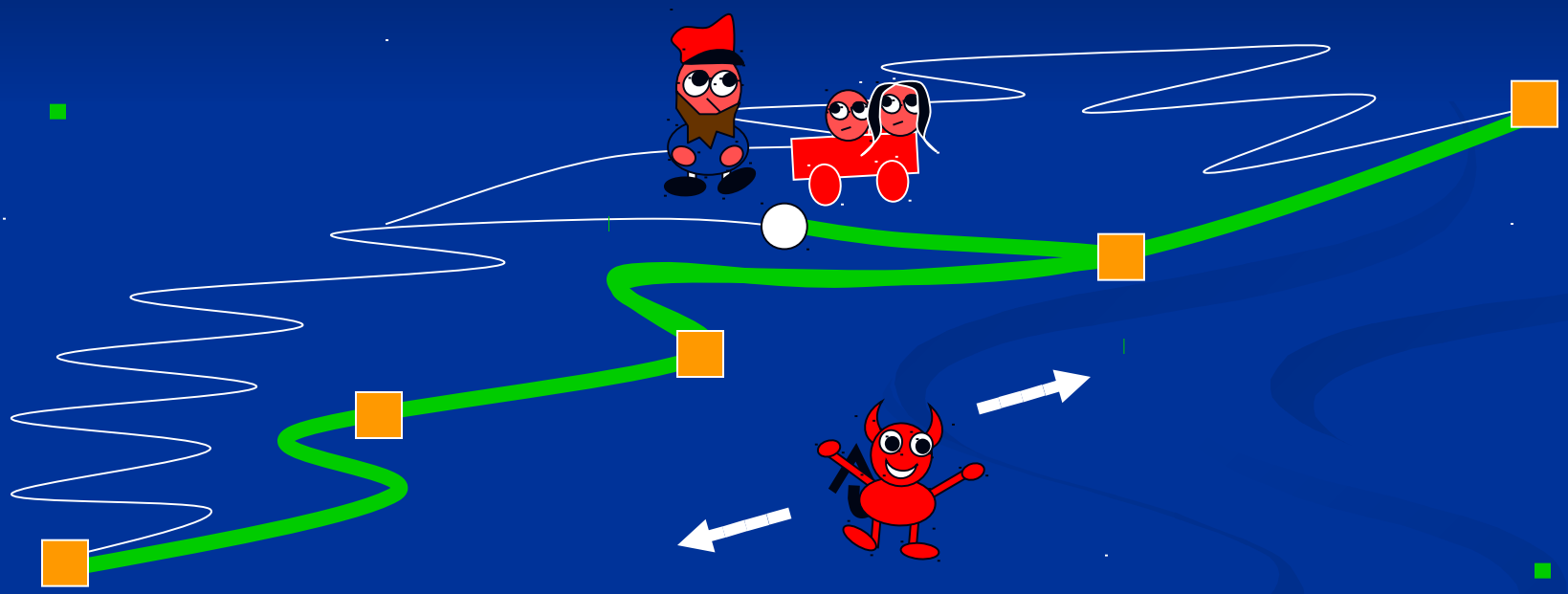
Asking for Directions



Best Route

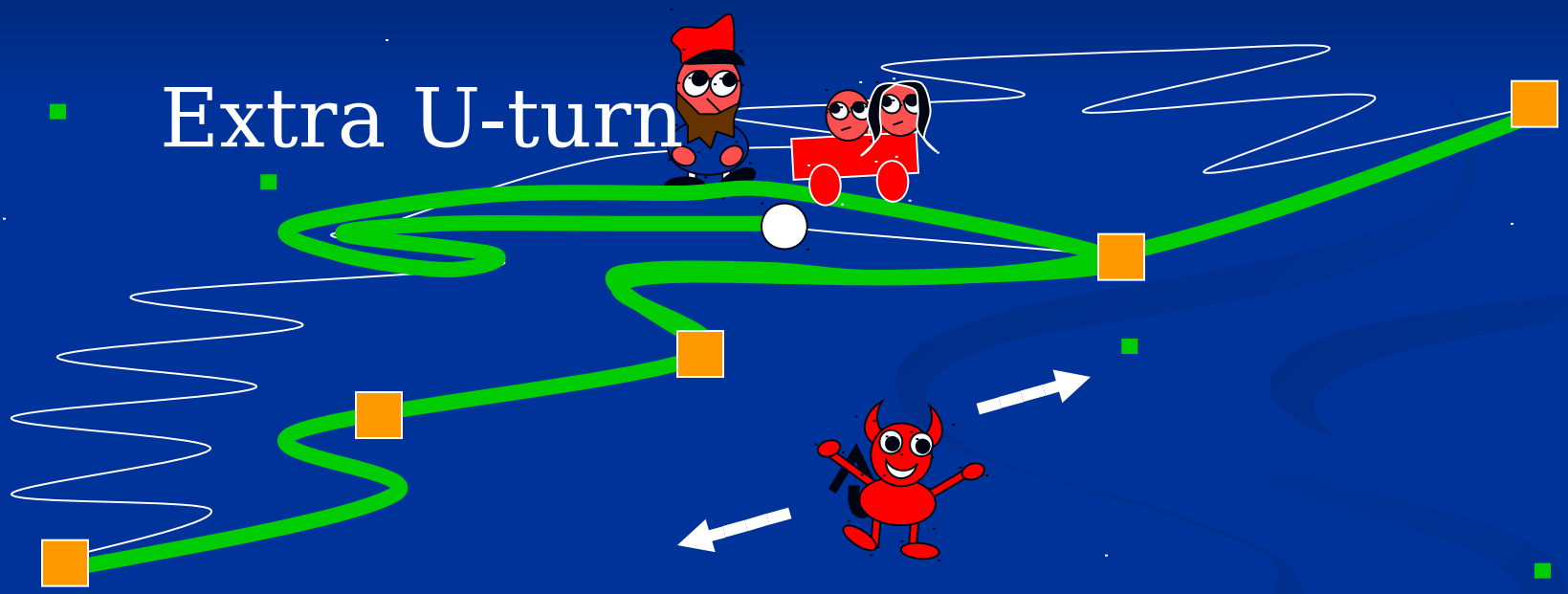


Best Route to Any Goal



Disregarding Advice is Bad

- Extra U-turn



Best Route to Any Goal

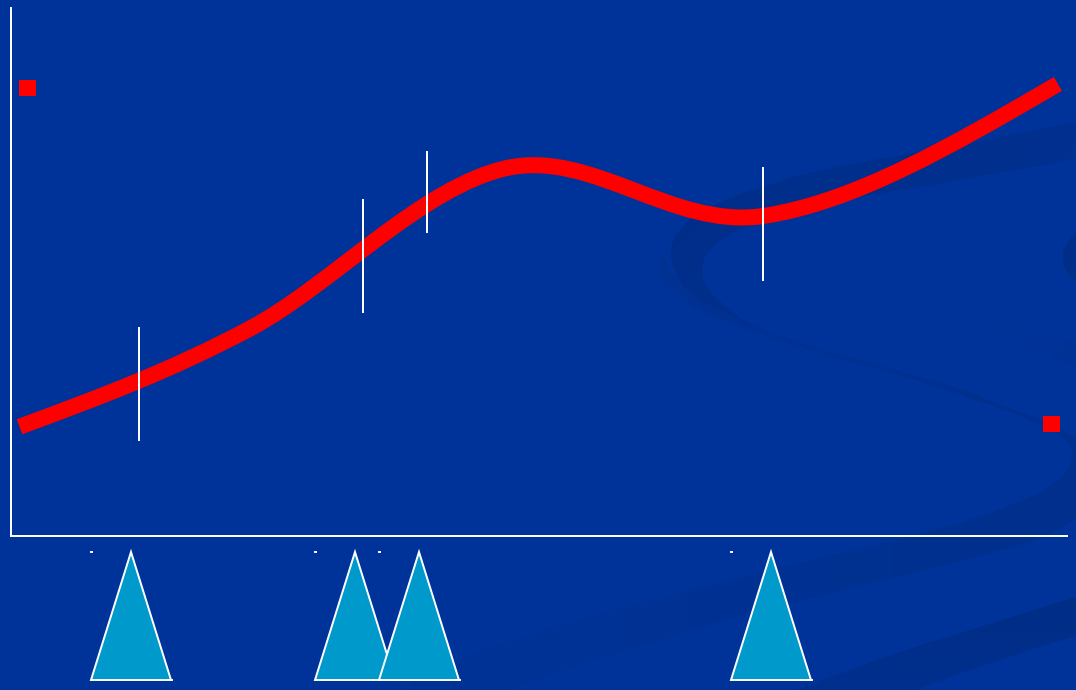
…so fixed advice can help you reach a hidden goal without circles, evasions, or magic

Retraction = Epistemic U-turn

- Choosing T and then failing to choose T next.

Curve Fitting

- Truth = some polynomial $Y = f(X)$
- Data = increasingly precise open intervals around Y at finitely many specified values of X.

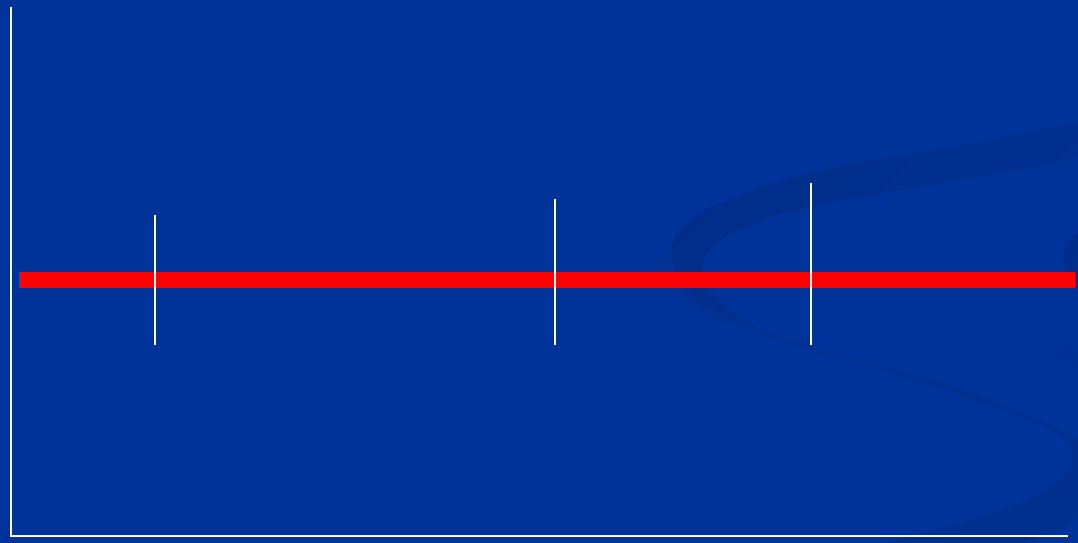


Epistemic Nesting

- **An arbitrary amount of constant data…**

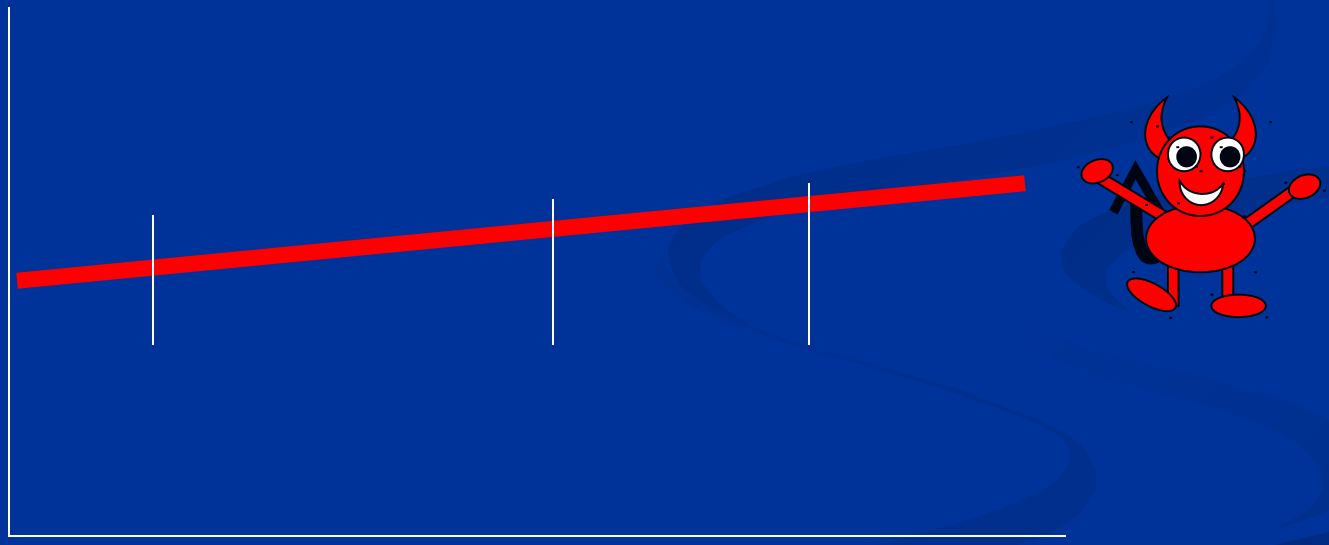
Epistemic Nesting

- **An arbitrary amount of constant data...**



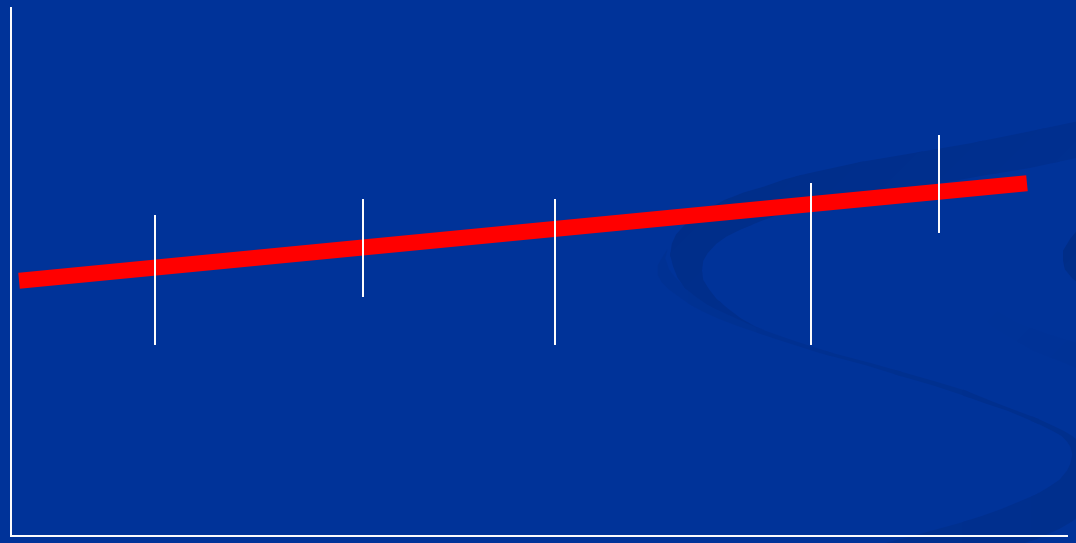
Curve Fitting

- **Is compatible with a linear (non-constant) law.**



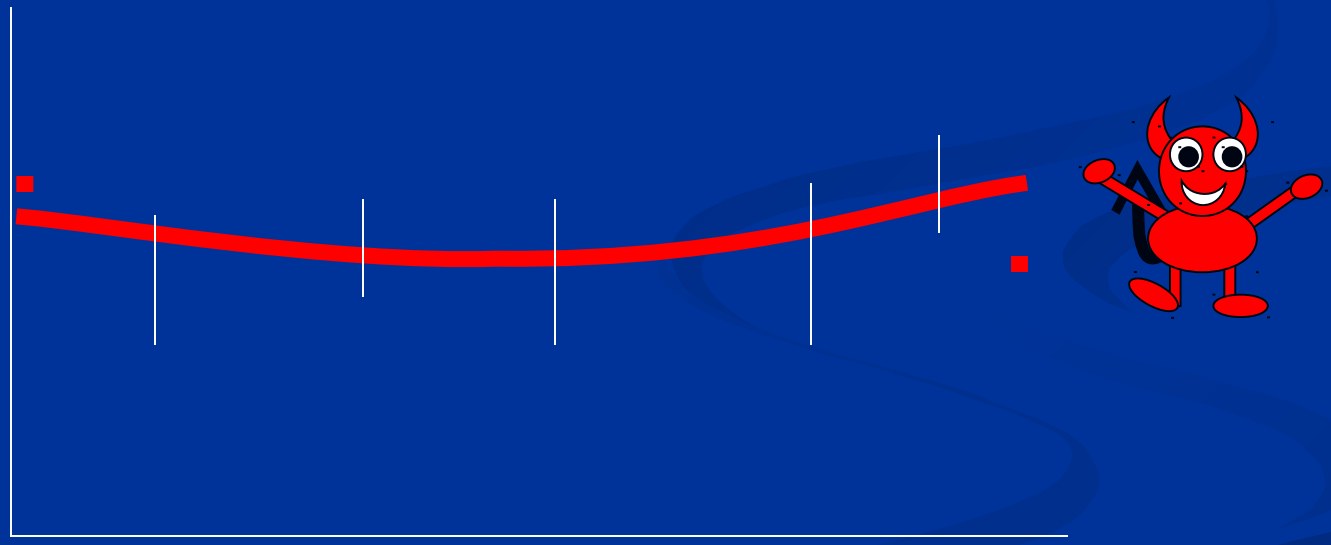
Curve Fitting

- **An arbitrary amount of linear data...**

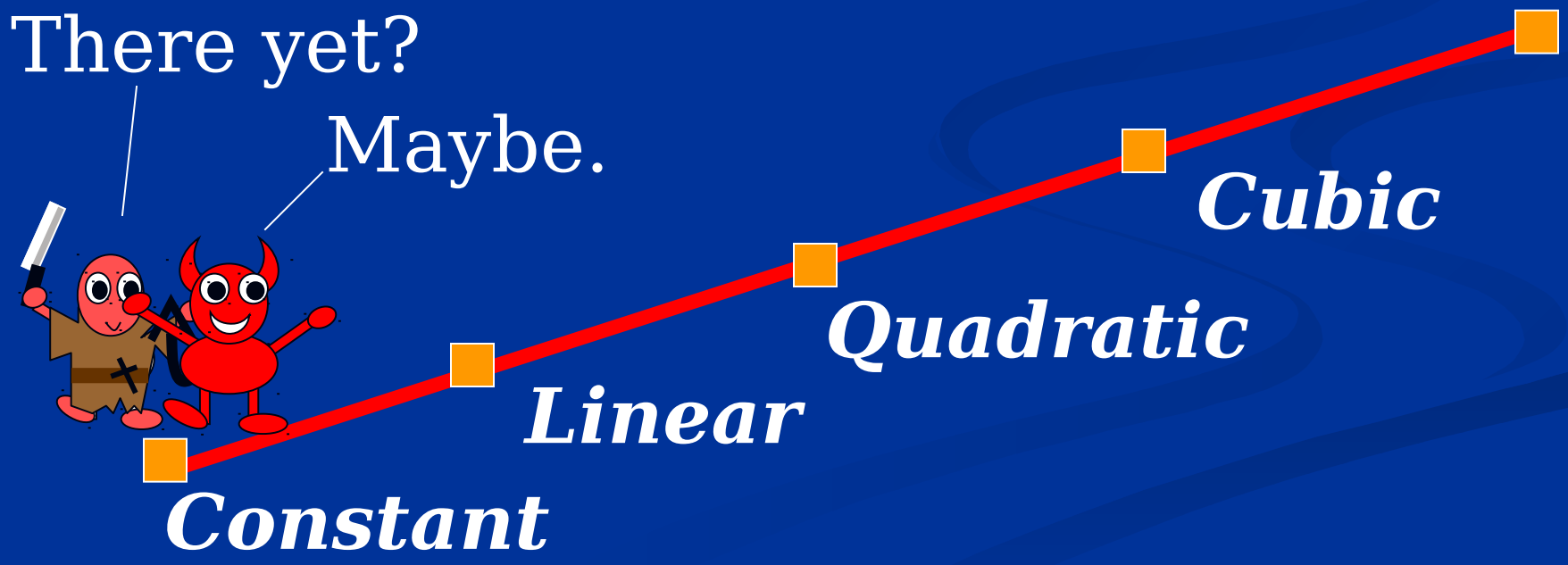


Curve Fitting

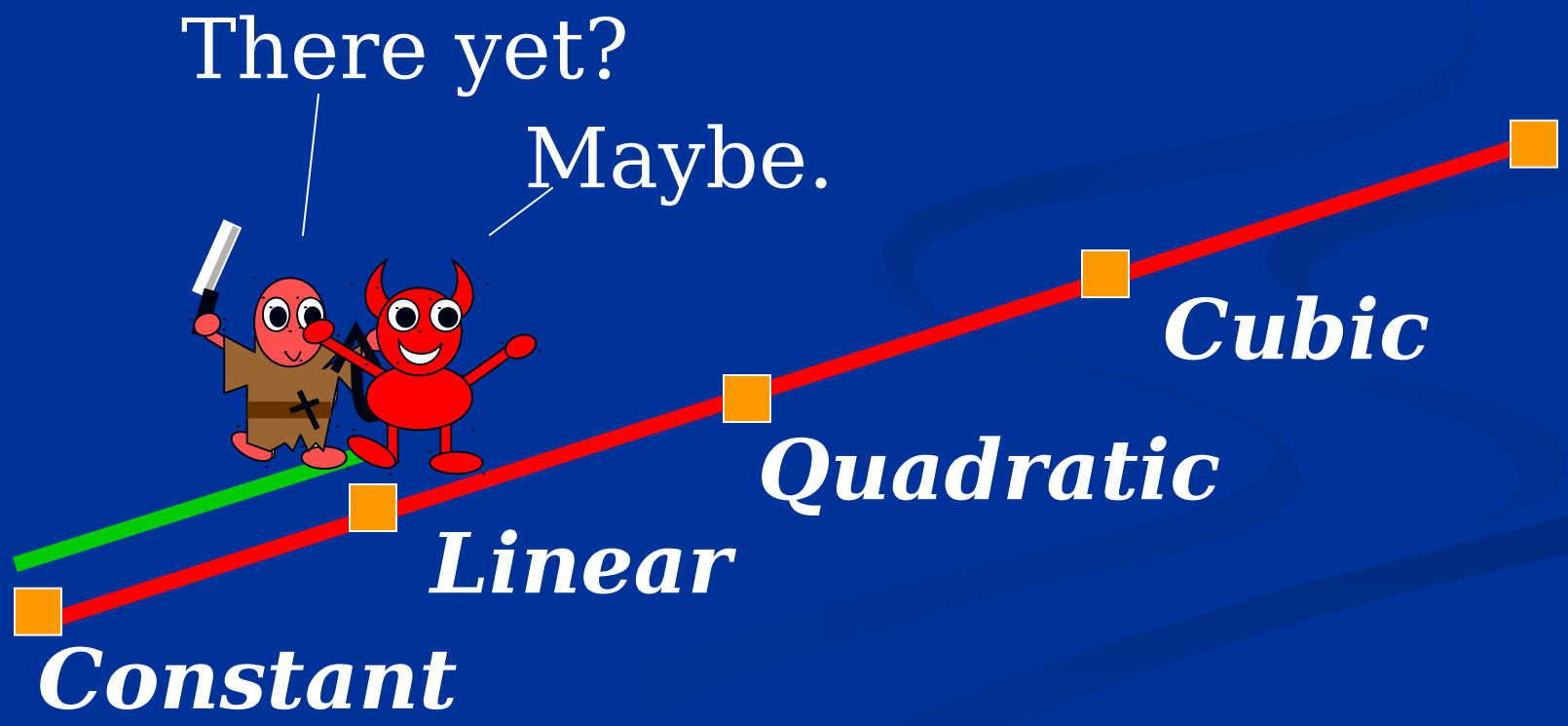
- **...is compatible with a quadratic law, etc.**



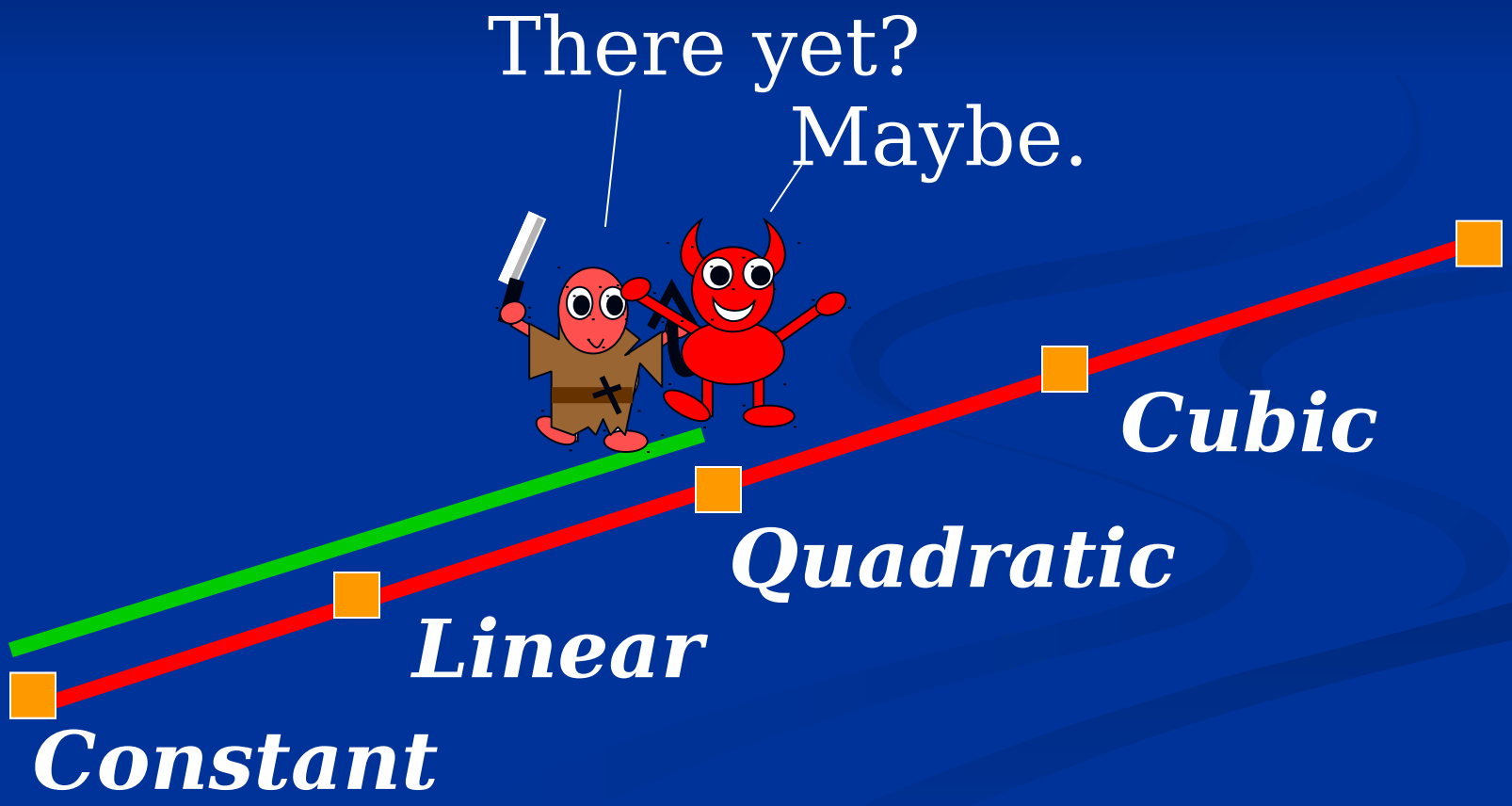
In Step with the Demon



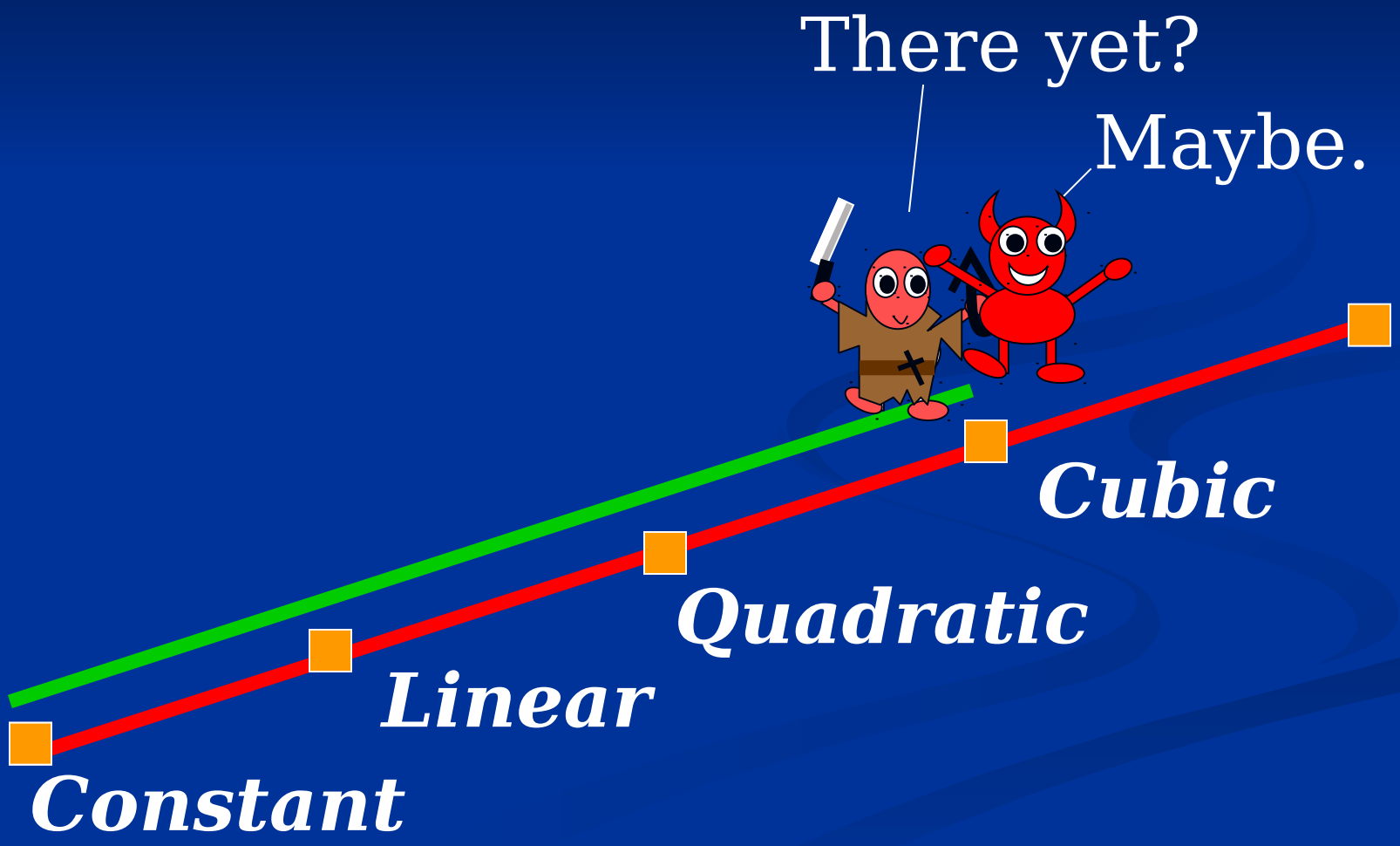
In Step with the Demon



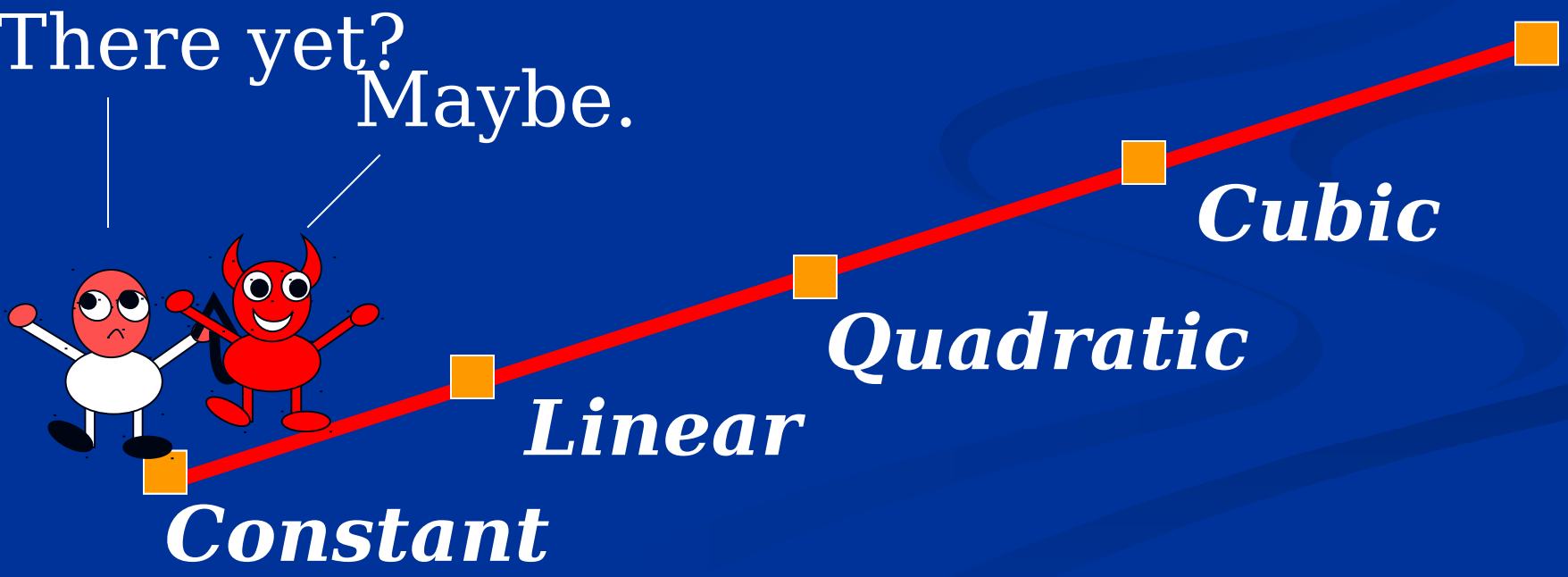
In Step with the Demon



In Step with the Demon

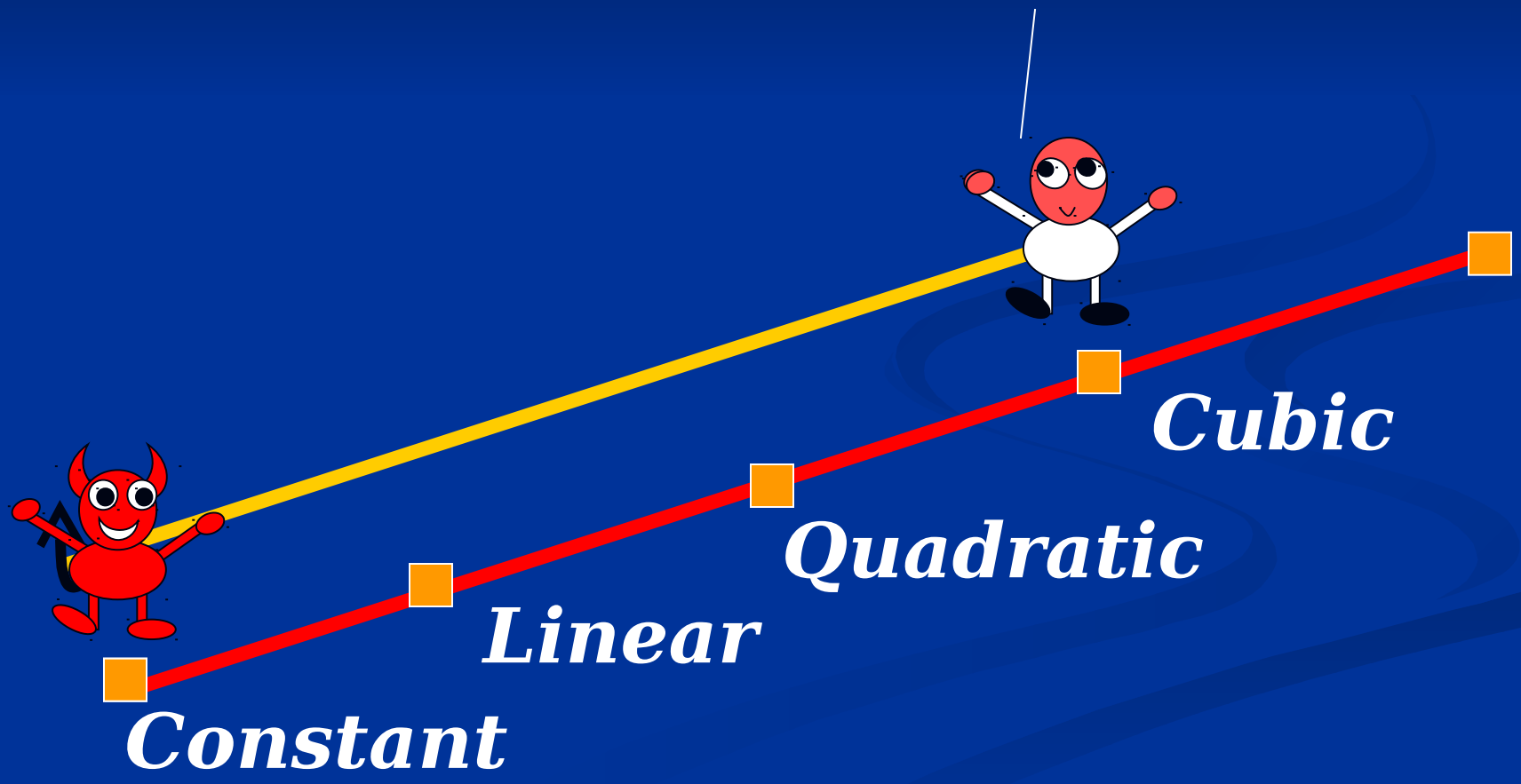


Ahead of the Demon

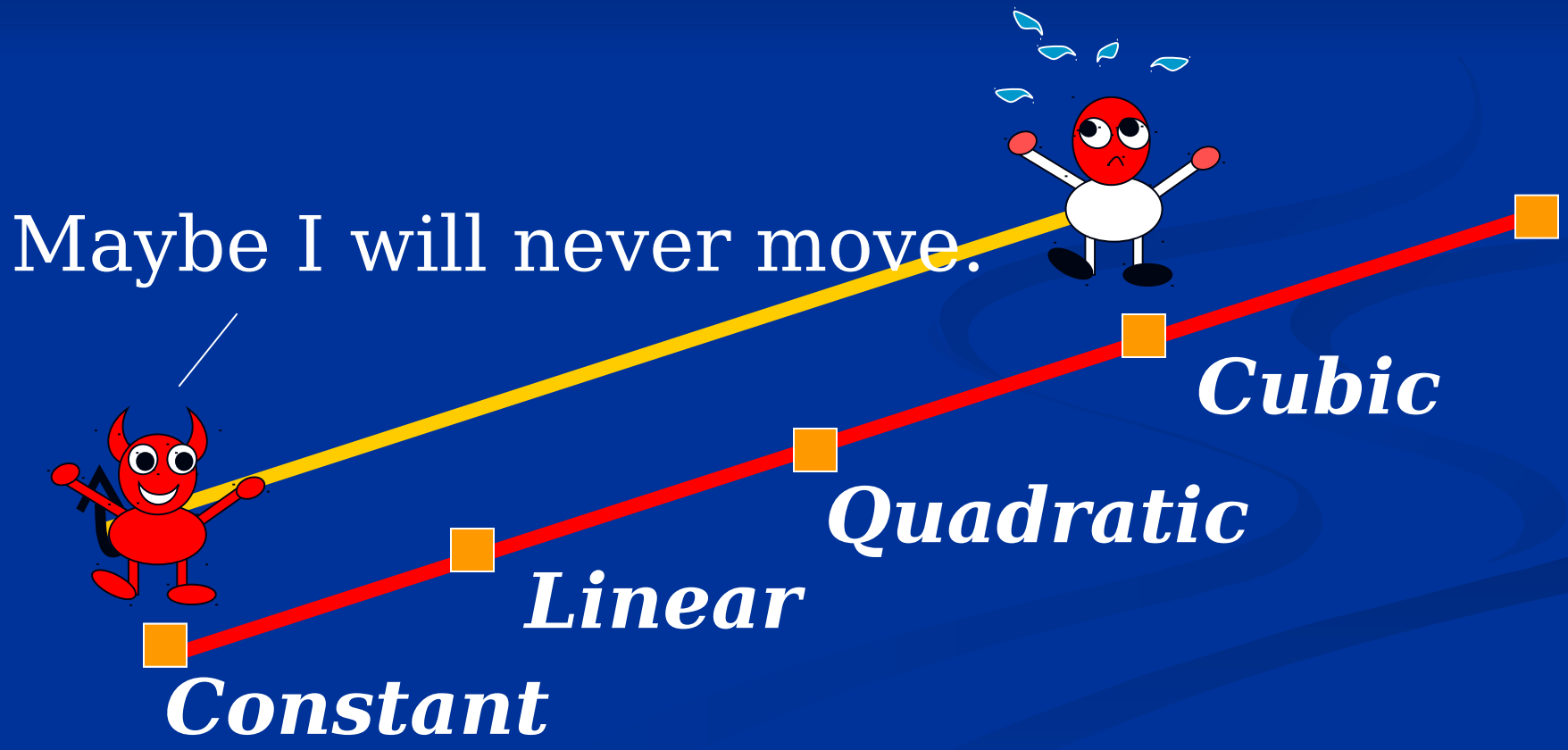


Ahead of the Demon

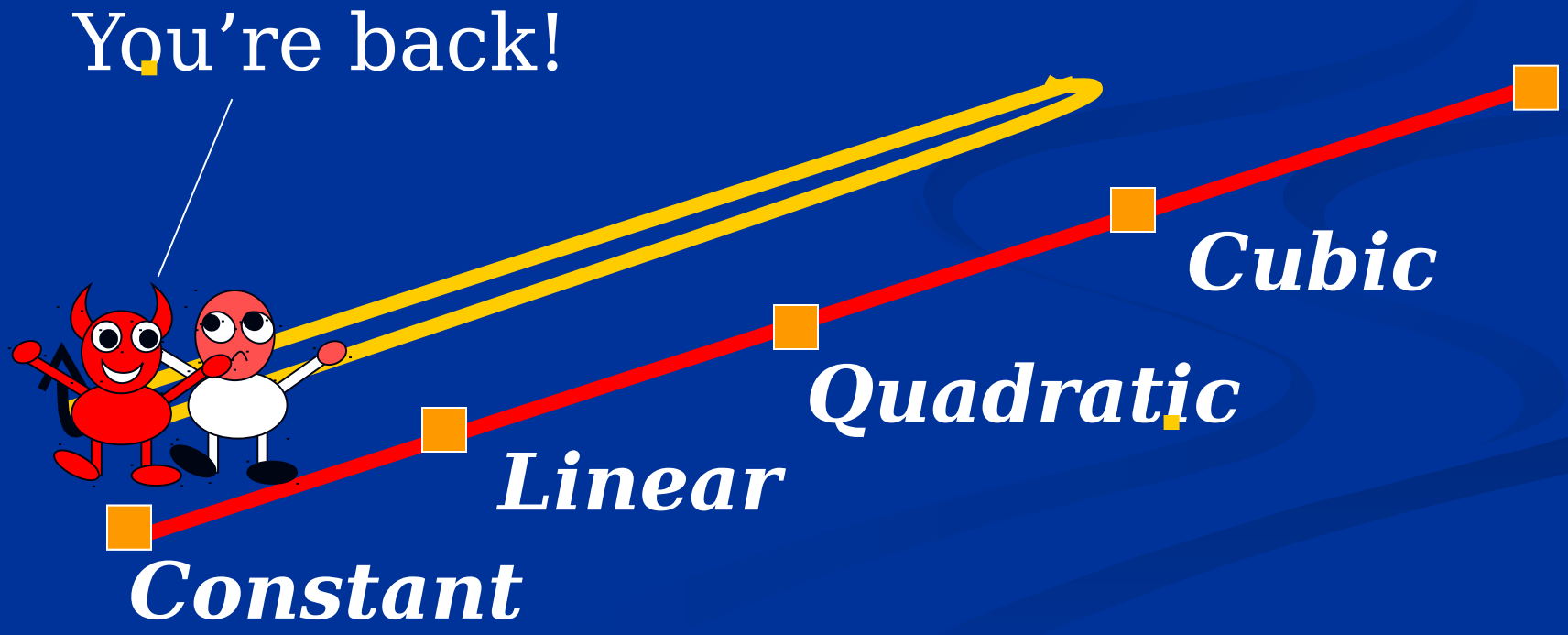
I think you will lead me h



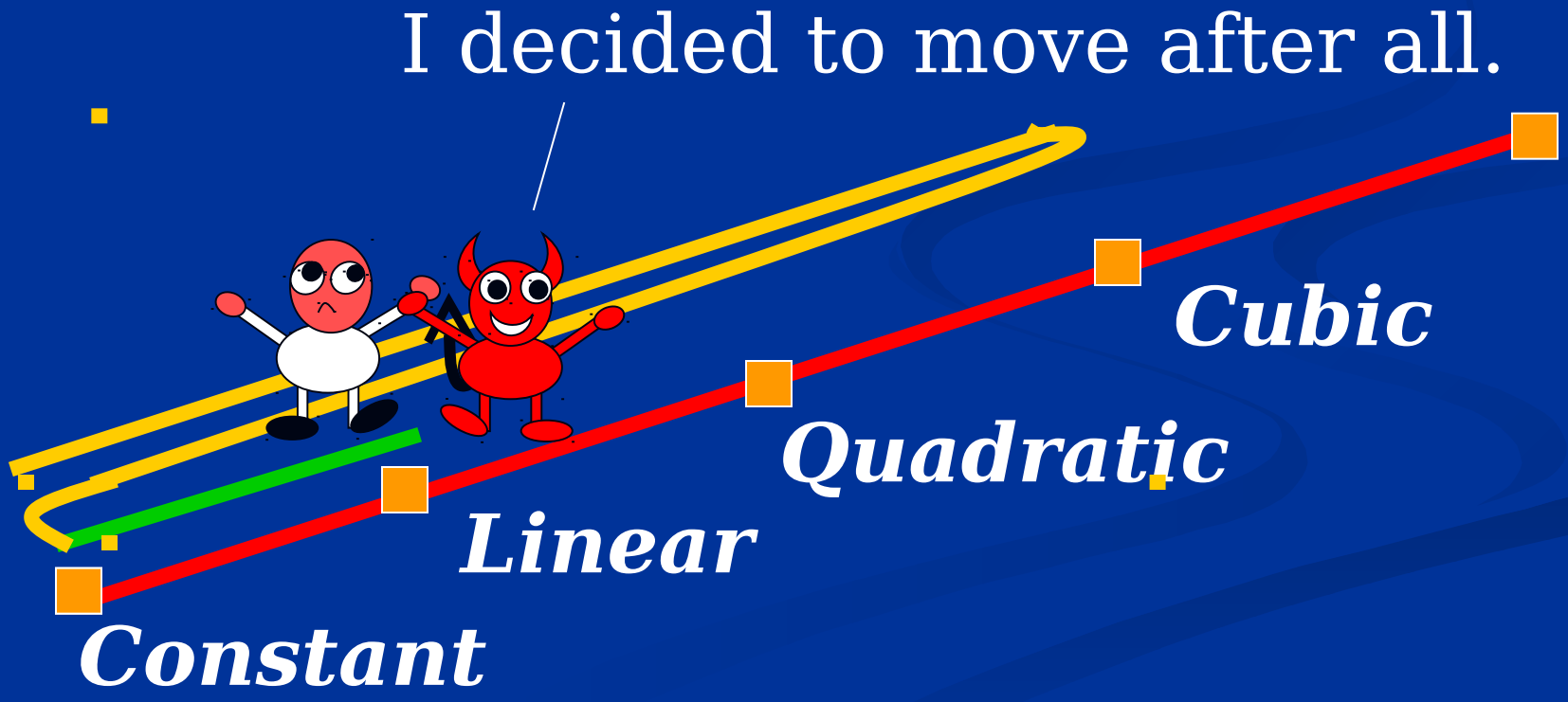
Pressure Builds



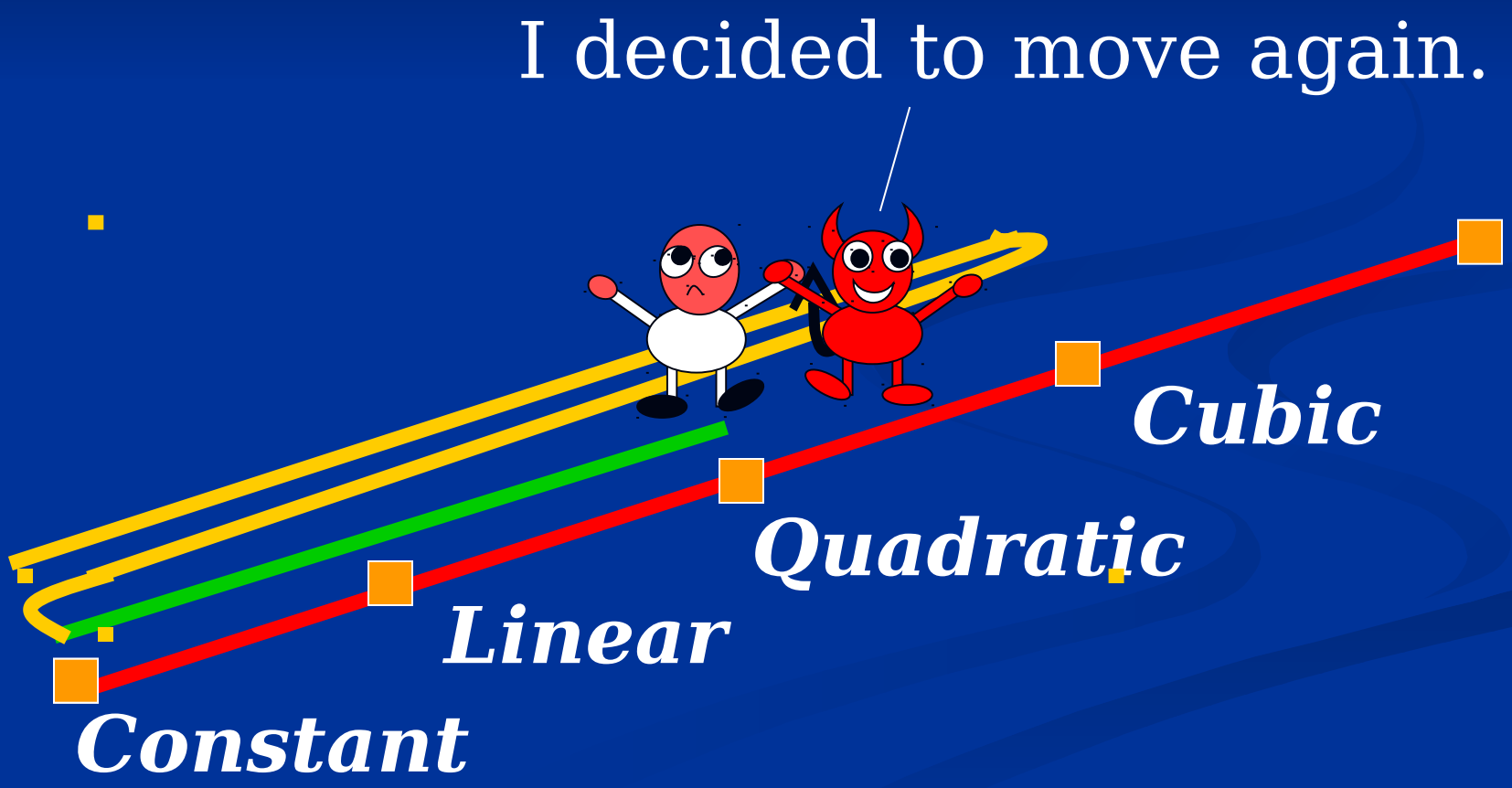
Retraction



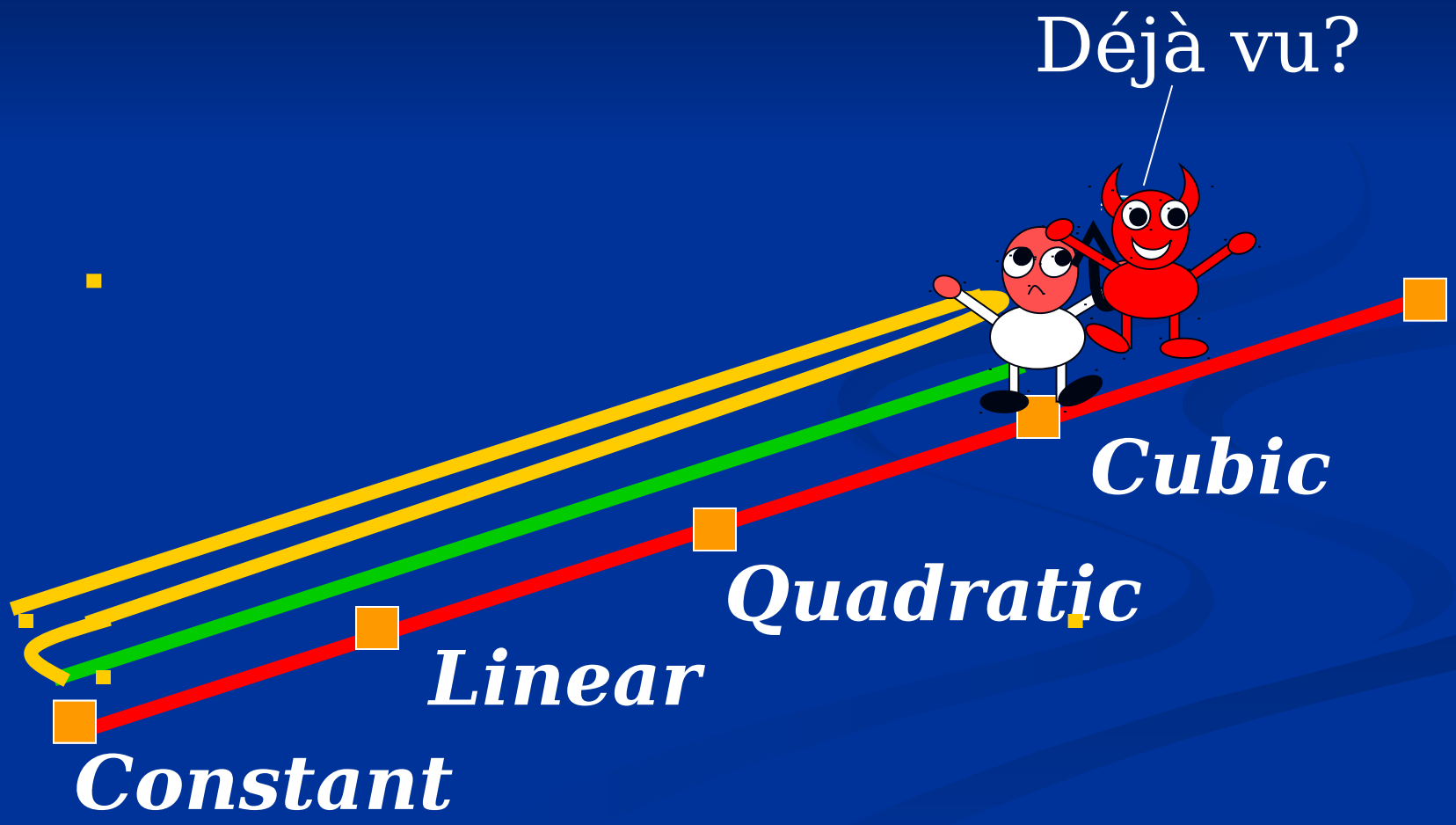
Another Retraction



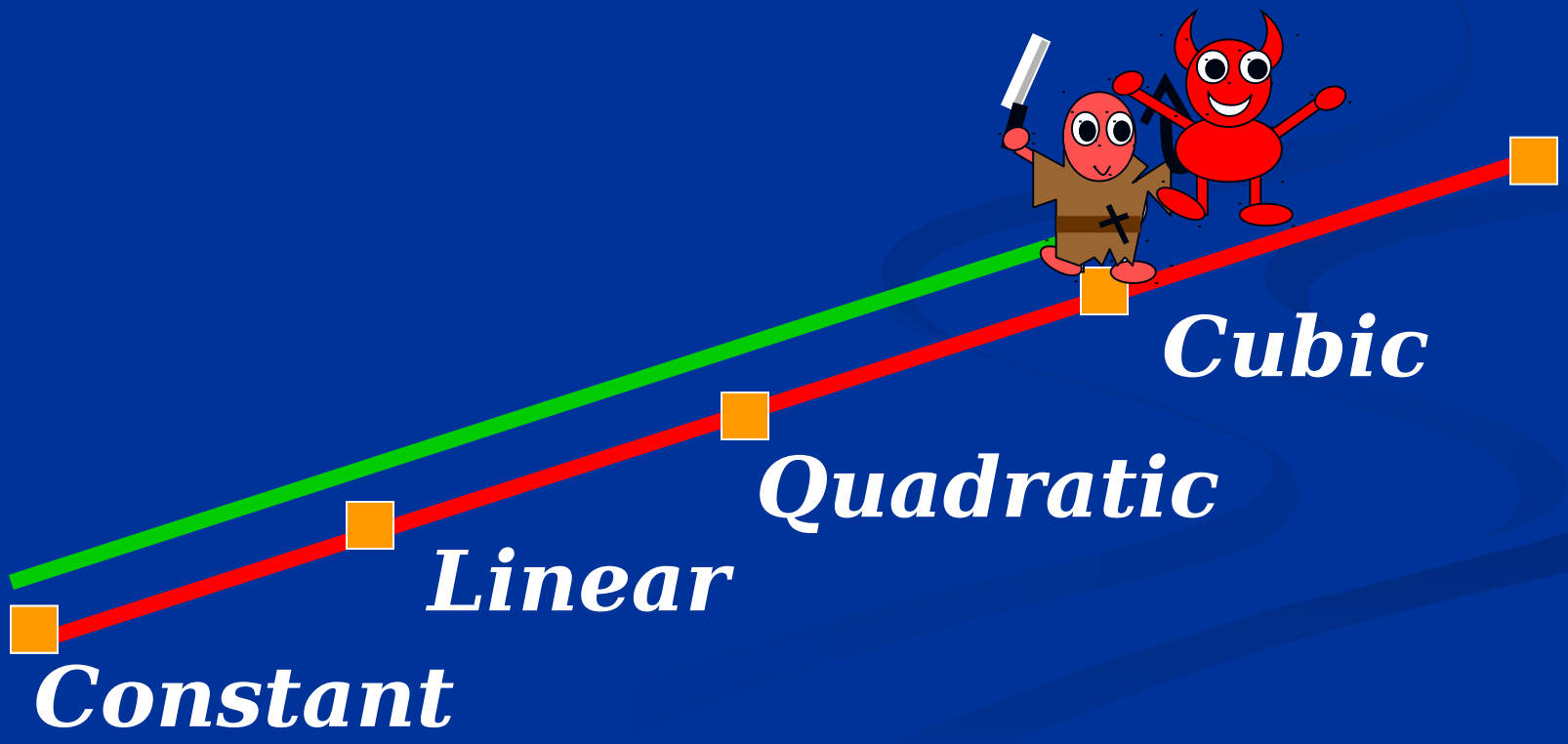
Another Retraction



Another Retraction



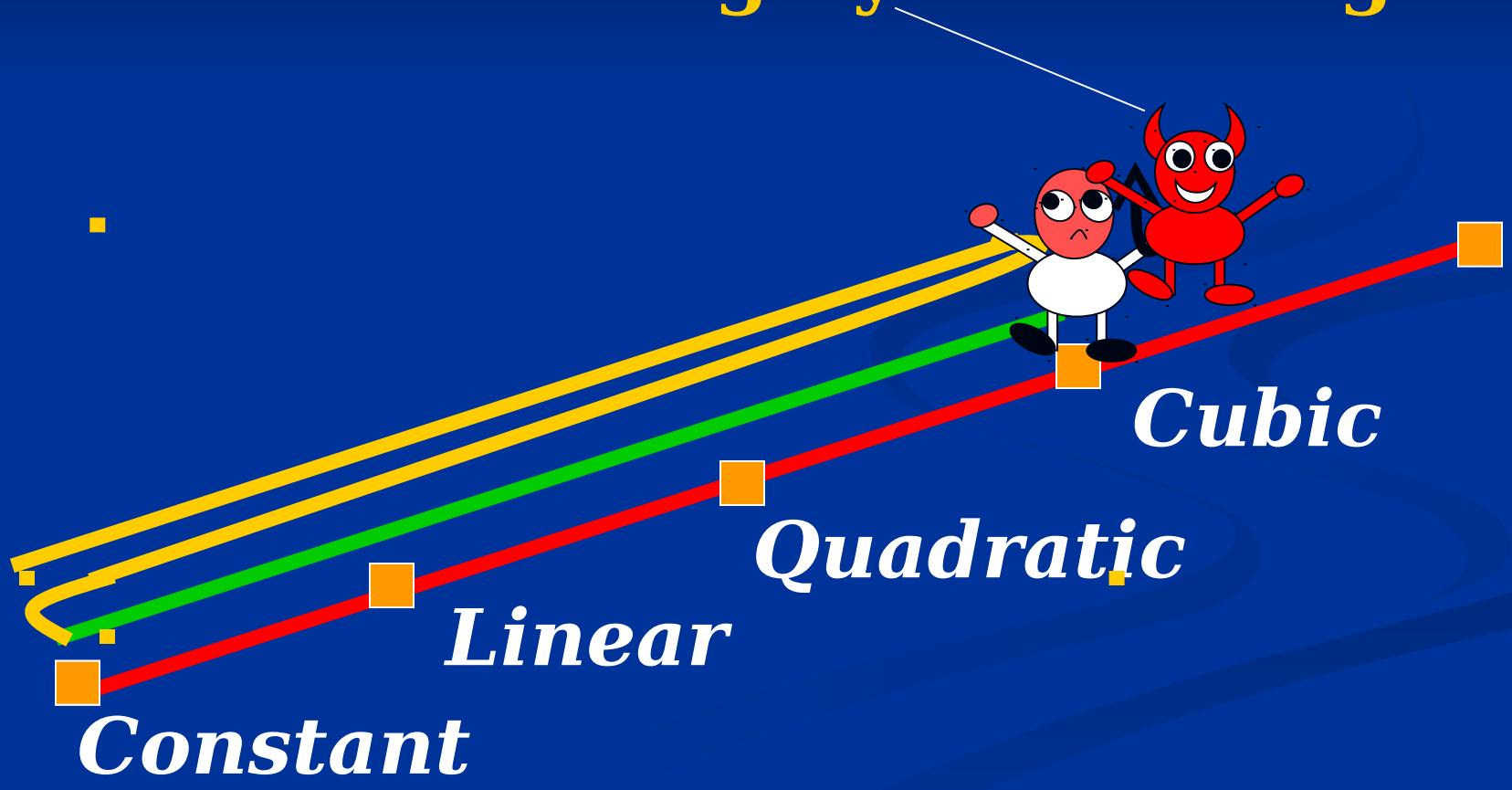
Ockham Path



Ockham Violator's Path

See, you shouldn't have run ahea

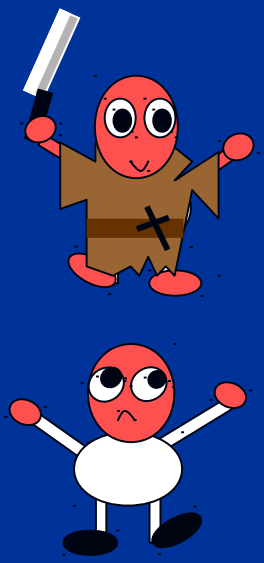
even though you were right!

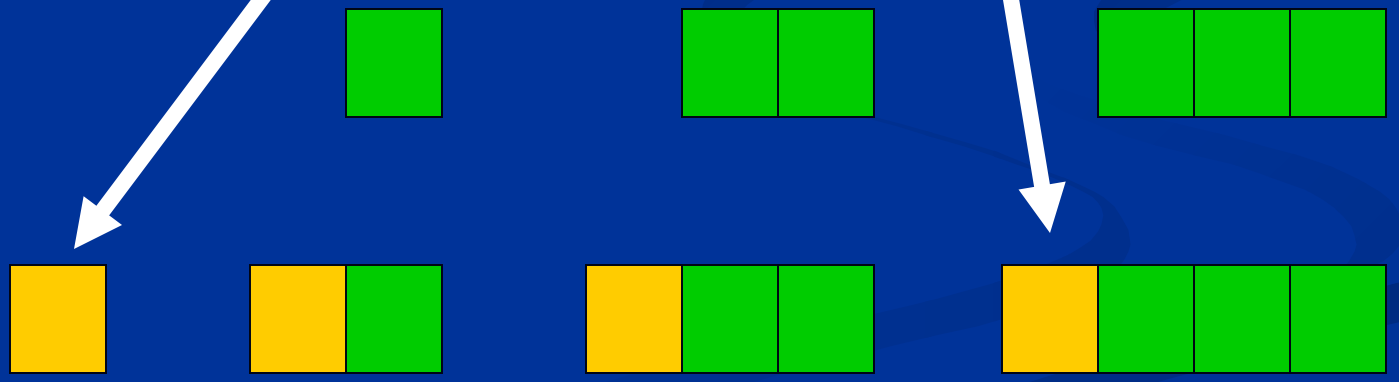


Worst-case Retraction Bounds for Ockham and Violator

Worse performanc
Even in the comple
class favored by th
Ockham violator

Extra first retraction

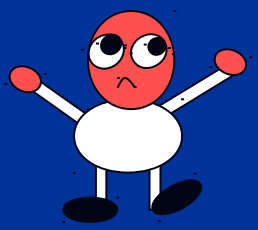




Constant *Linear* *Quadratic* *Cubic*

Ockham Efficiency Theorem:

- Ockham's razor is necessary for retraction-efficiency.



Best possible performance

Sub-optimal performance

Constant Linear Quadratic Cubic

IV. Back-and-forth Ockham Efficiency Theorem

Stalwartness Principle

- After selecting an answer, never drop it while it remains uniquely simplest in light of the data.
- Violation yields extra retraction.

Overly-simple Answers

- An overly simple answer counts as an error in simplest worlds that Ockham does not commit.
- Consider total number of retractions and total number of errors as independent costs.
- Make only easy (Pareto) comparisons between vectors of form:

  (total errors, total retractions).

Intermediate Violations

- When inputs *e* have been received, compare the violator only to methods that have always done the same just prior to the end of *e*.

Ockham Efficiency Representation

- **Theorem:** Among the convergent methods, the always stalwart and Ockham methods are exactly the methods that are always jointly efficient in terms of total errors and total retractions.

IV. The Nature of Empirical Simplicity

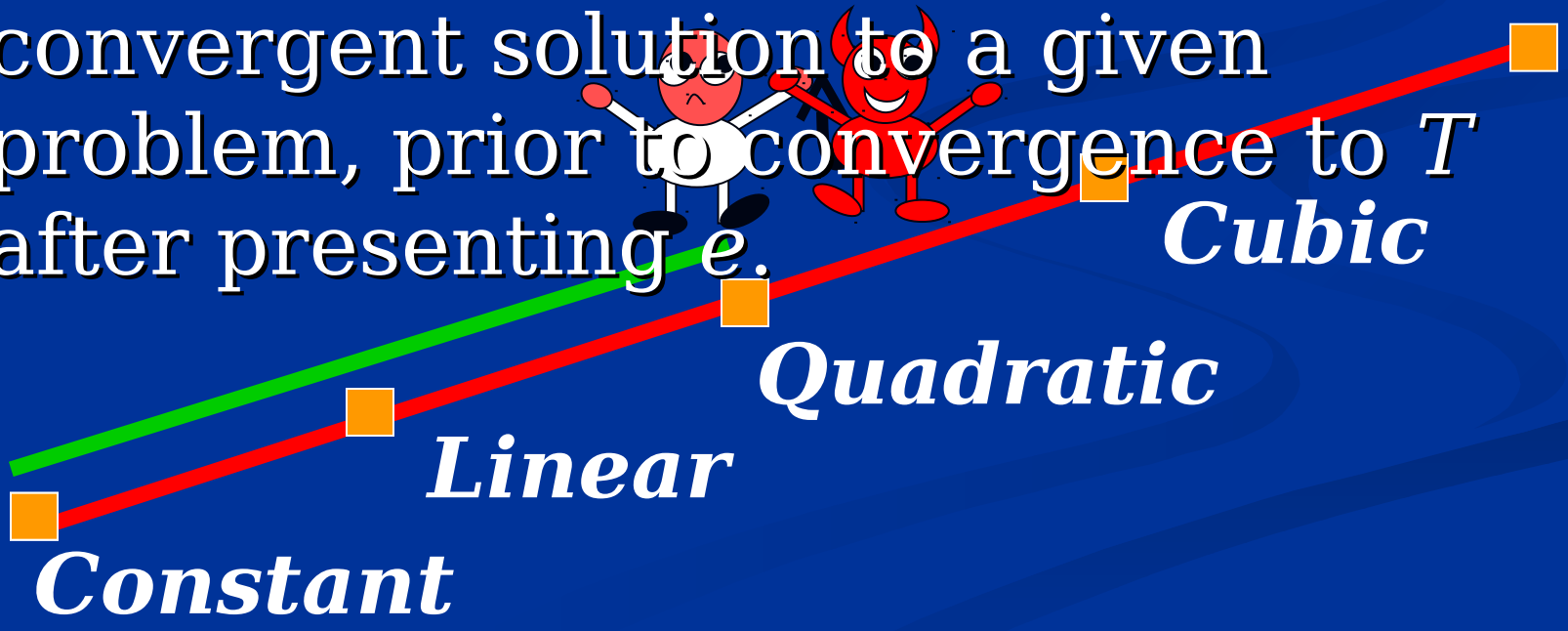
What Empirical Simplicity Is Not

- Syntactic brevity
- Program length
- Dimensionality
- Free parameters
- Fewer causes or entities
- Testability
- Prior probability

What Simplicity Is

- The simplicity of answer T given e in question Q =

  the maximum number of retractions nature can force from an arbitrary, convergent solution to a given problem, prior to convergence to T after presenting e.



Merits

- Depends on problem.
- Grue-proof (topological invariant).
- Immediate epistemological motive.
- Non-trivial for discrete parameter spaces.
- Non-trivial in deterministic

Can be Done Much More Generally

- Branching simplicity degrees.
- Hypotheses that lose and recover their status as uniquely simplest.
- Can be defined for sets of sampling distributions.

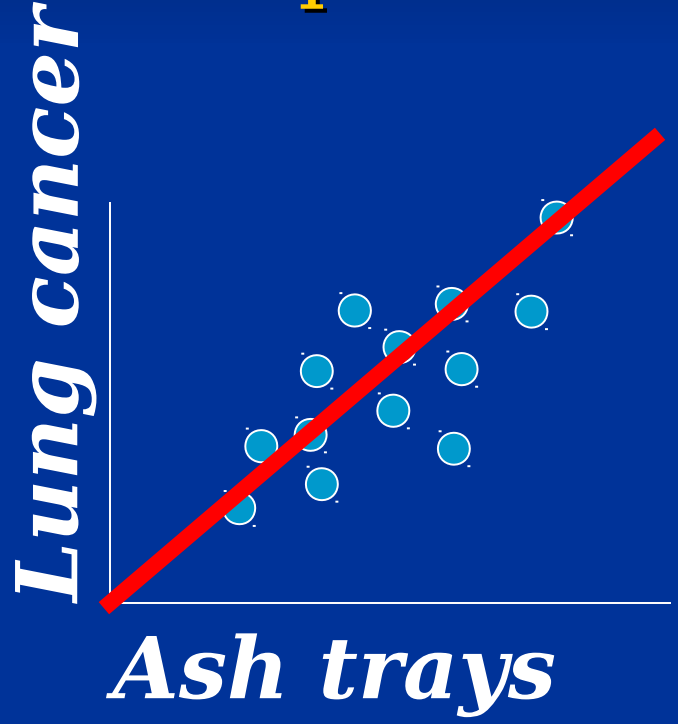
Ockham Efficiency Representation

- Theorem holds for a broad range of problems under the general simplicity definition.

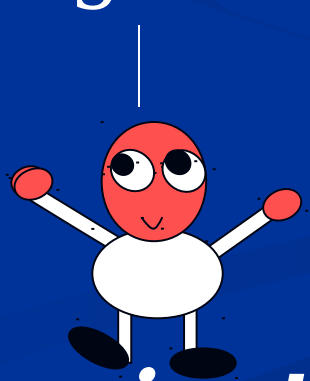
IV. Application: Non-experimental Policy Analysis

Predictive Links

- Correlation or co-dependency allows one to predict *Y* from *X*.



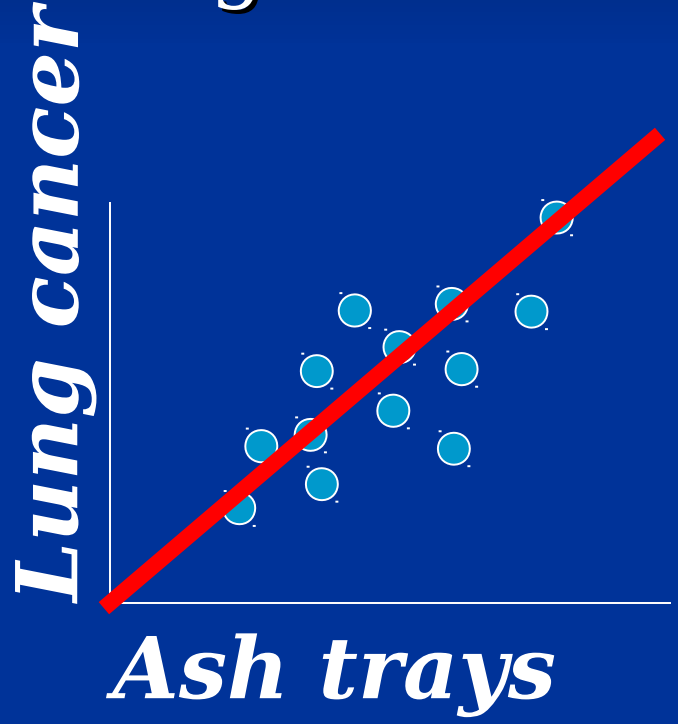
Ash trays
Linked to
Lung cancer!



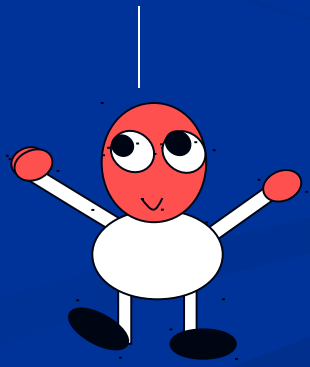
scientist policy make

Policy

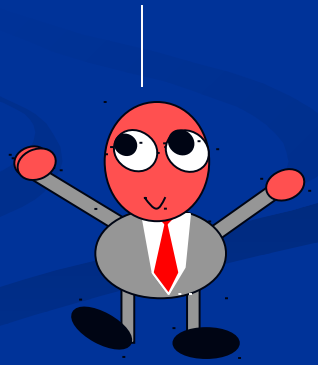
- Policy manipulates *X* to achieve a change in *Y*.



Ash trays Linked to Lung cancer!

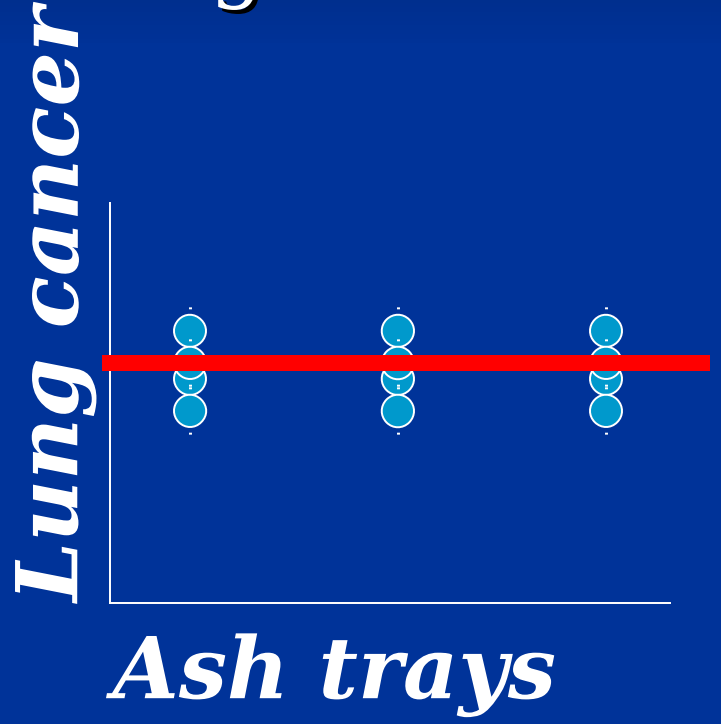


Prohibit ash trays!



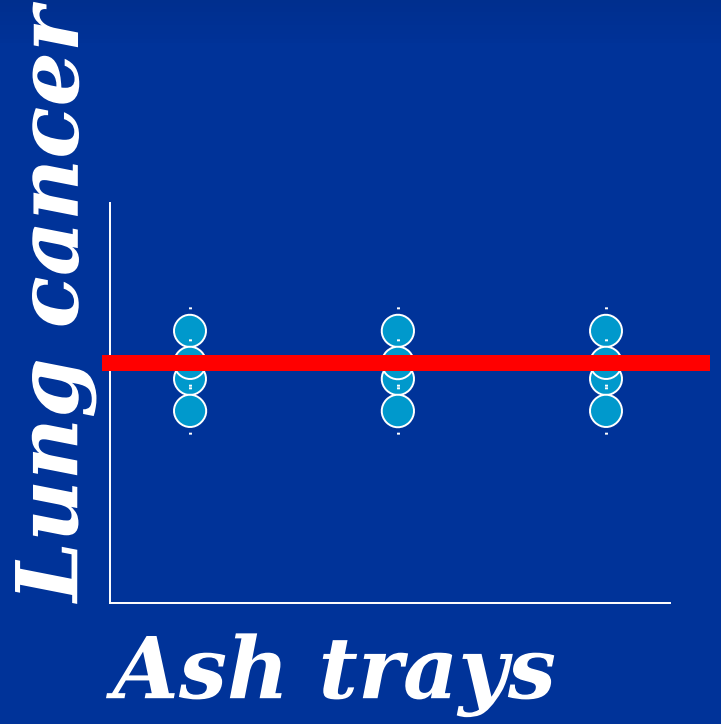
Policy

- Policy manipulates *X* to achieve a change in *Y*.



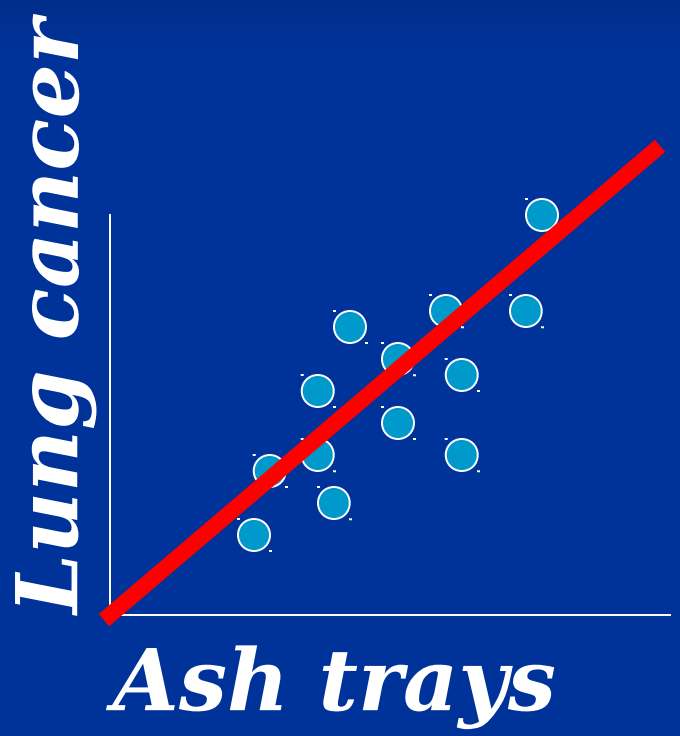
Correlation is not Causation

- Manipulation of X can destroy the correlation of X with Y.

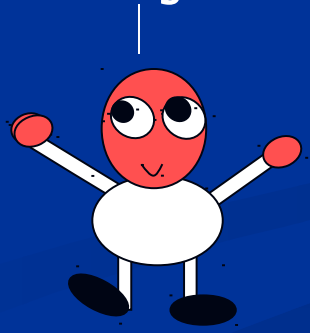


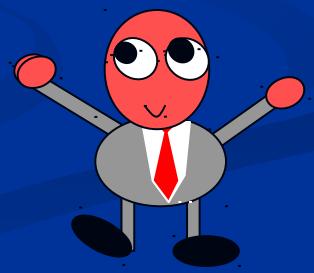
Standard Remedy

- Randomized controlled study



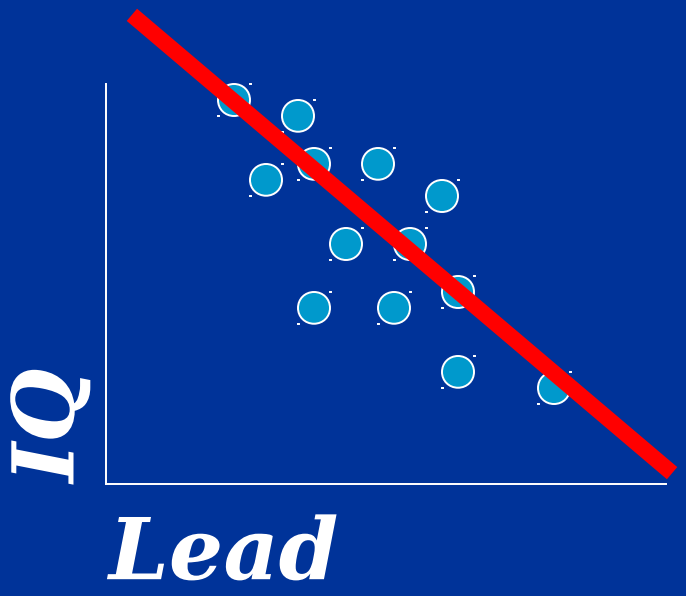
That's what happens if you carry out the policy.



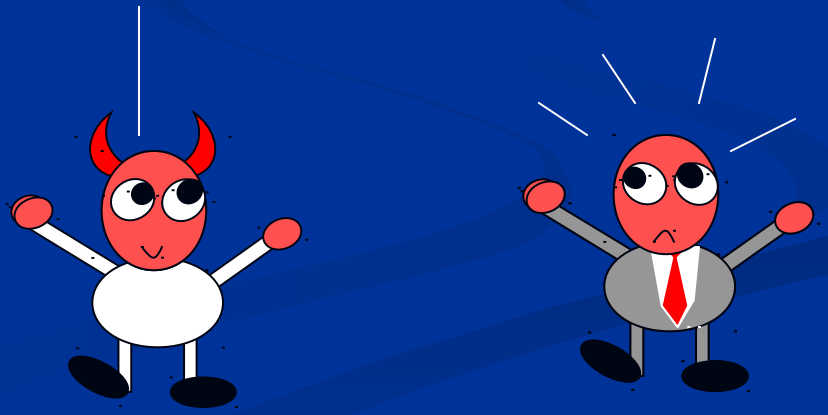


Experimental Infeasibility

- Expense
- Morality

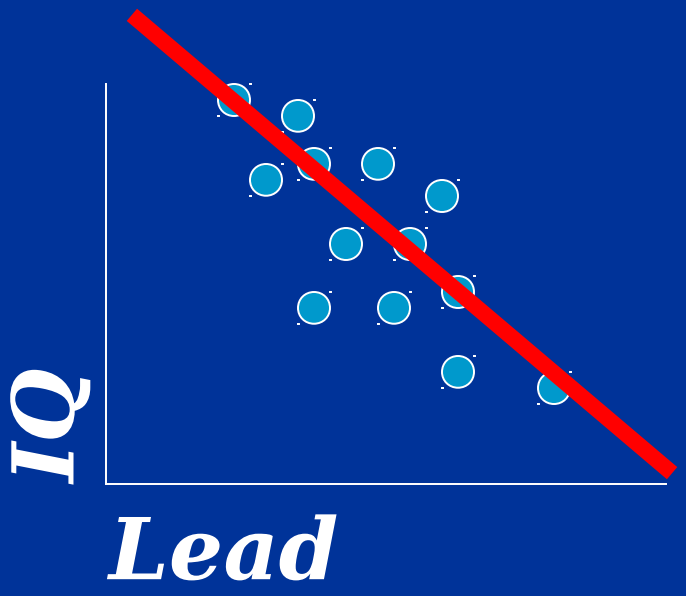


Let me force a few thousand children to eat lead.

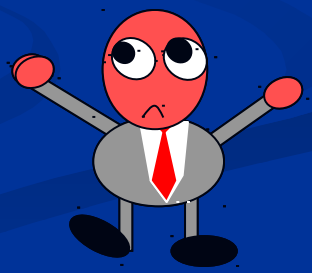


Experimental Infeasibility

- Expense
- Morality

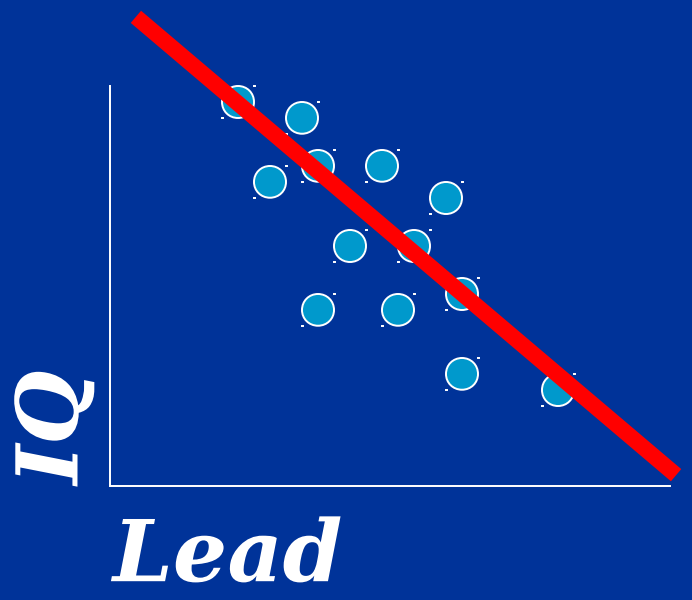


Just joking!



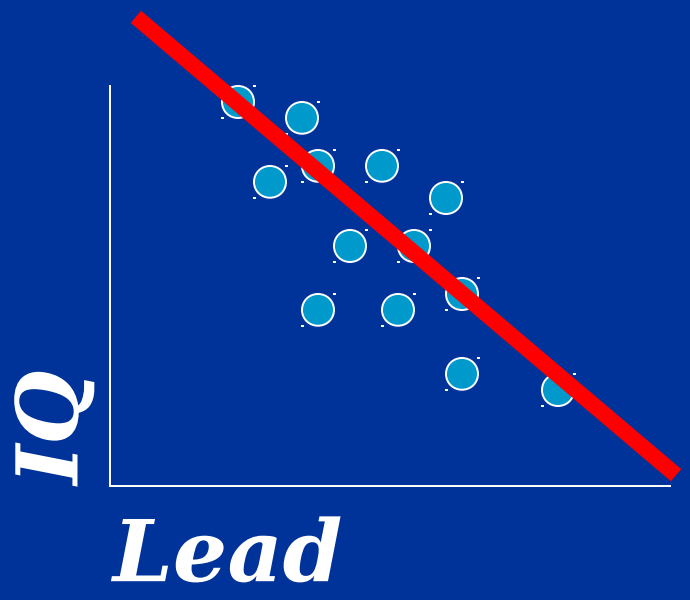
Ironic Alliance

Ha! You will never prove tha
lead affects IQ…



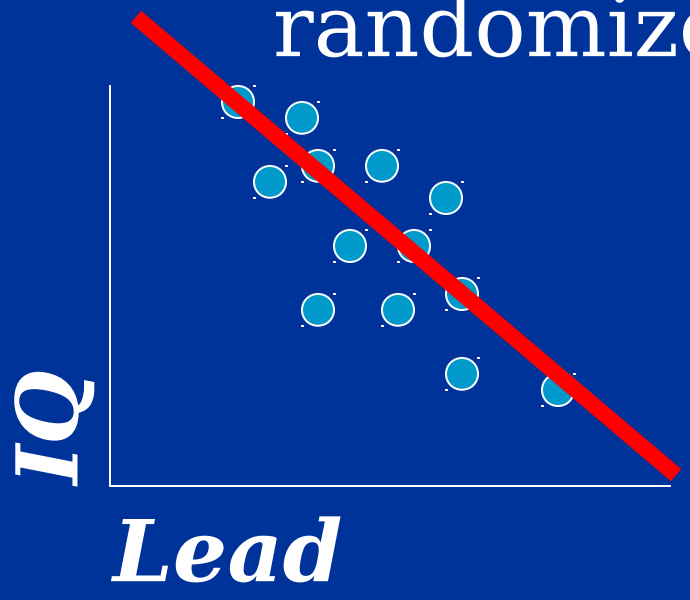
Ironic Alliance

And you can’t eliminate jobs on a mere whim.



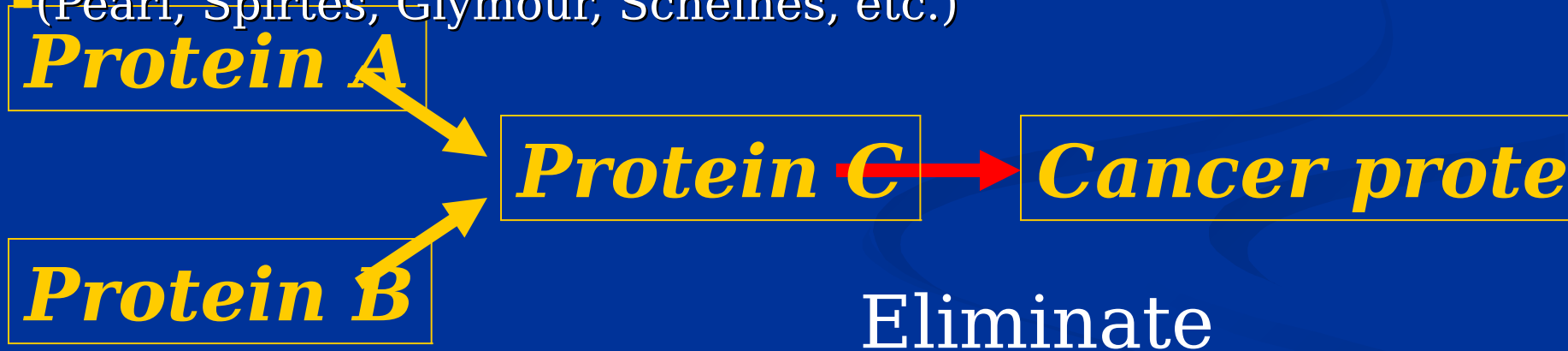
Ironic Alliance

So I will keep on polluting, which will never settle the matter because it is not a randomized trial.

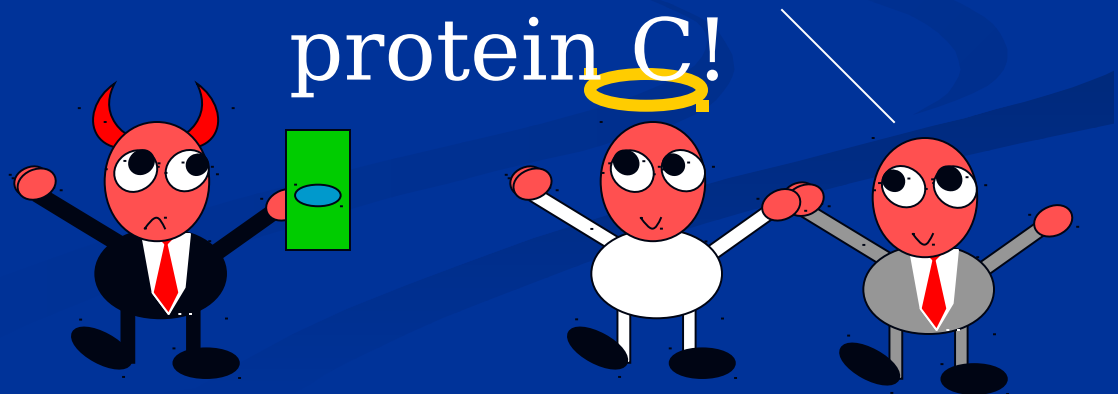


Remedy: Causes from Correlations

- Patterns of conditional correlation can imply unambiguous causal conclusions
  - (Pearl, Spirtes, Glymour, Scheines, etc.)

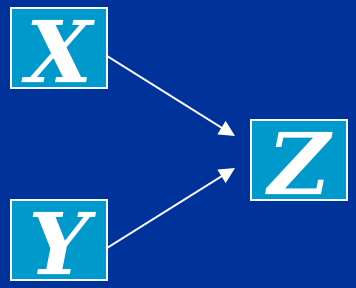


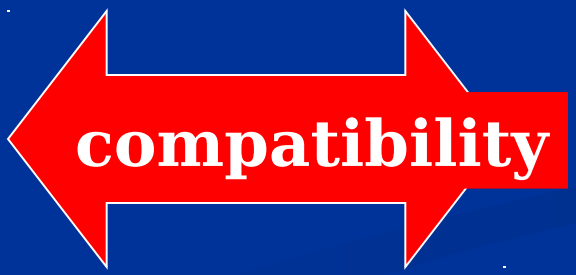
Eliminate protein C!

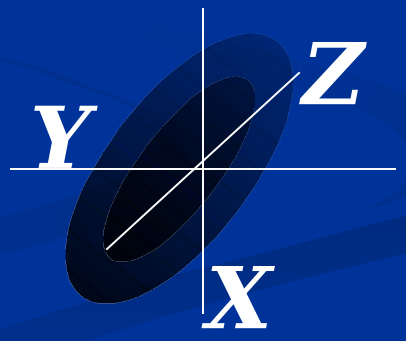


Basic Idea

- Causation is a directed, acyclic network over variables.
- What makes a network causal is a relation of compatibility between networks and joint probability distributions.



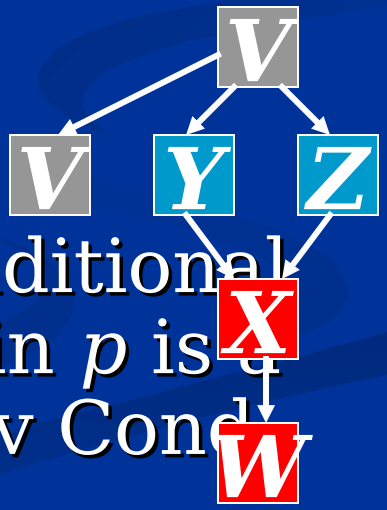




Compatibility

Joint distribution p is compatible with directed, acyclic network G iff:

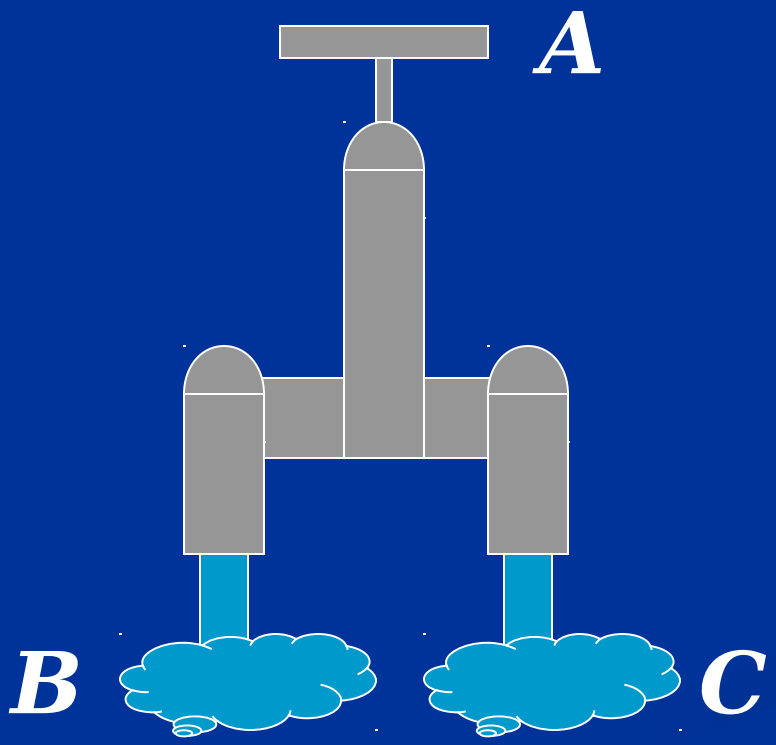
- Causal Markov Condition: each variable X is independent of its non-effects given its immediate causes.

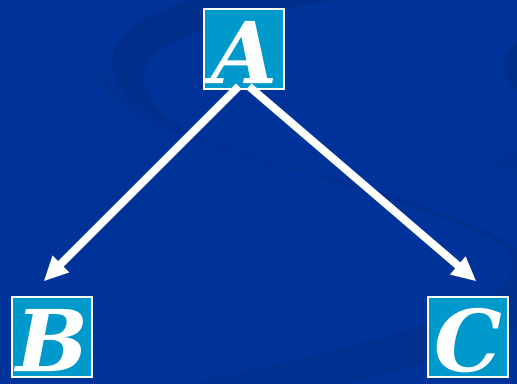


- Faithfulness Condition: every conditional independence relation that holds in p is a consequence of the Causal Markov Cond

Common Cause

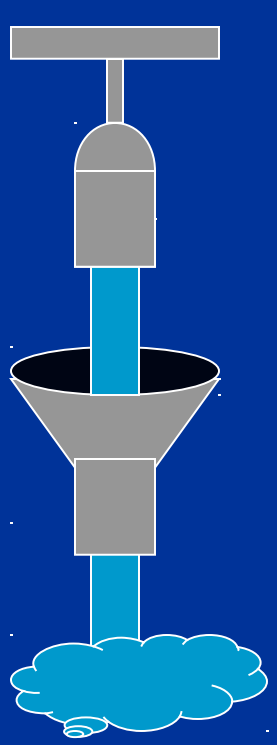
- *B* yields info about *C* (Faithfulness);
- *B* yields no further info about *C* given *A* (M





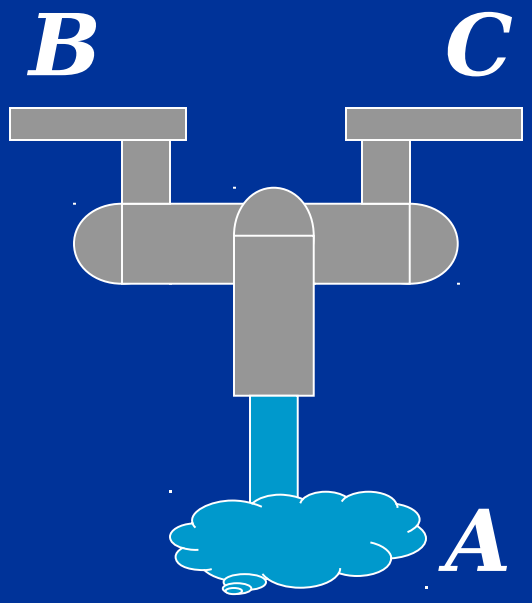
Causal Chain

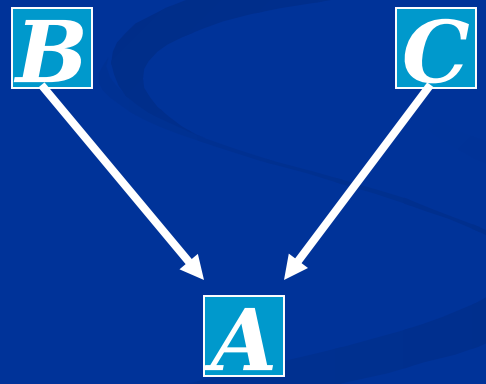
- *B* yields info about *C* (Faithfulness);
- *B* yields no further info about *C* given *A* (M



Common Effect

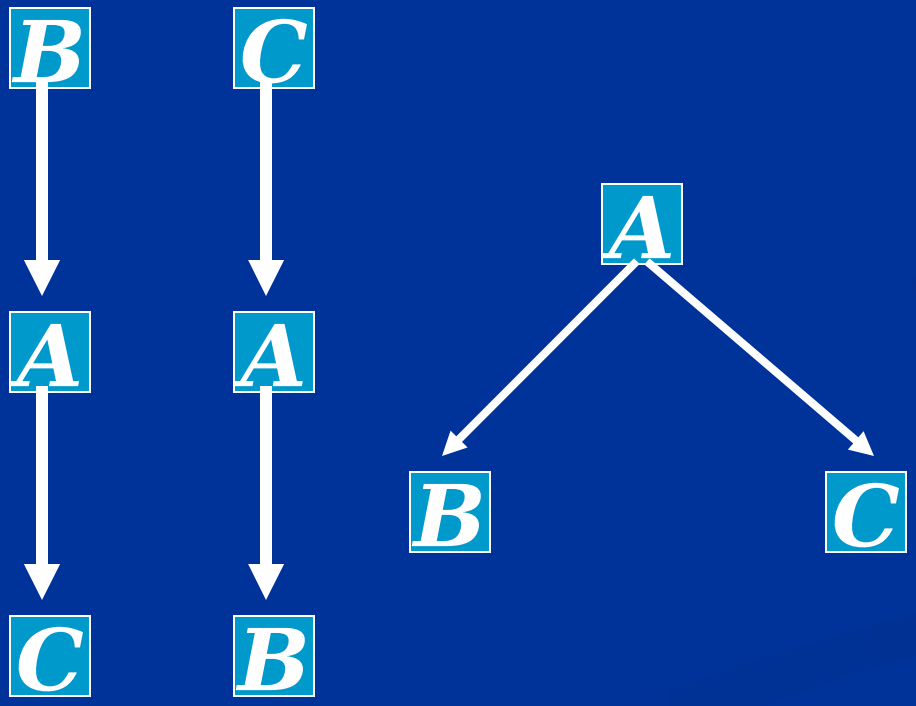
- B yields no info about C (Markov);
- B yields extra info about C given A (Faithfulness).



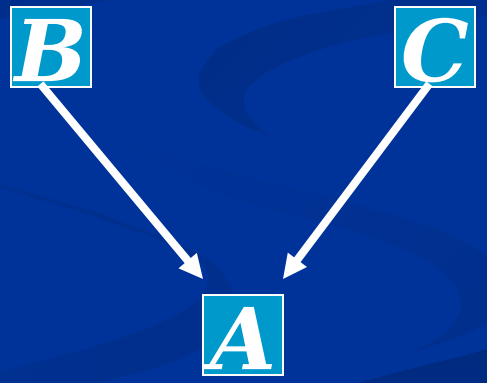


Distinguishability

indistinguishable

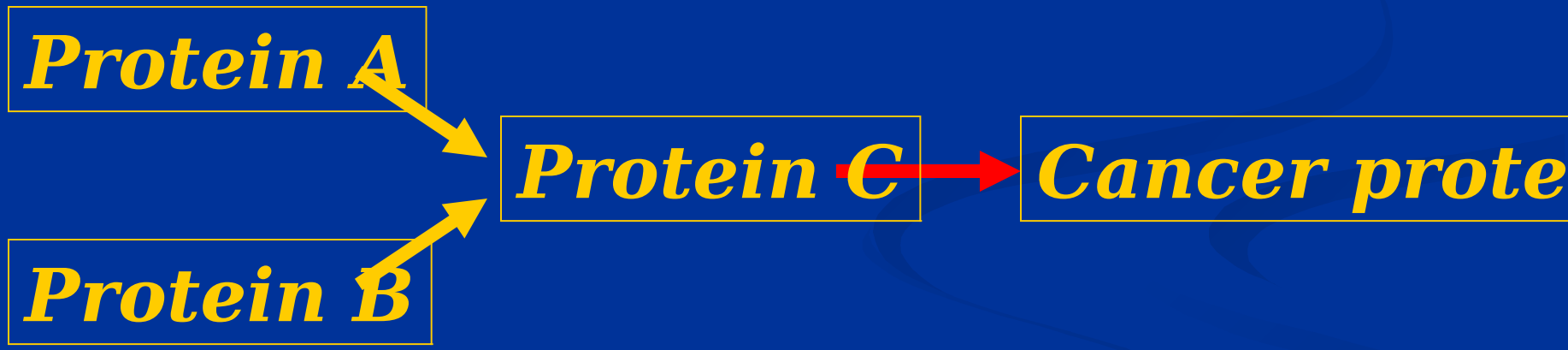


distinctive



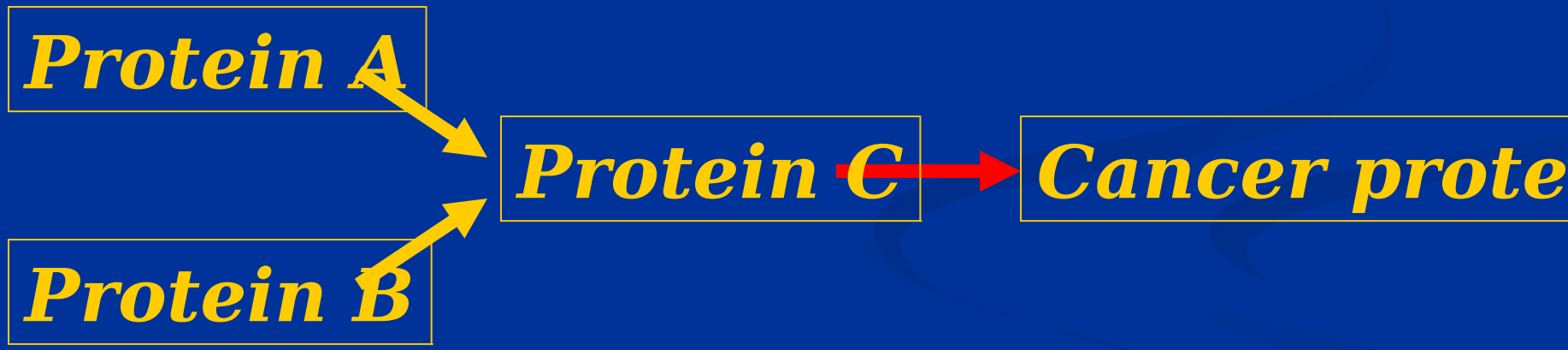
Causation from Correlation

- The following network is causally unambiguous if all variables are observed.



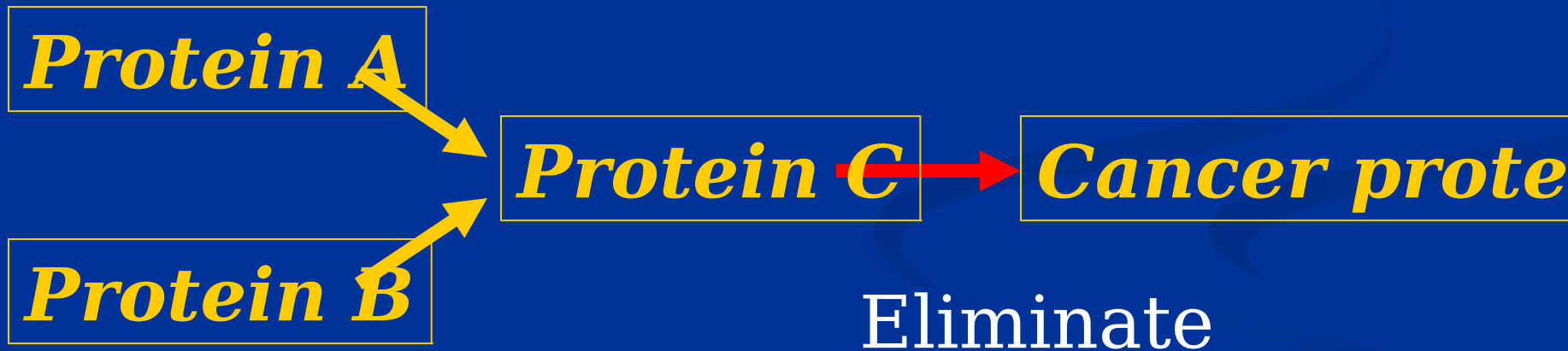
Causation from Correlation

- The red arrow is also immune to confounding by unobserved common causes.

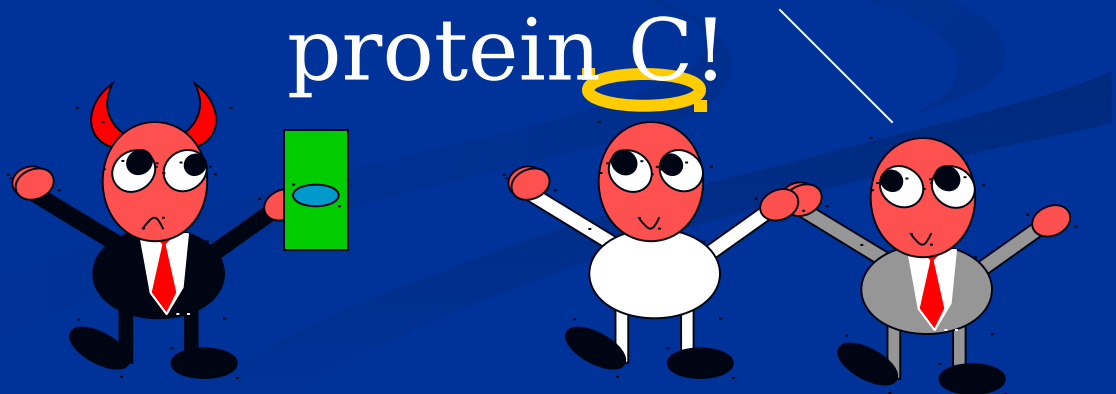


Brave New World for Policy

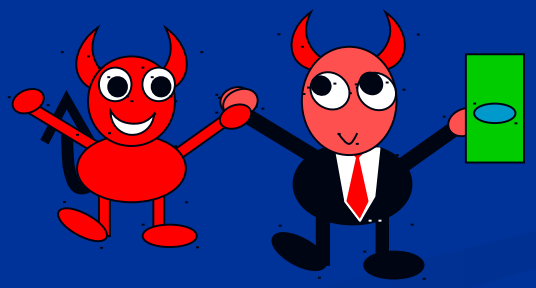
- Experimental (un-confounded) conclusions from non-experimental data!



Eliminate protein C!



V. The Catch

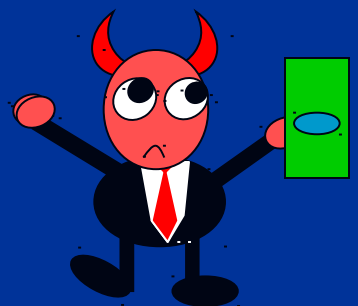


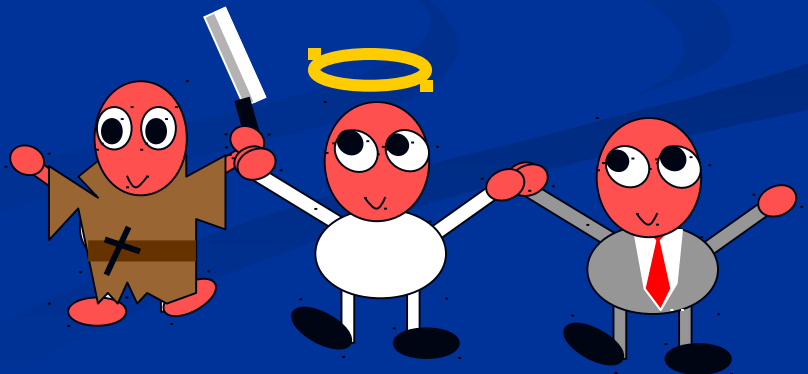
Problem of Induction

- Causal structure depends on statistical dependency discriminations that might be arbitrarily subtle.

Ockham's Razor is Crucial

- Ockham's razor: assume no more causal connections than necessary.





Causal Retractions

- No guarantee that small dependencies will not be detected later.
- Can have spectacular impact on prior causal conclusions.

Current Policy Analysis

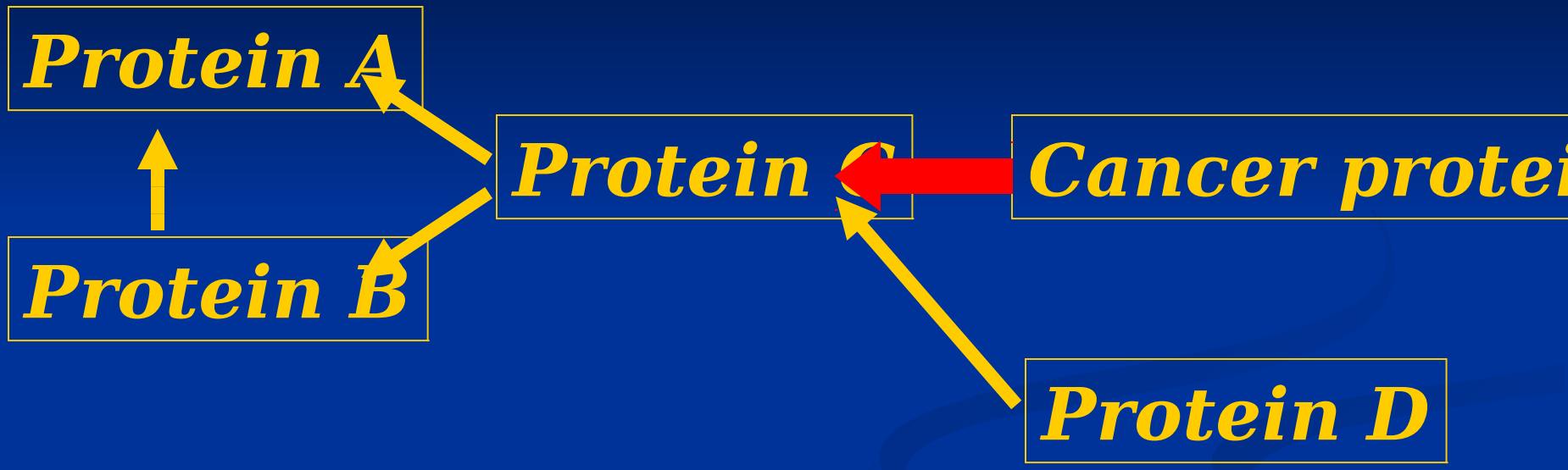
Protein A

Protein B

Protein C

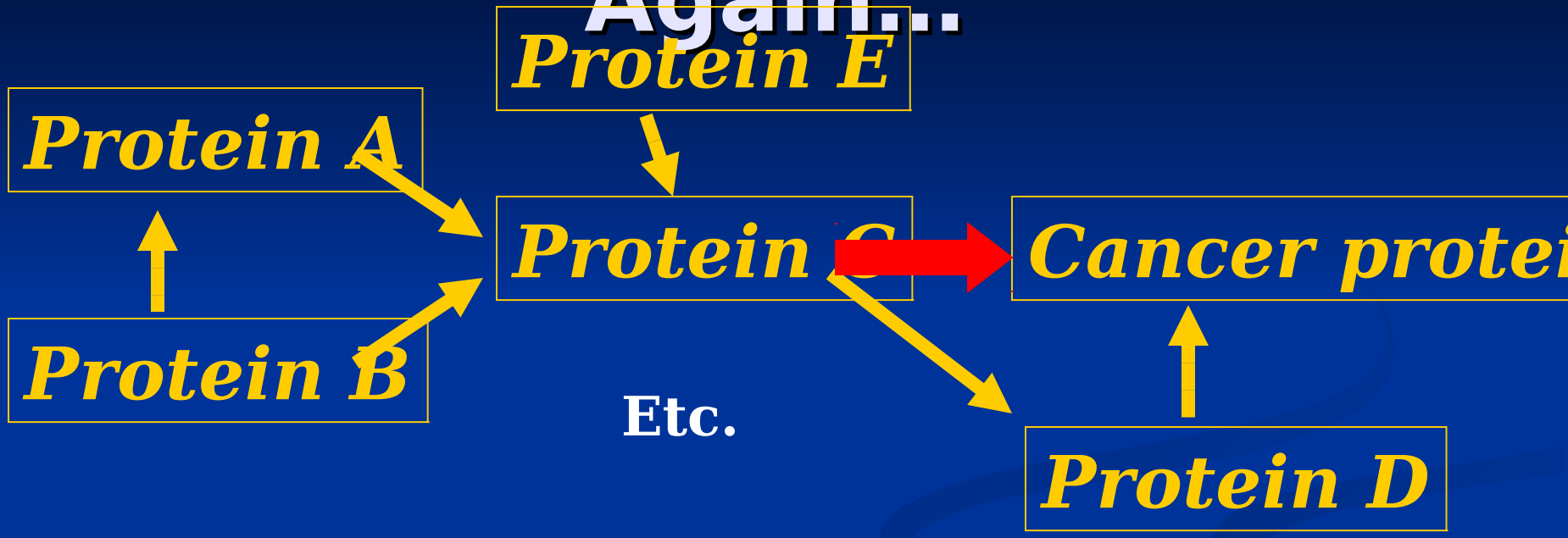
Cancer protei

As Sample Size Increases…

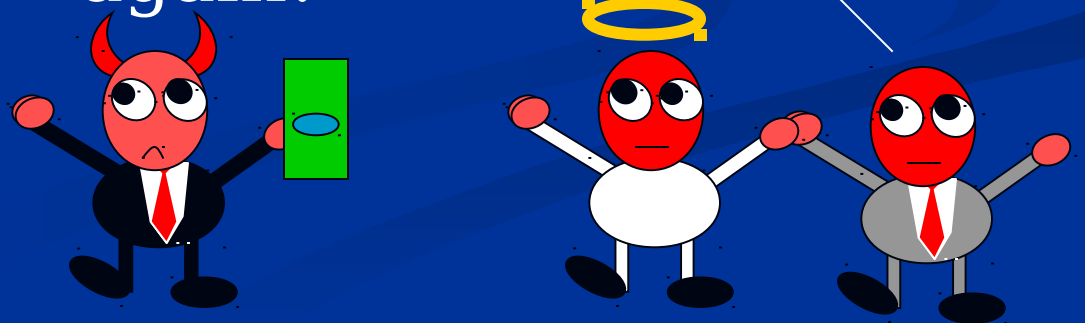


Rescind that order!

As Sample Size Increases Again…

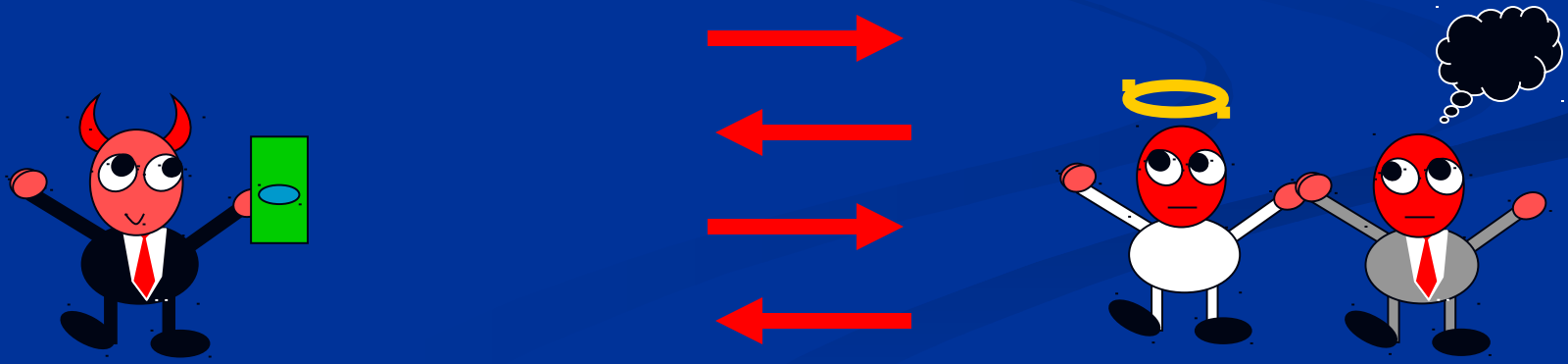


Eliminate protein C again!



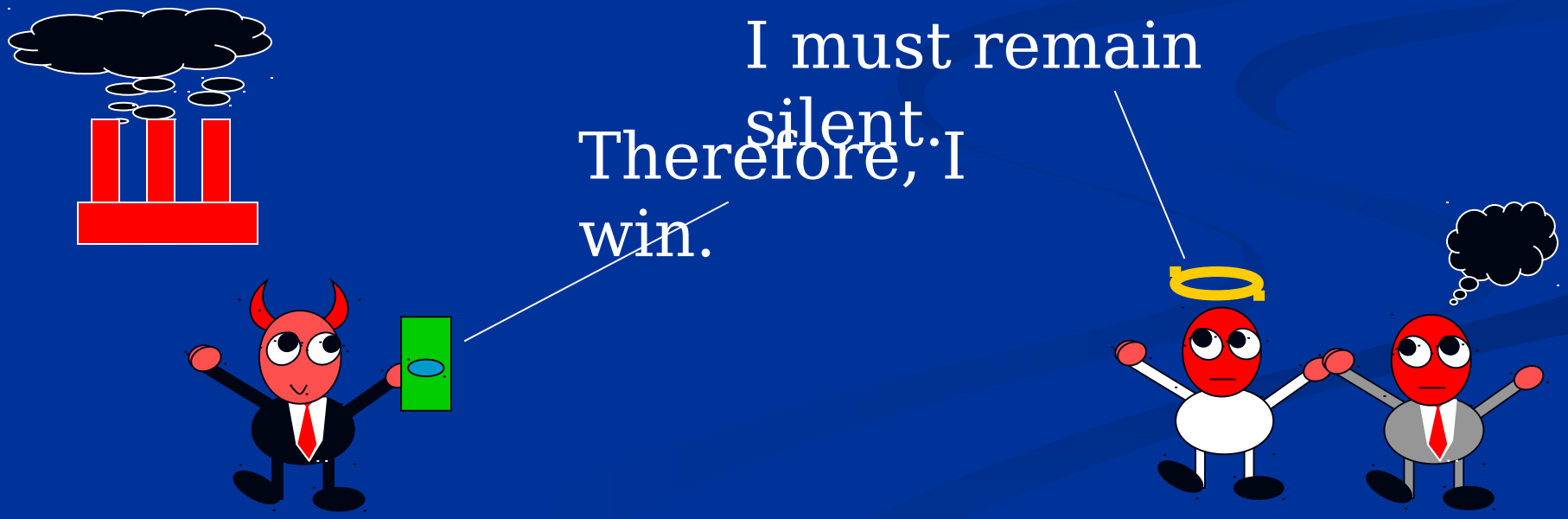
Causal Flipping Theorem

- Unamabiguous causal conclusions in a linear causal model can be forced by nature to flip any finite number of times as the sample size increases.



Extremist Reaction

- Since causal discovery cannot lead straight to the truth, it is not justified.



Moderate Reaction

- "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).

Skepticism Inverted

- Unavoidable retractions are justified because they are unavoidable.
- Avoidable retractions are not justified because they are avoidable.
- So the best possible methods for causal discovery are those that minimize causal retractions.
- The best possible means for finding the truth are justified.

New Directions

- Extension of unique efficiency theorem to mixed strategies, stochastic model selection and numerical computations.
- Latent variables as Ockham conclusions.
- Degrees of retraction.
- Ockham pooling of marginal Ockham conclusions.
- Retraction efficiency assessment of standard model selection methods.

Some Reading

- **"Ockham's Razor, Truth, and Information"**, in *Handbook of the Philosophy of Information*, forthcoming, J. van Benthem and P. Adriaans, eds.
- **"Ockham's Razor, Empirical Complexity, and Truth-finding Efficiency"**, *Theoretical Computer Science*, 383: 270-289, 2007.
- "A New Solution to the Puzzle of Simplicity", forthcoming, *Philosophy of Science*.

Pre-prints at:

www.hss.cmu.edu/philosophy/faculty-kelly.php