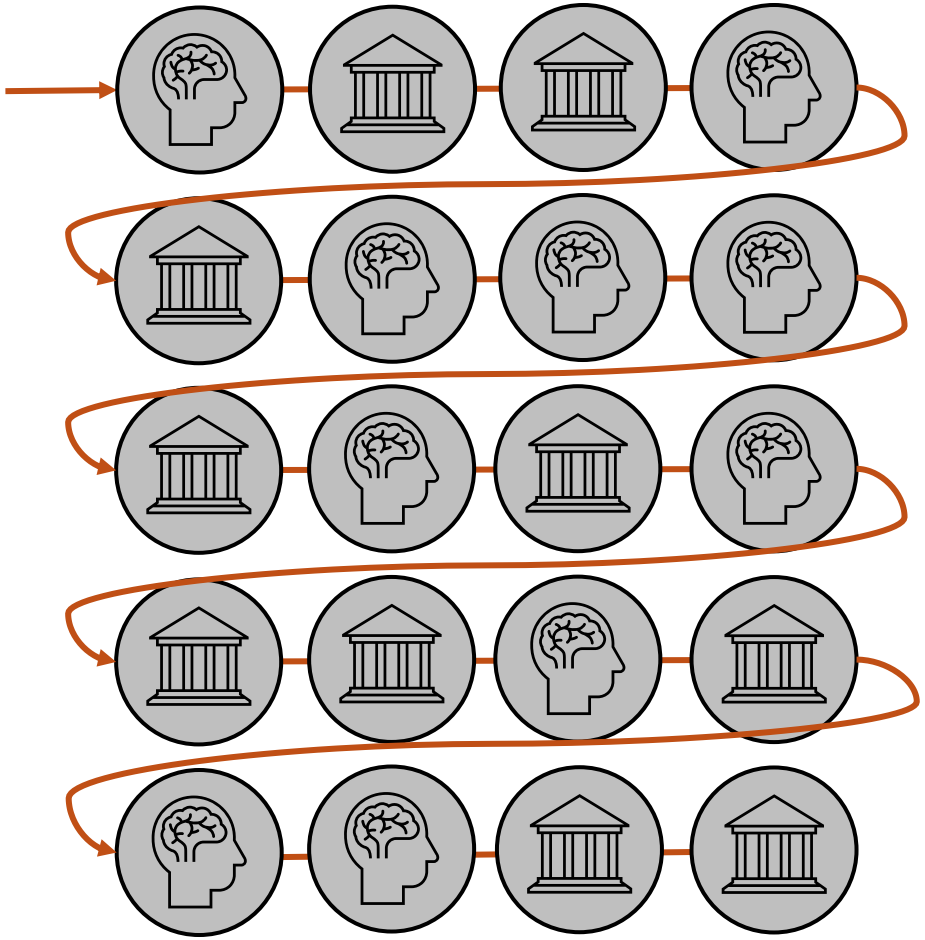
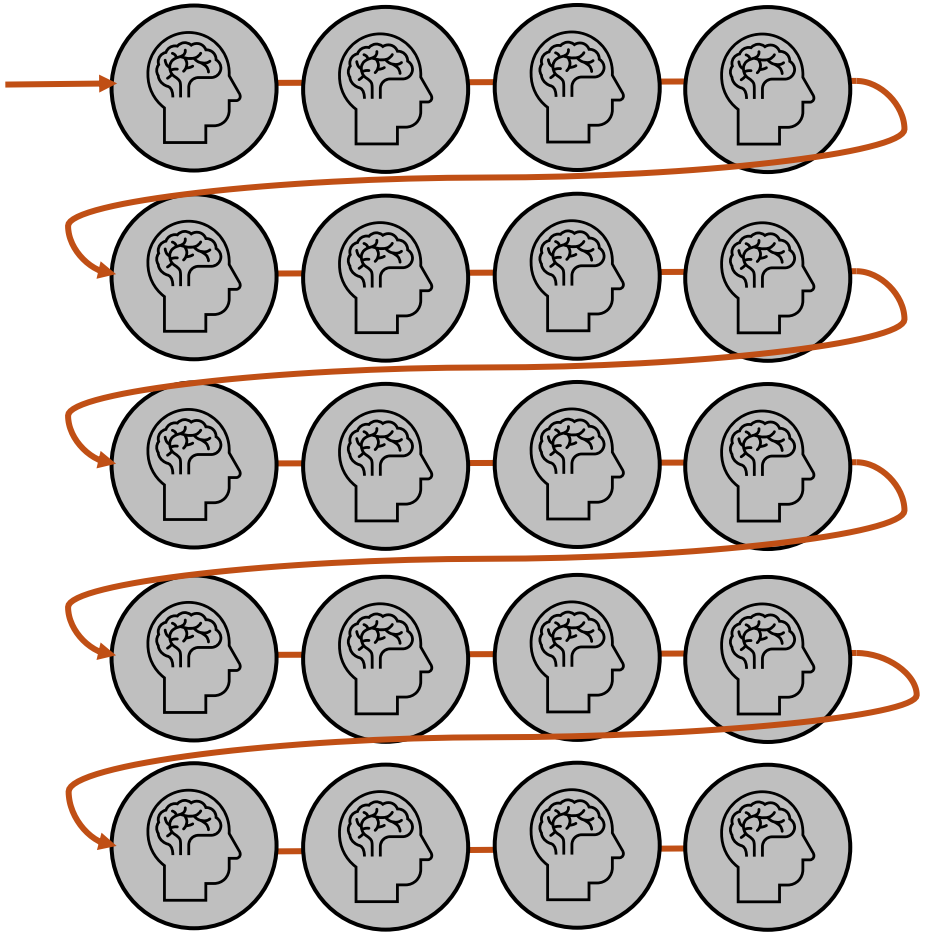
Principal of Parsimony

in statistics

Hannah Butler

Which is more likely?





Which is more plausible?

- Moon Landing Hoax
  - **H1:** The moon landing really happened.
  - **H2:** The moon landing was faked.

- Harris' AI airport crowd size
  - **H1:** A large group of people gathered at an airport in support of Kamala Harris.
  - **H2:** AI was used to generate a fake image of a large group of people at an airport in support of Kamala Harris.

Parsimony (Occam's Razor)

William of Ockham
(14th century English philosopher and theologian)

"Entities must not be multiplied beyond necessity"

- **Not** William of Ockham

350-ish BC

"We may assume the superiority *ceteris paribus* of the demonstration which derives from fewer postulates or hypotheses."

– Aristotle

130-ish AD

"We consider it a good principle to explain the phenomena by the simplest hypothesis possible."

– Ptolemy

1250-ish AD

"It is superfluous to suppose what can be accounted for by a few principles has been produced by many."

– Thomas Aquinas

Now-ish

"The simplest explanation is the best one"

- Many English-speaking scientists

Critics of Occam's Razor

1320-ish AD	**"If three things are not enough to verify an affirmative proposition about things, a fourth must be added, and so on."** – Walter Chatton
1600-ish AD	**"If one really wanted to start from a small number of entities, one could always consider the letters of the alphabet as the fundamental entities, since once could construct the whole of human knowledge out of them"** – Wikipedia (paraphrasing Galileo Galilei)
1680-ish AD	**Principle of Plentitude** (Gottfried Wilhelm Leibniz)
1760-ish AD	**"The variety of beings should not rashly be diminished."** – Immanuel Kant
1940-ish AD	**Law Against Miserliness: "Entities must not be reduced to the point of inadequacy" or "It is vain to do with fewer what requires more."** (Karl Menger)
1890-ish AD	**Pataphysics, "science of imaginary solutions"** (Alfred Jarry)
1988 AD	**"While Occam's razor is a useful tool in the physical sciences, it can be a very dangerous implement in biology."** – Francis Crick
Now-ish	**Inverse Occam's razor: "The simplest possible explanation is usually rejected"** - Igor Mazin

(Mis)Use in Statistics

Data

- Assume normality
- Assume independence
  - Observations
  - Variables (informally)
- Exclude/disregard outliers
- Categorize
- Assume inaccuracies are random

Statistical Modelling

- Bias-variance tradeoff
  - Demonize overfitting
  - Underemphasize underfitting
- Penalize complexity
  - Fewer variables are preferable
  - Regularization
  - Information Criterion

Parsimony in Statistics

- Since Occam's Razor ought to be invoked only when several hypotheses explain the same set of facts **equally well**, in practice its domain will be very limited…[C]ases where competing hypotheses explain a phenomenon equally well are comparatively rare (Holsinger 1980, pp. 144–5)
- When used in statistics, a model with fewer variables will ALWAYS explain the variability worse – can we really use the principal of parsimony to justify choosing a model with fewer variables? (replace "equally well" with "above a certain threshold")
- When invoking parsimony, the smaller set of variables is typically a subset of the larger set. How do you compare two models with distinct variables?
- Are all variables created equal? Can't some variables be more complex than others? Are some variables more parsimonious than others. If X1 is more complex than X2 and X3 combined, wouldn't $Y = mX1 + b$ be less parsimonious than Y = mX2 + kX3 + b?
- We most frequently measure the complexity of a model by the number of parameters, but how do we measure the complexity of a single parameter?
- Parsimony masks bias, bad sampling, bad data collection, bad model choice, lack of knowledge/understanding/research, etc., by generalizing
- Parsimony and confirmation bias
- https://medium.com/@qjbqvwzmg/occams-razor-in-machine-learning-the-principle-of-parsimony-de110ce7fe13

Parsimony Vs Elegance

- Ontological Simplicity vs Syntactical Simplicity
- Number of entities posited vs Number of assumptions
- In stats: we usually start with more assumptions and gradually reduce them
- Elegance and Parsimony interactions?
- Note: philosophically, parsimony requires knowledge of any and all involved entities (variables)?
- Principle of elegance
- Principle of Explosion

- “Occam's razor is a great tool when used as originally designed. Unfortunately, many scientists have turned this simple tool into a fetish object. They pursue simple explanations of complex phenomena as though parsimony were an end in itself, rather than a tool to be used in the pursuit of truth.”
- “the pursuit of parsimony is sometimes an obstacle to the pursuit of truth.” – Jonathan Haidt (2014)

Statistical Tools Against Parsimony

- **Nonparametric methods** address non-normality
- **Latent variable analysis/PCA/FA** address unobserved or unobservable factors
- **Bayesian Inference** addresses phenomena with subjective interpretations. Allows for updating when new data is available.
- **Machine Learning, Neural Nets, and AI??????**

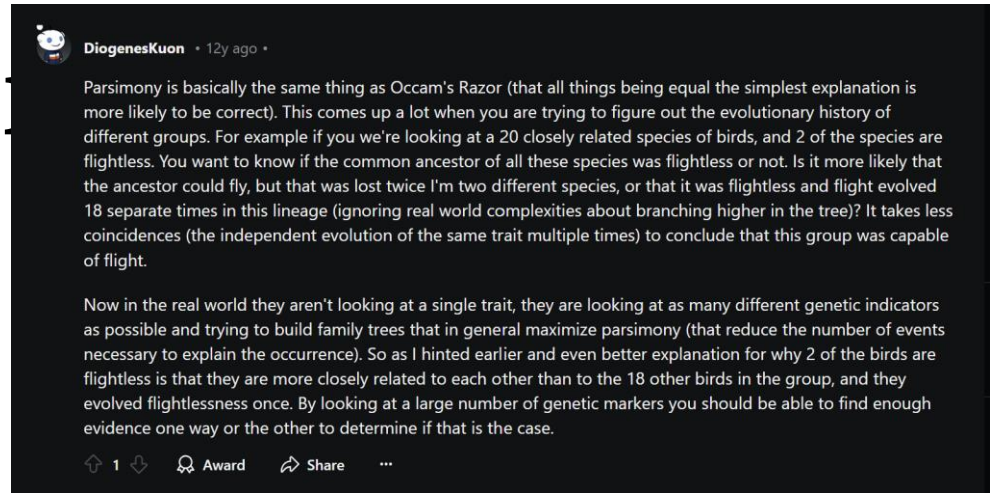
Phylogenetic Trees

History leading to Phylogenetics

- 14th century: Ockham's Razor
- 1763: Bayesian Probability (Rev. Bayes Version)
- 18th century: least abs. deviations (kind of like least squares)
- 1866: Ernst Haeckel
- 1912: Maximum likelihood (Ronald Fisher)
- Maximum parsimony: the shortest possible tree that explains the data is considered best
- Phylogenetic inference (it all goes back to stats baby)

Maximum Parsimony in Trees

- Maximum parsimony
- Phylogenetics
- Ernst Haeckel
- Scientific Racism/Eugenics

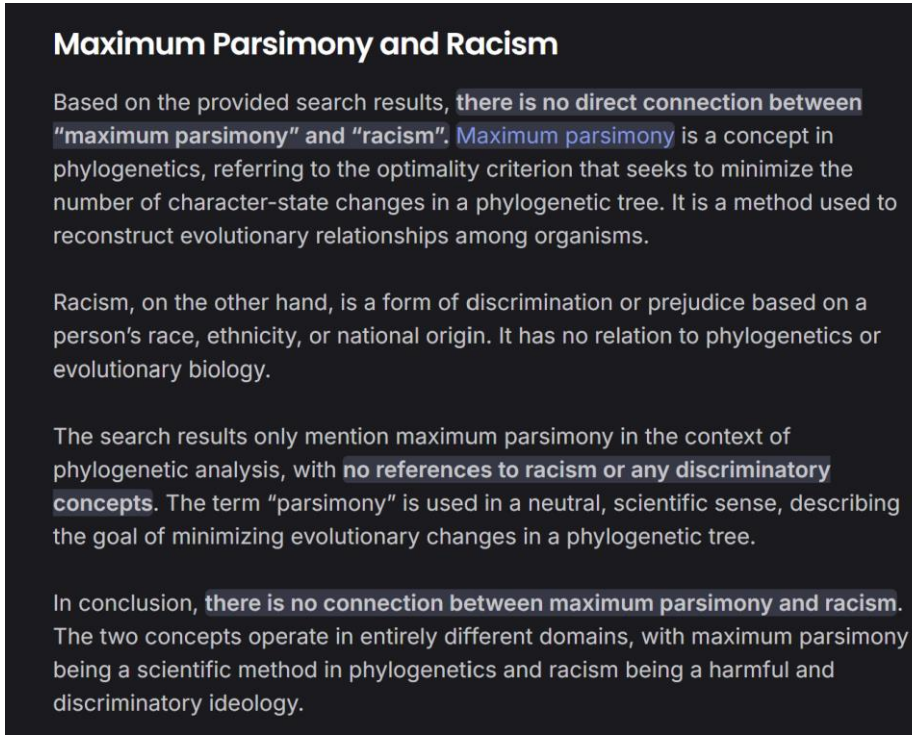


DiogenesKuon • 12y ago •

Parsimony is basically the same thing as Occam's Razor (that all things being equal the simplest explanation is more likely to be correct). This comes up a lot when you are trying to figure out the evolutionary history of different groups. For example if you we're looking at a 20 closely related species of birds, and 2 of the species are flightless. You want to know if the common ancestor of all these species was flightless or not. Is it more likely that the ancestor could fly, but that was lost twice I'm two different species, or that it was flightless and flight evolved 18 separate times in this lineage (ignoring real world complexities about branching higher in the tree)? It takes less coincidences (the independent evolution of the same trait multiple times) to conclude that this group was capable of flight.

Now in the real world they aren't looking at a single trait, they are looking at as many different genetic indicators as possible and trying to build family trees that in general maximize parsimony (that reduce the number of events necessary to explain the occurrence). So as I hinted earlier and even better explanation for why 2 of the birds are flightless is that they are more closely related to each other than to the 18 other birds in the group, and they evolved flightlessness once. By looking at a large number of genetic markers you should be able to find enough evidence one way or the other to determine if that is the case.

1 Award Share ...



Maximum Parsimony and Racism

Based on the provided search results, **there is no direct connection between "maximum parsimony" and "racism".** Maximum parsimony is a concept in phylogenetics, referring to the optimality criterion that seeks to minimize the number of character-state changes in a phylogenetic tree. It is a method used to reconstruct evolutionary relationships among organisms.

Racism, on the other hand, is a form of discrimination or prejudice based on a person's race, ethnicity, or national origin. It has no relation to phylogenetics or evolutionary biology.

The search results only mention maximum parsimony in the context of phylogenetic analysis, with **no references to racism or any discriminatory concepts**. The term "parsimony" is used in a neutral, scientific sense, describing the goal of minimizing evolutionary changes in a phylogenetic tree.

In conclusion, **there is no connection between maximum parsimony and racism**. The two concepts operate in entirely different domains, with maximum parsimony being a scientific method in phylogenetics and racism being a harmful and discriminatory ideology.

Joseph Felsenstein

- The effect of assortative mating on the genetic composition of a population in Eugenics Quarterly https://felsenst.github.io/eugenics.html
- PhD from University of Chicago Zoology (Charles Davenport)
- Publication in Heredity (founded by Fisher)

- “we should not go looking for more complex explanations when a simple one will do.”
- “the simplest explanation of the data must be correct”
- – David Hone (2008)

Parsimony for Resource Allocation

Thoughts

- Occam's razor has evolved from a basis for forming hypotheses to a basis for forming conclusions.
- In statistics we more often apply Occam's razor to final models, rather than the model assumptions.
- Occam's razor is characterized by elegance, intuition, and what "just makes sense."
- Occam's razor, by nature, will reinforce existing biases because of societal/cultural/individual notions of the above characteristics.
- Parsimony used to justify conclusions halts scientific progress by suggesting that there is no need for deeper understanding.
- In statistics we make little distinction between modelling physical and social phenomena.

Brain rot: Word of the Year

- Henry David Thoreau used the term to describe the scourge of oversimplification (Walden, 1854)
- “Thoreau criticizes society’s tendency to devalue complex ideas, or those that can be interpreted in multiple ways, in favour of simple ones” (Oxford University Press)
- Contrasts(?) with his values on living a simple life
- Thoreau: 19th century Brain rot = 21st century Parsimony?
- Thought: social media feeds us endless 1-variable explanations for issues; The brain rot we are fed is determined by a highly complex statistical model

Is Complexity Accessible?

- How do we reconcile an ideal of complexity with accessibility
- How can something so inaccessible by most (science) be parsimonious?
- Do we try to mask parsimony with false complexity?

Sources

- https://en.wikipedia.org/wiki/Occam's_razor
- https://scienceandsociety.duke.edu/occams-razor-and-the-virtue-of-science/
- https://www.recorder.com/Guest-columnist-Joralemon-56653580
- https://math.ucr.edu/home/baez/physics/General/occam.html
- https://fs.blog/occams-razor/
- http://pespmc1.vub.ac.be/OCCAMRAZ.html
- https://www.jstor.org/stable/27113318
- Chappell, Bill. "Writer Thoreau Warned of Brain Rot in 1854. Now It's the Oxford Word of 2024." *NPR*, December 2, 2024, sec. Books. https://www.npr.org/2024/12/02/nx-s1-5213682/writer-thoreau-warned-of-brain-rot-in-1854-now-its-the-oxford-word-of-2024.
- https://effectiviology.com/parsimony/