Knowledge and certainty

- Traditionally, philosophers regarded information as knowledge only when it was 100% certain (both justified and true).
- To rationalists, only ideas that could be deduced from certain first principles was true knowledge

  - e.g. Descartes' *I think therefore I am*

- To empiricists, only observations provided certainty
- Hume was the first to suggest that many things we "know" are not "real" knowledge!

Was Hume an empiricist?

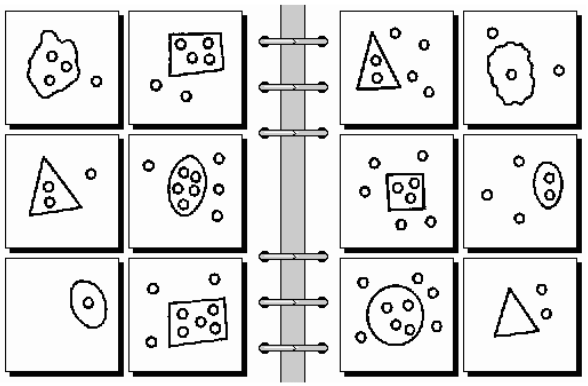
- Hume is usually classified as an empiricist because he assumed that knowledge can only come from observation

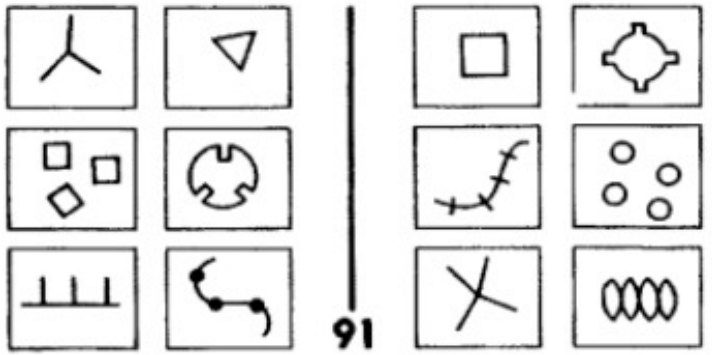
  - E.g.: that the only way we know the sun will rise tomorrow is as an extrapolation from past experiences with the sun

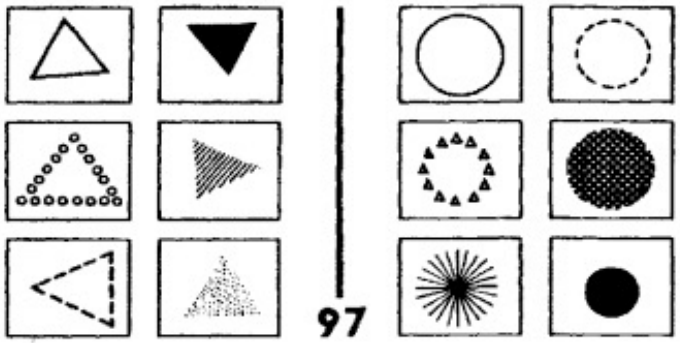
- However his main contribution was to show that without assumptions, pure observation does not provide any knowledge!

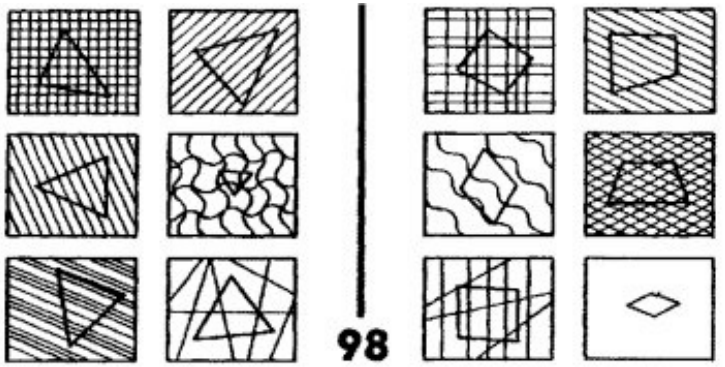
  - I.e. that "the sun will rise tomorrow" does not actually follow from our past observations

Bongard Problems









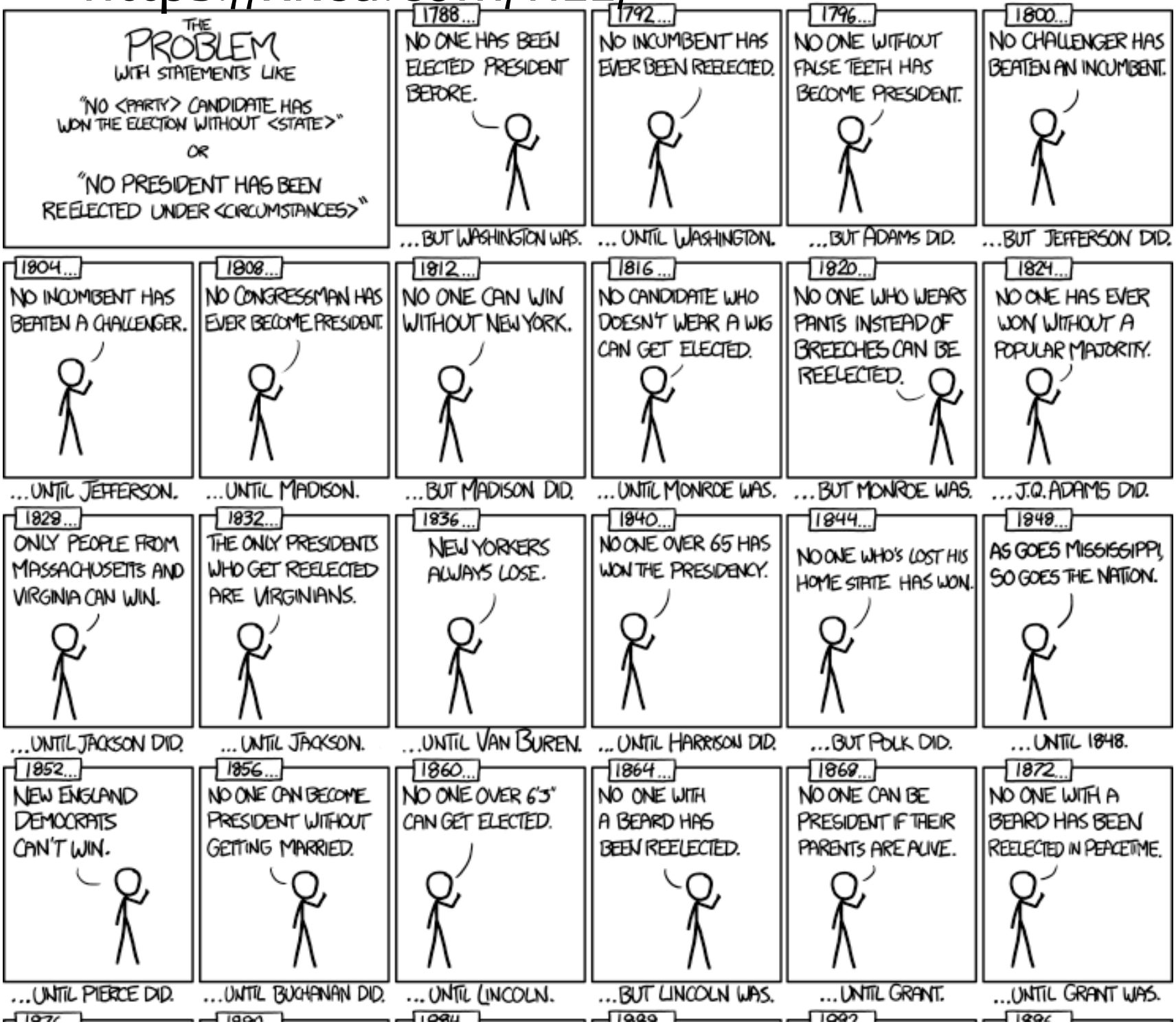
- Inductive generalizations assume that future (unknown) data will be "similar" to past (known) data.
- But what does similar mean?

  - After all, the future doesn't match the past *exactly*!

- The principle of Uniformity of Nature (Hume)

  - The world is a fairly regular place

https://xkcd.com/1122/

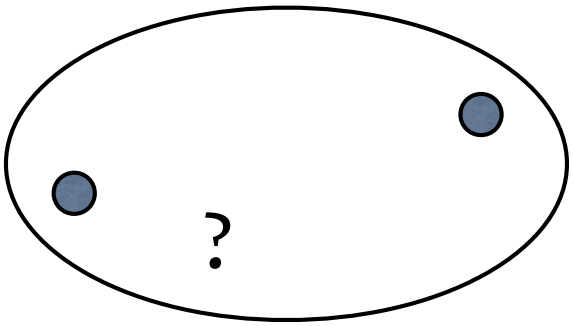


Factors influencing induction

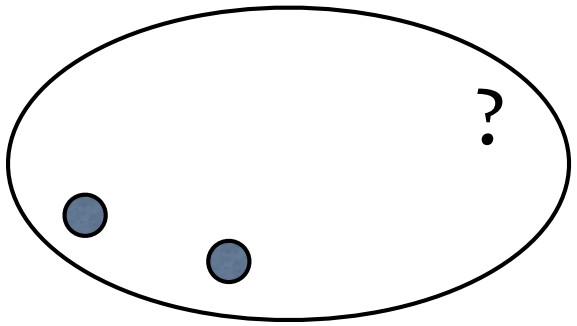
Similarity of premises

- *cats have spleens*
- *whales have spleens*
- Therefore *all mammals have spleens*

- *deer have omenta*
- *elk have omenta*
- Therefore *all mammals have omenta*



vs



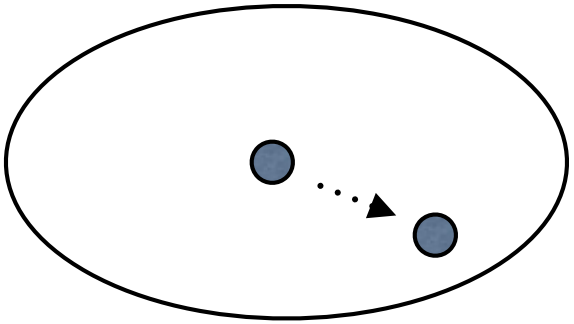
Stronger

Weaker

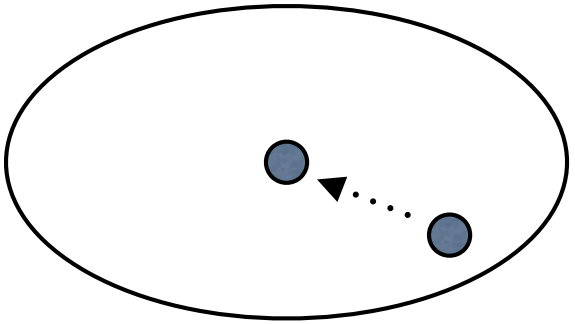
Typicality of premises

- *Ellen Robinson has a microwave oven in her kitchen*
- therefore *all Americans have a microwave in their kitchens*

- *Queen Elizabeth has a private helicopter*
- therefore *All English people have a private helicopter*



vs



Stronger

Weaker

Monotonicity

- Monotonic inference

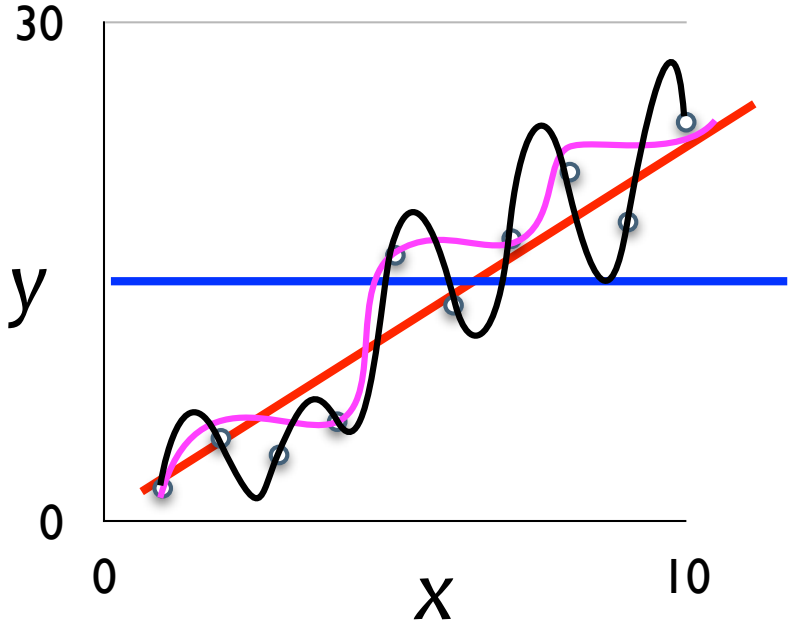
  - Once you conclude *P* is true, it remains true. Likewise if you conclude *P* is false.

- Deduction is monotonic

- Induction is nonmonotonic

  - After poking around, Sherlock homes thought the butler did it. But after considering more evidence, he changed his mind.

Data-fitting



$y = b$ [underfits]

$y = ax + b$ [fits well]

$y = ax^7 + bx^6 + cx^5 + dx^4 + ex^3 + fx^2 + gx + h$ [overfits]

A model with more parameters [= "knobs" = "fudge factors"] that are fitted to the data can always fit better. (Even perfectly.)

But this leads to terrible generalization because the model is fitting noise along with signal.

An underfit model is too far from the data. It's missing regularities.

An overfit model is too close to the data. It's fitting noise.

The best model fits all the regularities but none of the noise.

- In general there is no way to tell what is "real" regularity and what is noise.