The Book of Why
Judea Pearl

NHS-R Book group,
Intro and chapter 1
15-01-2021

Causal analysis

- Asks questions like:
    - How effective is a particular treatment at preventing disease?
    - Did changes to tax or advertising cause sales to go up?
    - Should I quit my job?

The problem with science

- We measure air pressure with a barometer
- We can write this as:

$$B = kP$$

- (Barometer reading = constant x air pressure)
- We can, e.g., rewrite as:

$$B / k = P$$

- But we haven’t clarified whether B causes P or P causes B!

A mathematical language of causality

- We lack a formal language of causality
- In simple cases it is not needed:
    - Hot weather => ice cream sales *not*
    - Ice cream sales => hot weather
- It is necessary in more complex cases
- There are historical parallels
- E.g. nobody formalised probability until interest was provoked in gambling in C17th

Statistics and causality

- Galton and Pearson explicitly turned away from causality at the birth of statistics
- Statisticians learn
  - Correlation is not causation
  - But then what is it?
- 100 years ago even asking if tobacco causes cancer would have been considered “unscientific”

A calculus of causality

- The effect of drug on lifespan

  $$p(L \mid do(D))$$

- (probability that someone lives L years given that they take drug D)
- The “do” indicates an intervention as has no parallel in statistics
- The “do” rules out confounding
- Without “do”:
  - $p(L \mid D)$
  - We have no idea about confounds

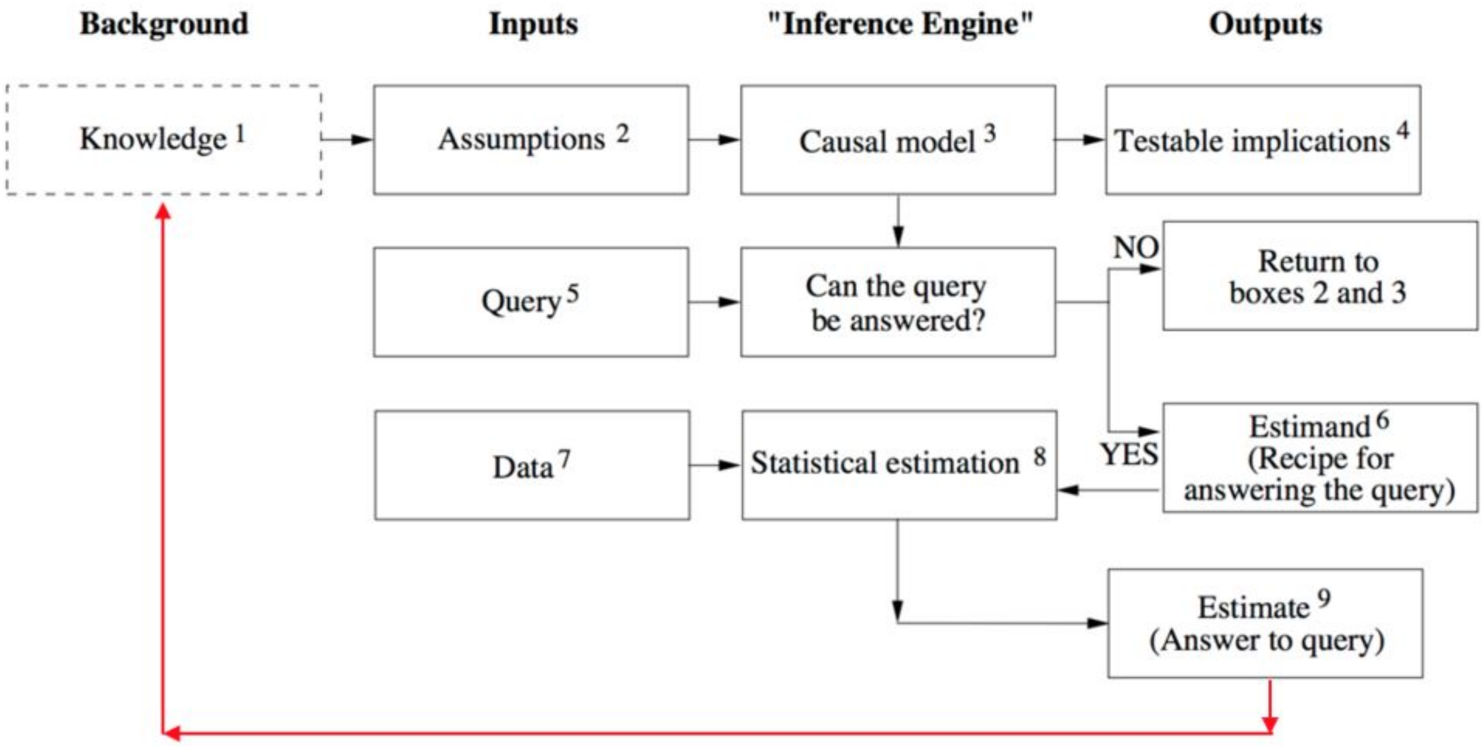
Counterfactuals and causal reasoning

- Another key part of causal reasoning is counterfactuals

$$p(L \mid do(\text{not-}D))$$

- Counterfactuals are fundamental to reasoning and moral behaviour
- A language of causality and counterfactuals can be used in artificial intelligence
- Grammar and language are building blocks for human reasoning as well as “strong” AI

Inference engine



Inference engine

- Inputs:
  - Assumptions
  - Queries
  - Data
- Outputs:
  - Can the given question be answered?
  - Estimand (a recipe for generating the answer)
  - An actual answer (including uncertainty)

The success of causal reasoning

- A deep learning model trained on one patient group in one hospital
- Can be retrained in another hospital
- But can only learn new *associations*, not new causal models
- Google flu trend’s failure

Human success is based on causal reasoning

- For example reasoning about the success of a hunt:
  - Number of hunters
  - Size of mammoth
  - Weather
  - Terrain
  - Direction of attack
  - …
- A counterfactual layer is important
  - Imagine success with twice as many hunters

The causal ladder

- Seeing
- Doing
- Imagining
- The book proposes a model based on what causal reasoning at each level is capable of
    - Parallels with Turing’s reasoning about the types of statements that are computable by different types of computer

Seeing. Association

- An owl watches a mouse's behaviour and predicts where it goes next
- A deep learning model learns to play Go
- What is the probability that someone who buys toothpaste will buy floss

$$p(\text{floss} \mid \text{toothpaste})$$

Doing. Intervention

- What happens if we double the price of toothpaste?
- What happens if we give drug B?
- Statistics is no help, relying on intervention
- We can go from level 1 to level 2 with a strong causal model

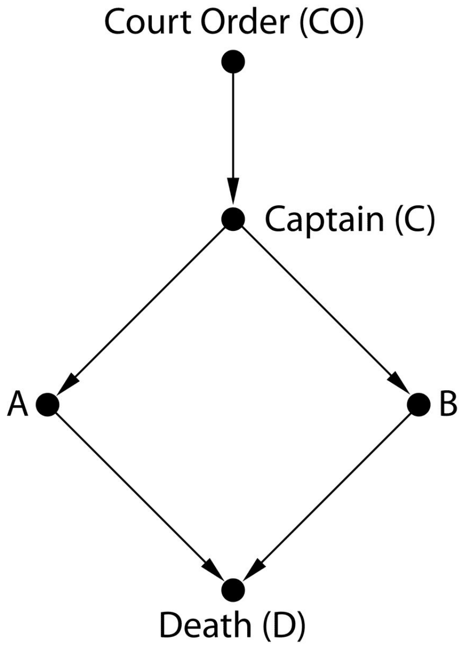
Imagining

- Level 2 does not say why
- What would have happened if we hadn’t had a sale on toothpaste?
- The laws of physics exist at this level
- Level 3 laws are backed by lots of level 2 (experimental) observations
  - Backed by a theory
- Harari (and Pearl) theorise that this level was a fundamental plank of the dominance of our ancestors over all other hominids existing at the time

The mini Turing test

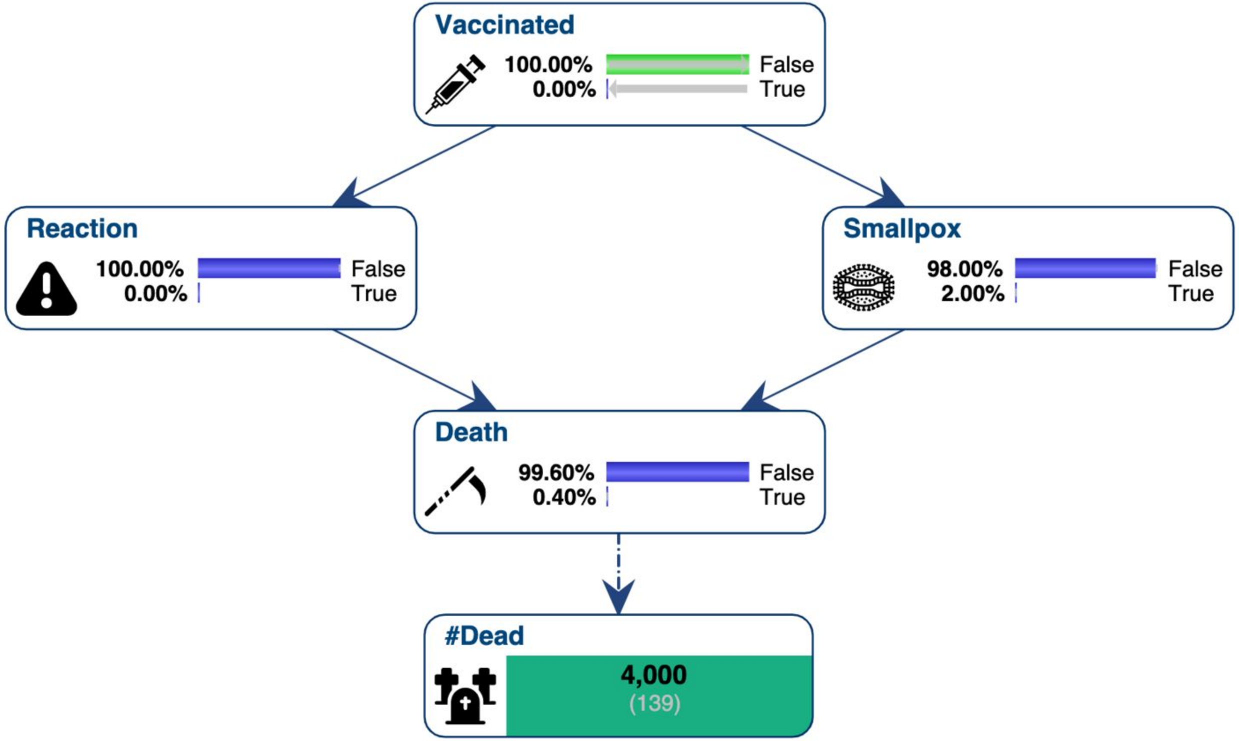
- Teach a computer a story, encoded in any suitable fashion
- Get the computer to answer causal questions about the story, in the way a human might
- The Chinese room experiment

The firing squad



- Can begin to reason causally
- If prisoner is dead, did CO occur?
- If A fires, does B fire?
- What if we *make* A fire?
  - Counterfactual reasoning- remove the lines that cause A

“Vaccines cause death!”



The power of causal diagrams

- Causal analysis of this kind is invariant
  - A new, safer vaccine
  - Better treatment of smallpox
  - Reduced incidence of smallpox
- The model still works with the new probabilities
  - (which can be estimated with statistics)
- Crucially, “X raises the probability of Y” *cannot* be written:
  - $p(Y \mid X) > p(Y)$
  - This is a conditional probability and ignores confounding

An attempt at repair

- Philosophers have attempted to deal with confounding with

$$p(Y \mid X, K = k) > p(Y, k = K)$$

- The problem remains at defining what goes in K
- *Anything* could theoretically be in there
- Apparent cheating in an election was caused by a cosmic ray
- But you can answer:

$$p(Y \mid do(X)) > p(Y)$$