

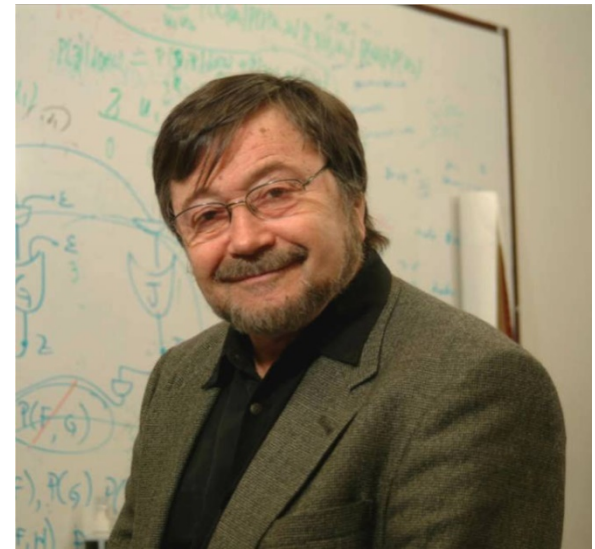
# Why did the ...?

## Judea Pearl and his quest for Causality

Deepak Khemani

# Judea Pearl

Judea Pearl was born on September 4, 1936, in Tel Aviv, which was at that time administered under the British Mandate for Palestine. He grew up in Bnei Brak, a Biblical town his grandfather went to reestablish in 1924.



From [The Daniel Pearl foundation](#)

In 1956, after serving in the Israeli army and joining a Kibbutz, Judea decided to study engineering. He attended the Technion, where he met his wife, Ruth, and received a B.S. degree in Electrical Engineering in 1960.

[The Judea Pearl ACM Turing award page](#)

## A tragic event

Daniel Pearl (October 10, 1963 – February 1, 2002) was a journalist for the Wall Street Journal with American and Israeli citizenship. He was kidnapped by terrorists and later murdered in Pakistan.

Pearl was kidnapped while working as the South Asia Bureau Chief of the Wall Street Journal, based in Mumbai, India. He had gone to Pakistan as part of an investigation into the alleged links between British citizen Richard Reid (known as the "shoe bomber") and Al-Qaeda. Pearl was killed by his captors.



[https://en.wikipedia.org/wiki/Daniel\\_Pearl](https://en.wikipedia.org/wiki/Daniel_Pearl)

## A puzzle for homework

A desert traveler  $T$  has two enemies. Enemy 1 poison's  $T$ 's canteen, and enemy 2, unaware of 1's action, shoots and empties the canteen.

A week later,  $T$  is found dead and the two enemies confess to action and intention.

A jury must decide whose action was the *actual cause* of  $T$ 's death.

## Judea Pearl – influential books

- Heuristics: Intelligent Search Strategies for Computer Problem Solving, 1984.
- Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, 1988.
- Uncertainty management in AI systems, 1988
- **Causality**, 2000.
- Causal Inference in Statistics: A Primer, *with* Madelyn Glymour, and Nicholas P. Jewell, 2016.

ALL quotes in this talk, unless otherwise stated, are from the above.

## Some fundamental questions

What is **intelligence**?

What is **thinking**?

What is a **machine**?

Is the **computer** a machine?

Here on when we say machine we will mean a programmable computer system

Can a **machine** think ?

If yes are **WE** machines?!

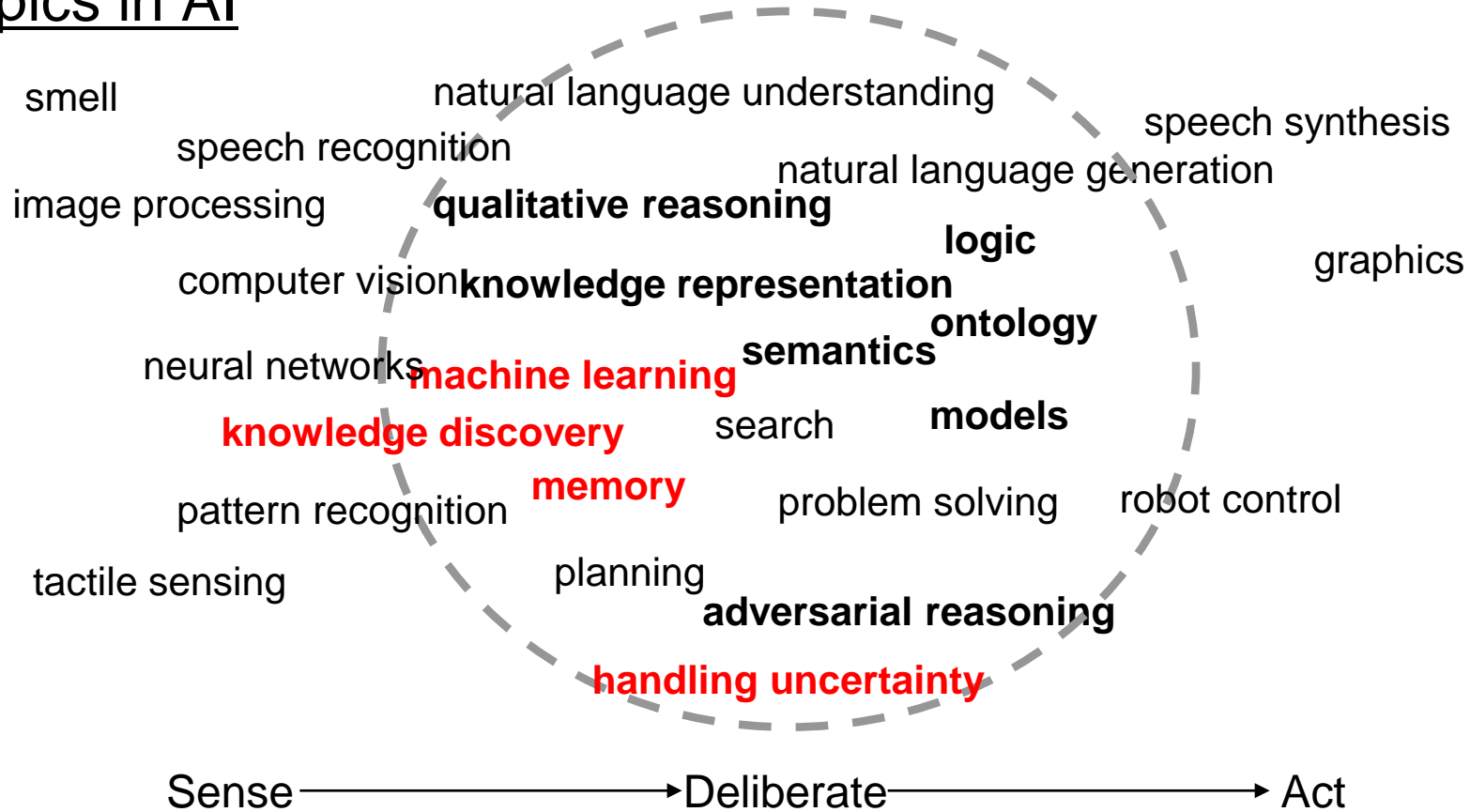
# Creating a model of the world

- An AI agent needs to create a **model** of the world that it is operating in.
- This model should help the agent understand the world, diagnose the cause of events and states, predict the future, and plan to achieve its goals.
- Judea Pearl believes that **causal relationships** are fundamental both to **representing and reasoning** about physical reality, and also the way humans reason about it.
- Reasoning in a world with incomplete information has to contend with **uncertainty**, and **probabilistic models** have often come to our aid here.

# Do we need *reasoning* when we have ML?!

- Machine Learning helps us solve *similar* problems
  - by keeping a memory of past solutions, as in *Case Based Reasoning*
  - by tuning parameters used in game playing by studying human games and learning from them, as in *AlphaGo*
  - by learning to **classify** events, utterances, text documents and images by learning from data, as done by *Deep Neural Networks* nowadays.
  - by learning from scratch (tabula rasa) by playing millions of games against itself, as in *AlphaGo Zero*, and a similar version for Chess that became an expert in a day or so.
- Reasoning is needed when a problem is new, and when the task involves *patterns of inferences that do not repeat*.

# Topics in AI



Source: Deepak Khemani, A First Course in Artificial Intelligence

# The many difficulties in automated reasoning

- Logical inference is local. But finding a proof, a chain of inferences is hard.
- Default reasoning is non-local. It involves something like “unless you can show otherwise”. For example,
  - If Jimmy is a professor then (it is reasonable to conclude that) he has a PhD degree, unless you know otherwise.
- Probabilistic reasoning likewise needs look at the entire data in the Naïve Bayes inference
  - Is reasoning in Bayesian Belief Networks local?
  - Where do the probabilities come from?

# The Compactness of Logical Reasoning

- To infer  $Q$  from  $P$  and  $(P \rightarrow Q)$ , one does not have to look at the rest of the knowledge bases.
  - but one does have to find the statements  $P$  and  $(P \rightarrow Q)$ .
- But universals in logic can create a problem of inconsistency
  - From  $\text{Bird}(\text{peppy})$ ,  $(\forall x (\text{Bird}(x) \supset \text{Flies}(x)))$  and  $\neg \text{Flies}(\text{peppy})$  one can derive anything
- Solution: Default Reasoning (for example, Circumscription)
  - If Peppy is a bird, and it is consistent to believe that Peppy can fly, then it is reasonable to infer that Peppy can fly.
  - consistent = it cannot be shown that Peppy cannot fly
    - this requires one to scan the entire KB and lose locality

# Logic

- Classical logic focuses on valid arguments.
- Connecting true premises to valid conclusions.
  - does NOT bother on the truth of premises.
  - the onus is on the user who represents the knowledge
  - sloppy representation can lead to fallacious conclusions, for example
    - If the grass is wet, then it rained last night. (abduction)
    - If we water the lawn, the grass will be wet. (causal relation)
    - Therefore, if we water the lawn then it rained last night.
- The *implication* statement does *not* capture causality  
( $P \rightarrow Q$ ) is equivalent to ( $\sim P \vee Q$ )  
either P is *false* or Q is *true*
- Intuitionistic logics .... (?)

# Data

- All reasoning and learning is from data.
- In modern times both data and the ability to crunch it has increased rapidly.
  - Erik Brynjolfsson and Andrew McAfee, in *The Second Machine Age* say that “.. we are now in the second half of the chessboard”.
- Question: Can we derive the Cause-Effect relationship from data?
  - statisticians have played hide and seek with the notion of causality
  - though it is of great importance in economics, epidemiology, sociology and psychology
  - what is the relation, if any, between correlation and causality?
  - physics seems to have abandoned causality in favour of equations and algebra

# Causality or correlation?

- Correlation does not always imply causality / causation
- Correlation is when two things (events) are related
  - an oft quoted example - the number of people drowning is correlated with the sale of ice creams
  - but is their a causal relation? No
  - If the weather is nice (which means warm in many countries) then
    - more people buy ice cream, and also,
    - more people go swimming
  - another example – married men live longer than single men
  - but maybe the high life expectancy features like wealth and health make it more likely to find a spouse?
- Sometimes when two variables are correlated, one of them causes the other (e.g. milk consumption and bone strength)
  - more on this later...

# Causality or correlation?

- Correlation does not always imply causality
  - Ice cream and murders (In 2009 a *New York Times* article on summer killings in New York City cited a CDC study that found a national increase in homicide numbers between the months of July and September.
  - Rooster crooning does not cause the sunrise (or vice versa?)
- In many situations we need to find the cause of something
  - Legal reasoning (was his cancer caused by smoking?)
  - Medical diagnosis (are his symptoms indicative of cancer?)
- We often do abductive inferences for medical diagnosis, and qualify our arguments by things like “it is very likely”, for example,
  - If the heel is painful it is very likely that you have Plantar Fasciitis
  - Doctors rely on confirmatory tests
  - MYCIN used *certainty factors*

# The erosion of Causality through the ages

Galileo in his 1638 book “Discorsi” states two maxims –

1. Description first, explanation second

– that is, the “how” precedes the “why”

2. Description is carried out in the language of mathematics

– Algebra

Ask not, said Galileo, whether an object falls because it is pulled from below or pushed from above.

Ask how well you can predict the time it takes for an object to travel a certain distance...

# The erosion of Causality through the ages

Philosophers still searched for the origin and explanation of Galilean equations.

- Descartes ascribed cause to eternal truth.
- Leibniz made cause a self evident logic law.
- But, Hume, extended maxim 1 even further,
  - “the *why* is not only second to the *how*, but the *why* is totally superfluous as it is subsumed by the *how*”

## David Hume, *A Treatise of Human Nature*, 1739

“Thus we remember to have seen that species of object we call *flame*, and to have felt the species of sensation we call *heat*. We likewise call to mind their constant conjunction in all past instances. Without any further ceremony, one can call one *cause* and the other *effect*, and infer the existence of the one from that of the other.”

- causal connections are a product of observations
- a learnable habit of the mind
- almost as fictional as optical illusions
- and as transitory as Pavlov’s conditioning

from *The Art and Science of Cause and Effect*, in *Causality*.

# David Hume

- *“We may define a cause to be an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed”.*
- A bit like the 5 step syllogism used in Nyaya Sutra
  - Smoke fire example (next slide)
  - When you see smoke you infer fire.
  - But the fire did not cause the smoke. Rather the other way around.
- Causality is hard to determine from data/statistics

# The 5 Step Syllogism

The *Nyāyasūtra* describes the structure (Mohanty, 2000) of a good argument as a five step process.

- 1.a statement of the thesis (*Pratijñā*): *there is fire on the mountain*
- 2.a statement of reason (*Hetu*): *because there is smoke on the mountain*
- 3.an example of the underlying rule (*Udahāraṇa*):  
*where there is smoke there is fire, like the culinary hearth*
- 4.a statement that (*Upānaya*): *this case is like that*
- 5.finally the assertion of the thesis proven (*Nigamana*):  
*therefore the mountain is on fire*

The derived piece of knowledge is known as *anumāna* (after cognition).

## Logical Arguments

This is in contrast to the three step syllogism exemplified by the Socratic argument. It has also been observed by Müller (1853; 1859) that the Indian philosophers used the five step reasoning only when the task was to convince others about their conclusions. When the task was to infer something for oneself the simpler three step process was used, as follows.

1. There is smoke on the mountain
2. Wherever there is smoke there is fire
3. Therefore, there is fire on the mountain

This is precisely the form of reasoning, the Aristotelian syllogism, which is fundamental to western logic.

# Causality in science

Bertrand Russell : “*All philosophers imagine that causation is one of the fundamental axioms of science, yet oddly enough, in advanced sciences, the word ‘cause’ never occurs...*”

Patrick Suppes: “*There is scarcely an issue of ‘Physical Review’ that does not contain at least one article using either ‘cause’ or ‘causality’ in its title.*”

Do physicists talk, write and think one way  
and formulate physics in another?  
Do they continue to write equations in the office  
and discuss cause–effect in the cafeteria?

from *The Art and Science of Cause and Effect*, in *Causality*.

# Equations

Newton's law:  $f = ma$

Does force cause acceleration?

For example, in a body in free fall.

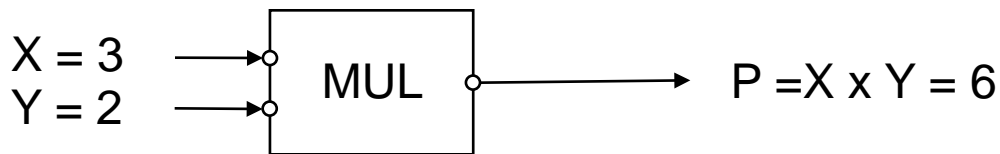
or

Does acceleration cause force?

For example, on a passenger in an accelerating car.

The equation is (only) used to compute the relation between  $f$  and  $a$

# A Multiplier



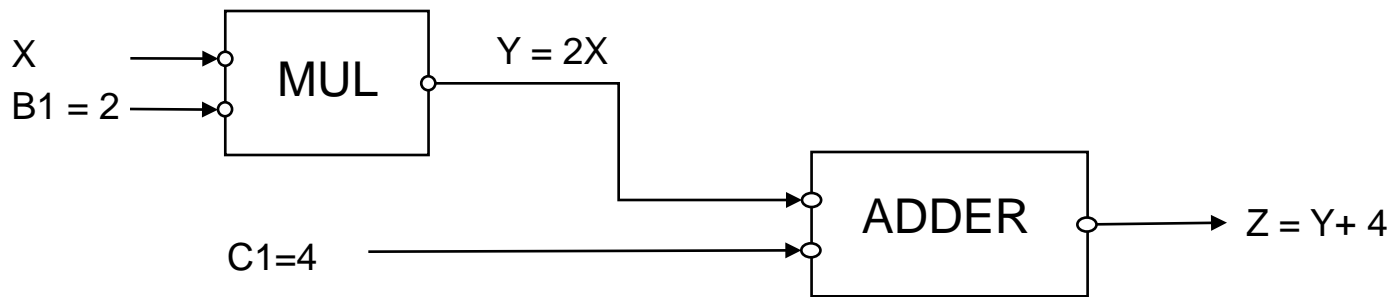
A multiplier can be described by an equation that says that  $P = XY$

but the equation does not capture causality.

A assignment statement  $P \leftarrow X * Y$

does assert that  $X$  and  $Y$  are the independent variables and  $P$  the dependent variable

# Cause has directionality, equations do not



What if we clamp  $Y$  at 10?

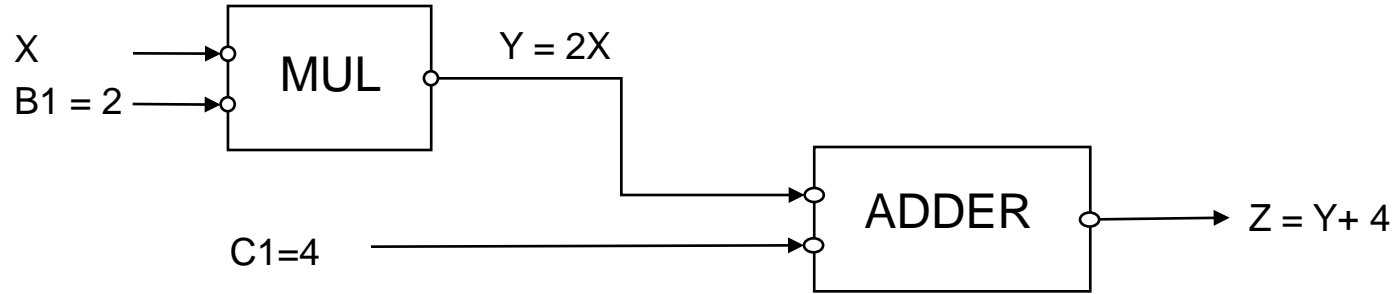
We get  $Z = 10 + 4 = 14$  ( $Z \leftarrow Y + 4$ )

But it is not true that  $X = 5$ .

We have to jettison the equation  $Y = 2X$ , because  $Y$  is no longer dependent on  $X$ .

- surgery or intervention

# Diagrams



The **causal relation** between  
the independent or manipulator variable  
and the dependent or manipulated variable  
is **best captured by diagrams.**

If X increases, so does Y, and in turn so does Z

# Deep Understanding : How things work

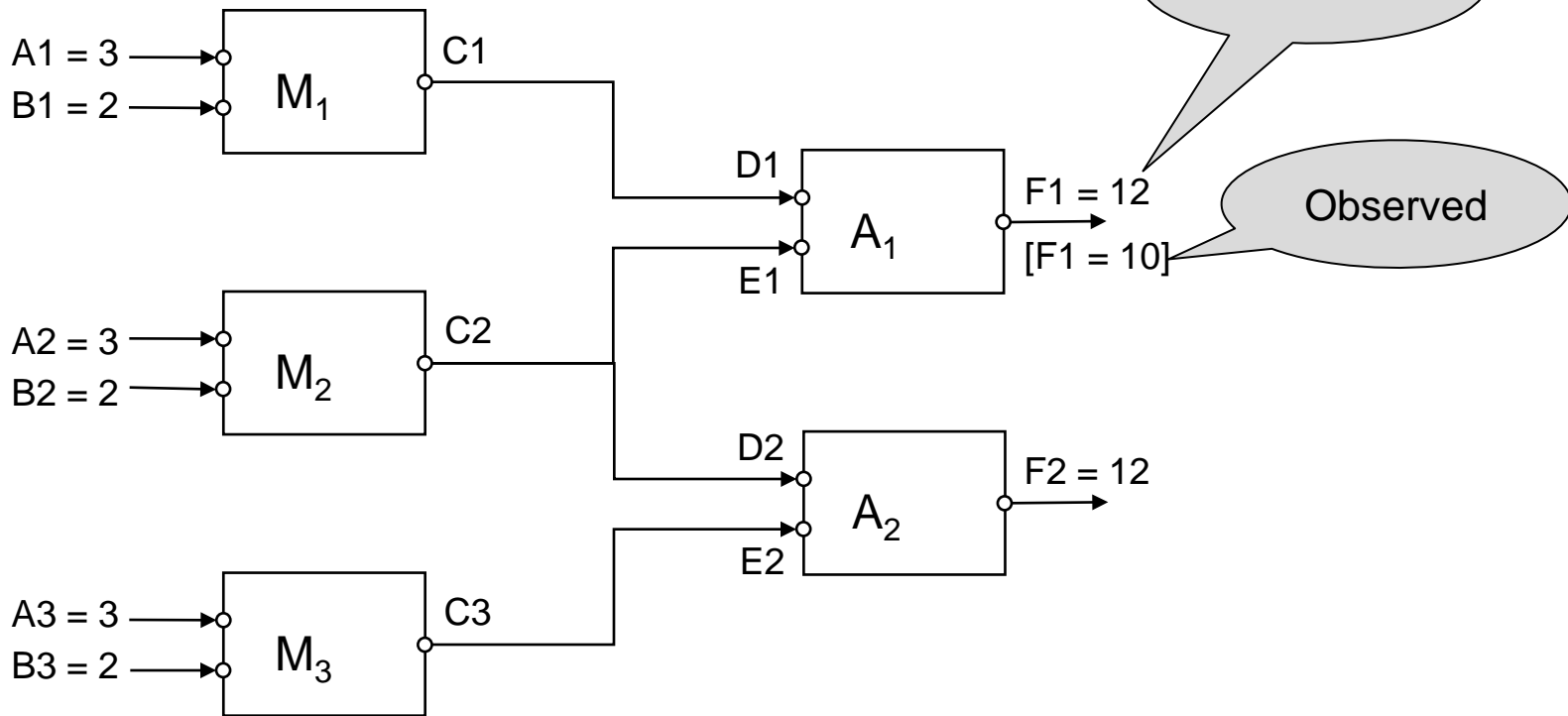
- “Deep understanding means knowing not merely how things behaved yesterday, but how things behave under new hypothetical circumstances.”
- When we have such understanding we feel “in-control” even if we have no practical way of controlling things.
  - we can predict the tidal effects of (say) the Moon being hit by a meteor, though we cannot control celestial events
- Causal models allow ‘deliberate’ reasoning
  - what-if scenarios
  - for example, what if we clamp  $Y$  to 10
- Engineers routinely do such reasoning

from *The Art and Science of Cause and Effect*, in *Causality*.

# Diagnosis = finding the cause

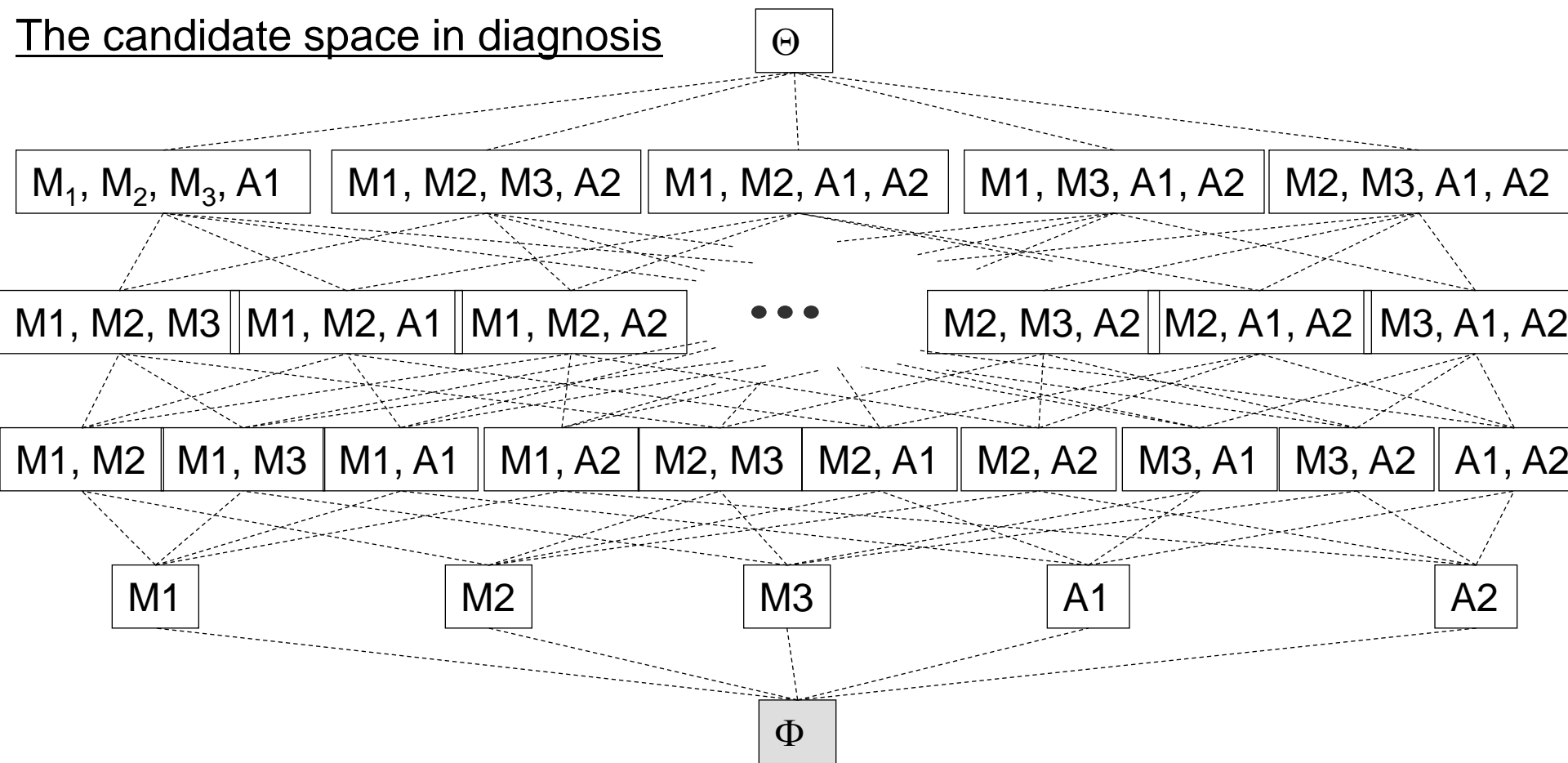
- Cause leads to Effect
  - for example, disease causes symptoms.
- Model based diagnosis uses deductive reasoning on a deep model of the artifact.
  - Adder multiplier example (next slide)
- Abductive reasoning attempts to infer the cause from the effect
  - The inference is not sound, and takes us into the realm of reasoning under uncertainty – often using probabilistic reasoning
  - Evidential reasoning strives to find the most likely hypothesis based on the available evidence. For example, the Dempster Shafer model.
    - Hypothesis: dinosaur extinction was due to a meteor striking the earth - proposed by Luis and Walter Alvarez in 1980.
    - Evidence: a meteor strike in Chicxulub in the Yucatan Peninsula at the appropriate time - Pälke [2013]

## A simple malfunctioning device



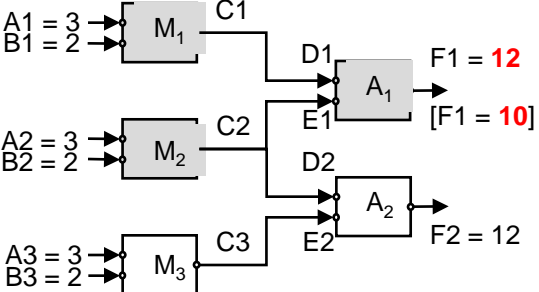
A simple device made of three multipliers and two adders. A fault has occurred because the observed value at  $F_1 = 10$  differs from the predicted value 12.

The candidate space in diagnosis



The *minimal candidate* when the device is working is the empty set.

# Inconsistency leads to conflict sets



The predicted value  $F1=12$  is based on the assumptions  $\neg Ab(M_1)$ ,  $\neg Ab(M_2)$  and  $\neg Ab(A_1)$ , which state that the corresponding components are working correctly and not broken.

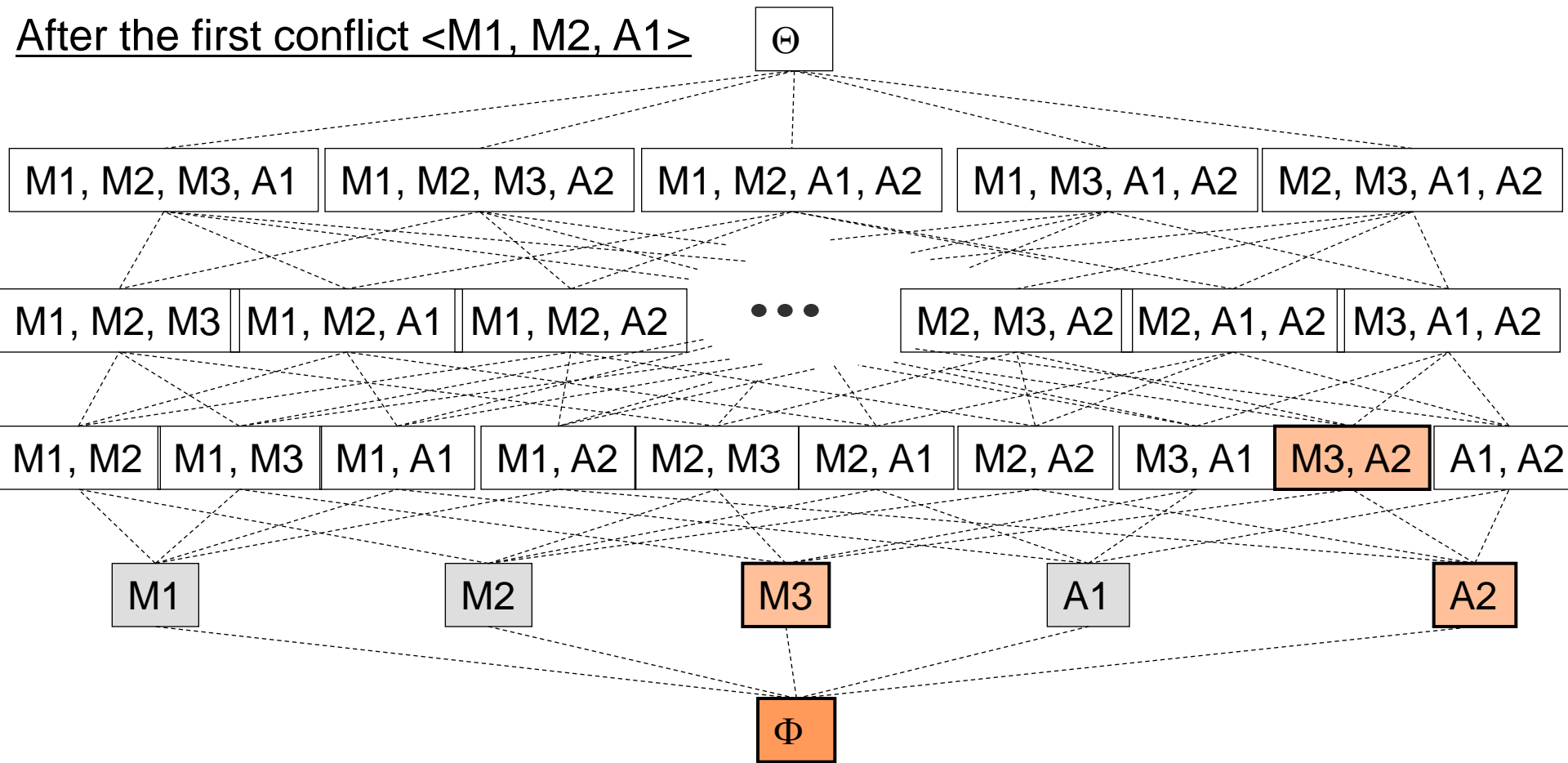
The other value  $F1=10$  is an observation, and is based on no assumption.

The cumulative assumptions from the two ways of arriving at the value are inconsistent together.

That is,  $\neg Ab(M_1)$ ,  $\neg Ab(M_2)$  and  $\neg Ab(A_1)$  cannot be true at the same time.

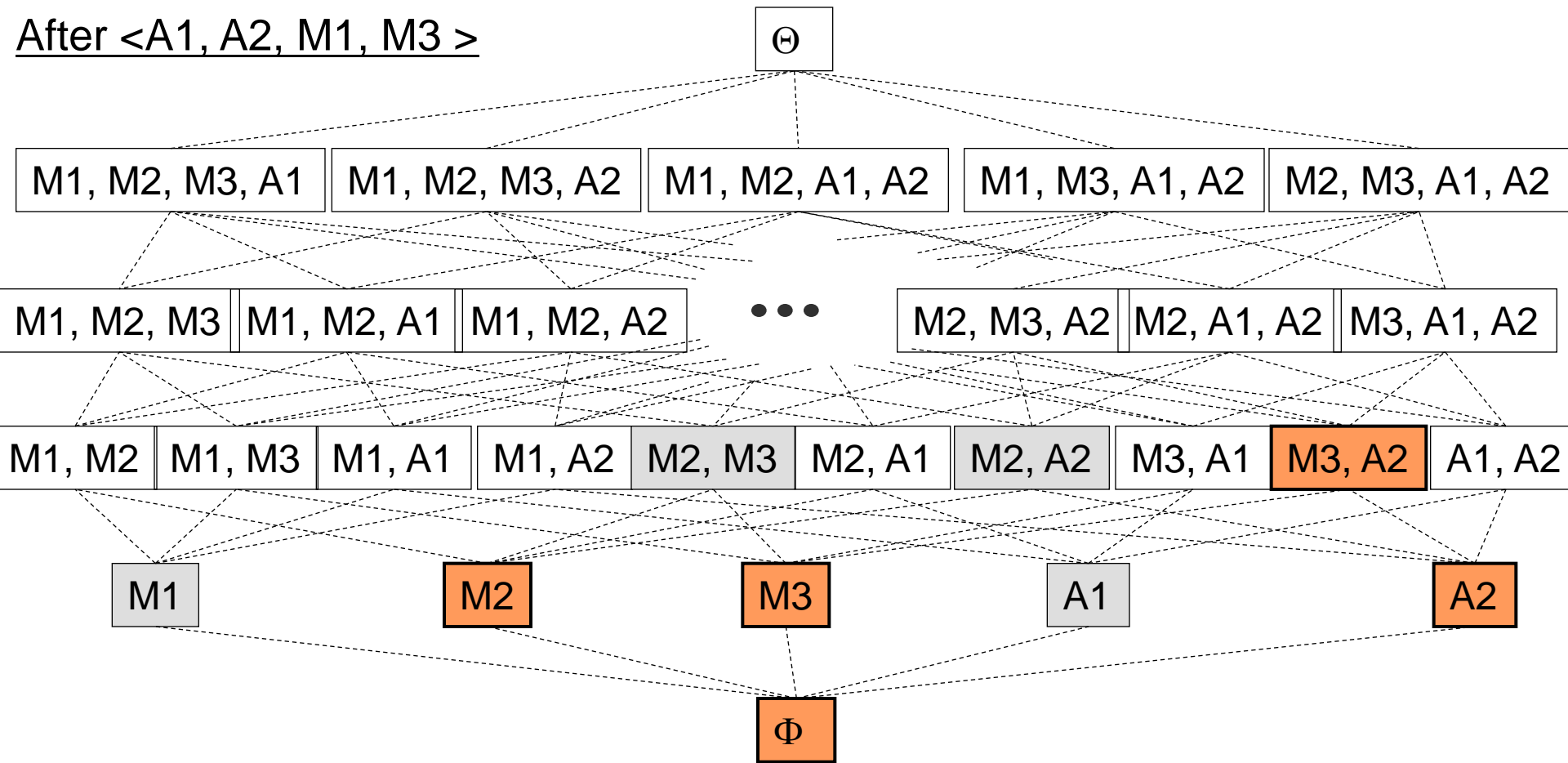
We represent this as the conflict  $\langle M_1, M_2, A_1 \rangle$ .

After the first conflict  $\langle M1, M2, A1 \rangle$

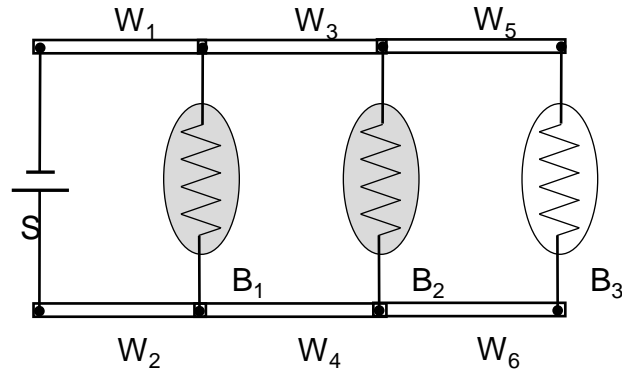


The minimal candidates are  $\{M1\}$ ,  $\{M2\}$ , and  $\{A1\}$ .  $\{\Phi\}$ ,  $\{M3\}$ ,  $\{A2\}$  and  $\{M3, A2\}$  eliminated.

After  $\langle A1, A2, M1, M3 \rangle$



## Fault Models



Three bulbs connected to a battery in parallel with six wire components. Bulbs  $B_1$  and  $B_2$  are not glowing.

The general diagnosis algorithm working without fault models generates *many* diagnosis, for example “the Battery  $S$  is broken and so is  $B_3$ ”.

The general algorithm cannot conclude that since  $B_3$  is lit *all* the wires and the battery *must* be okay.

# Logic, science, equations and causality

- Logic is concerned with valid arguments. It is like solving equations.
- Science often involves constructing causal models of the world.
- Russel's enigma – the clash between directionality of causal relations and the symmetry of physics equations
  - causal models have directionality. Does force create acceleration or does acceleration result in force being felt?
  - equations describe both in one model, but the distinction between the manipulator and manipulated vanishes.
- equations (and algebra) lead to computational facility (at the expense of causal reasoning)
  - they help analyze observations,
  - but do not support intervention (surgery in controlled experiments)

from *The Art and Science of Cause and Effect*, in *Causality*.

# Statistics and Probability

Francis Galton, inventor of fingerprinting and cousin of Charles Darwin, explored various ways of measuring how properties of one class of individuals or objects are related to those of another class.

In 1888, he measured the length of a person's forearm with the size of a person's head, and asked to what degree can one of these quantities predict the other.

*“It's easy to see,”* said Galton, *“that co-relation must be the consequence of the variations of the two organs being partly due to common causes.”*

His disciple, Karl Pearson, now considered to be the father of modern statistics, was profoundly impressed.

from *The Art and Science of Cause and Effect*, in *Causality*.

# Correlation vs. Causality

*“I interpreted....,”* said Pearson 45 years later, *“...Galton to mean that there was a category broader than causation, namely correlation, of which causation was only the limit, and that this new conception of correlation brought psychology, anthropology, medicine, and sociology in large parts into the field of mathematical treatment.”*

In his 1911 book *The Grammar of Science* he says, *“Beyond such discarded fundamentals as ‘matter’ and ‘force’ lies another fetish amidst the inscrutable arcana of modern science, namely, the category of cause and effect.”*

- he categorically denies a need for an independent concept of causal relation beyond correlation!

from *The Art and Science of Cause and Effect*, in *Causality*.

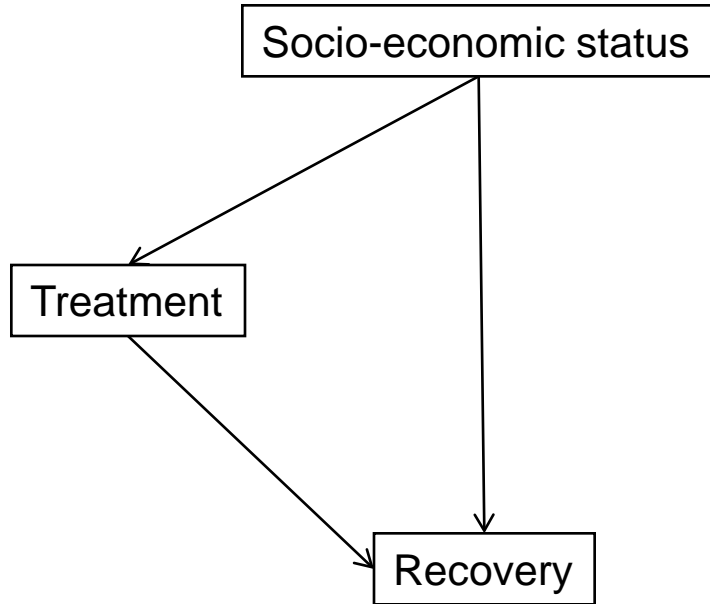
# Ronald Fisher and the randomized experiments

It took another 25 years for Ronald Fisher to devise the randomized experiment to investigate the cause-effect relationship from data.

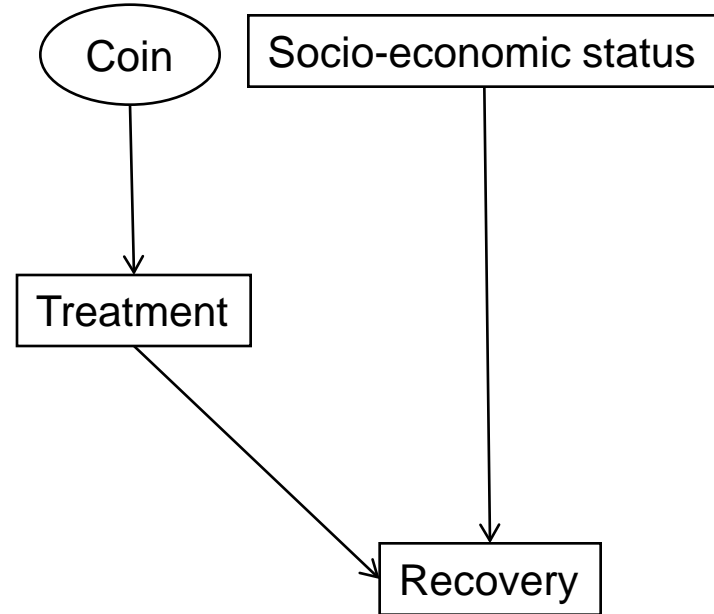
- Suppose we want to investigate whether a given treatment for an ailment is effective or not.
- If we do try the treatment on different subjects how do we know whether the recovery is *because* of the treatment, or due to some other factors like socio-economic status?
- We separate the subjects into similar groups, and then randomly give treatment to some, and a placebo to others.

# Intervention as surgery (controlled experiments)

Uncontrolled conditions



Experimental conditions



from *The Art and Science of Cause and Effect, in Causality.*

# Should we talk of causality or not?

- Philip Dawid, editor “Biometrika”, says “*Causal inference is one of the most important, most subtle, and most neglected of all problems in statistics.*”
- Terry Speed, former President of Biometric Society – “*..causality in statistics... preferably not at all, but if necessary, with great care.*”
- David Cox and Nanny Wermoth – “*We do not in this book use the word causal or causality...*”

Judea Pearl believes that this reticence has to do with “*the official language of statistics, namely the language of probability.*”

The word *cause* is not in the language of probability theory. We cannot express a sentence like “*the mud does not cause rain*”.

## Pearl's own views

Judea Pearl himself started off, while writing the book on Probabilistic Reasoning, with an *“empiricist tradition in which it is held that probabilistic relationships constitute the foundations of human knowledge”*.

In *Causality* he writes - *“Today my view is quite different. I now take causal relationships to be the fundamental building blocks both of physical reality and of human understanding of that reality, and I regard probabilistic relationships as but the surface phenomena of the causal machinery that underlies and propels our understanding of the world.”*

Further *“Accordingly, I see no greater impediment to scientific progress than the prevailing practice of all our mathematical resources on probabilistic and statistical inferences while leaving causal inferences to the mercy of intuition and good judgment.”*

from *The Art and Science of Cause and Effect*, in *Causality*.

# Probability

- Suppose you hear '12' from the next table at the casino. Are they rolling a pair of dice, or are they on a roulette wheel?
- Compute  $P(\text{dice} \mid 12)$  and  $P(\text{roulette} \mid 12)$ 
  - hard to estimate directly
- Easier to compute  $P(12 \mid \text{dice}) = 1/36$  and  $P(12 \mid \text{roulette}) = 1/38$
- Also  $P(\text{dice})$  and  $P(\text{roulette})$  can be estimated by doing a survey of the casino.
- Bayes rule now allows you to decide which of the two is more likely.
  - $P(\text{dice} \mid 12)/P(\text{roulette} \mid 12) = P(12 \mid \text{dice})P(\text{dice}) / P(12 \mid \text{roulette})P(\text{roulette})$

## Rain and wet lawns

What is the probability that it rained given that the grass is wet?

$$P(\text{rain} \mid \text{wet}) - P(\text{“happened” rain} \mid \text{“I see” wet})$$

$$P(\text{rain} \mid \text{see}\{\text{wet}(\text{grass})\}) = P(\text{wet} \mid \text{rain})P(\text{rain})/P(\text{wet})$$

But what if I had watered my lawn. The language of probability does **not** allow us to talk of **doing** things!

One should be able to say something like

$$P(\text{rain} \mid \text{do}\{\text{wet}(\text{grass})\}) = P(\text{rain})$$

--- a Do-algebra!

# Pearl's rules for Causal Calculus

Rule 1: Ignoring observations

$$P(y \mid \text{do}\{x\}, z, w) = P(y \mid \text{do}\{x\}, w)$$

ignore an irrelevant observation

Rule 2: Action/observation exchange

$$P(y \mid \text{do}\{x\}, \text{do}\{z\}, w) = P(y \mid \text{do}\{x\}, z, w)$$

replace an action with an observation the same fact

Rule 3: Ignoring actions

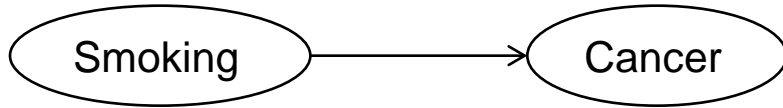
$$P(y \mid \text{do}\{x\}, \text{do}\{z\}, w) = P(y \mid \text{do}\{x\}, w)$$

ignore an irrelevant action

# Smoking and Cancer

In 1964 the Surgeon General issued a report linking cigarette smoking to cancer. The report was based on nonexperimental studies in which a strong correlation was found between smoking and lung cancer.

The claim was that the correlation found was causal, and hence one should ban smoking.

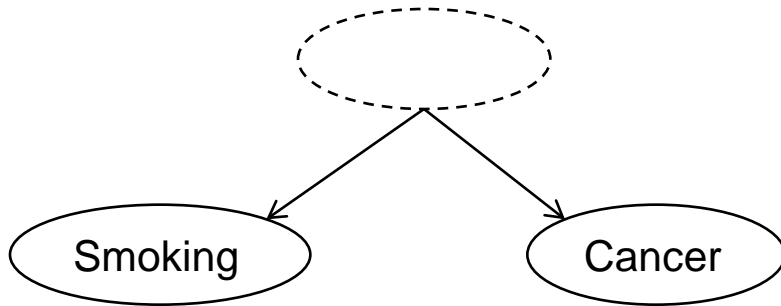


$$P(c \mid \text{do}\{s\}) = P(c \mid s)$$

# Tobacco Industry

The Surgeon General's report came under severe attack from the tobacco industry, backed by some very prominent statisticians, among them Sir Ronald Fisher.

The claim was that the correlation observed can be explained by a model in which there is no causal connection. Instead, an unobserved genotype might exist that simultaneously causes cancer and produces an inborn craving for nicotine.



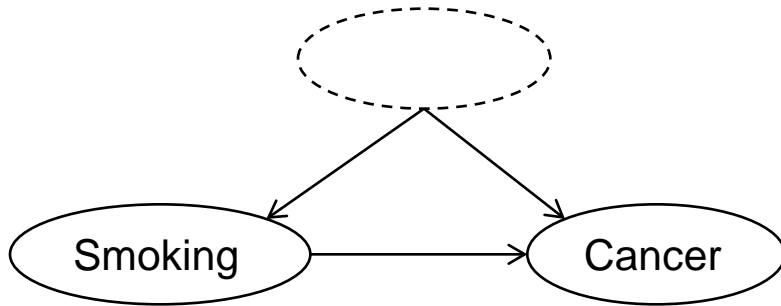
$$P(c \mid \text{do}\{s\}) = P(c)$$

from *The Art and Science of Cause and Effect*, in *Causality*.

## Could both - genetic and smoking – be the cause?

Supposing both sides concede that perhaps either or both could be a cause of cancer.

The strength of the edges then cannot be computed.



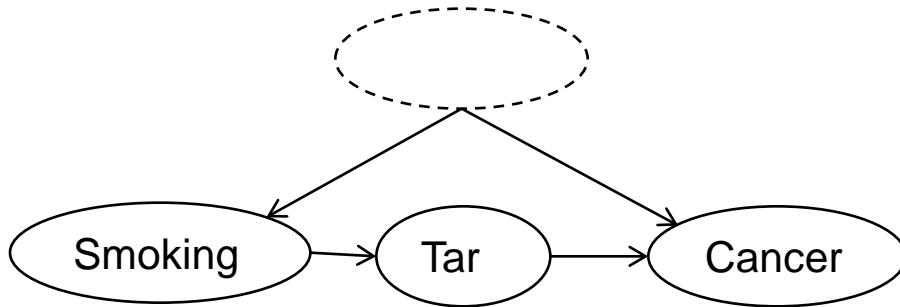
$$P(c \mid \text{do}\{s\}) = \text{noncomputable}$$

# Can we measure tar deposits?

Since the hypothesis is that smoking affects lung cancer through the accumulation of tar deposits, one could measure the tar deposits, and this might provide the necessary information for quantifying the links?

It turns out that this is possible, because we can ignore the  $do\{s\}$  action, and convert the  $do\{tar\}$  deposit to  $see\{tar\}$  deposit.

---- we skip the details



$$P(c \mid do\{s\}) = \text{computable}$$

## Simpson's paradox

First observed by Karl Pearson in 1899 – Every statistical relationship between two variables may be *reversed* by including additional factors in the analysis.

Students who smoke get higher grades,  
but if you consider *age*,  
then in *every age group* smoking predicts lower grades!

If you additionally consider *parental income*,  
the relationship *reverses again!*

The Adjustment Problem – which factors to consider?

# Graduate admissions

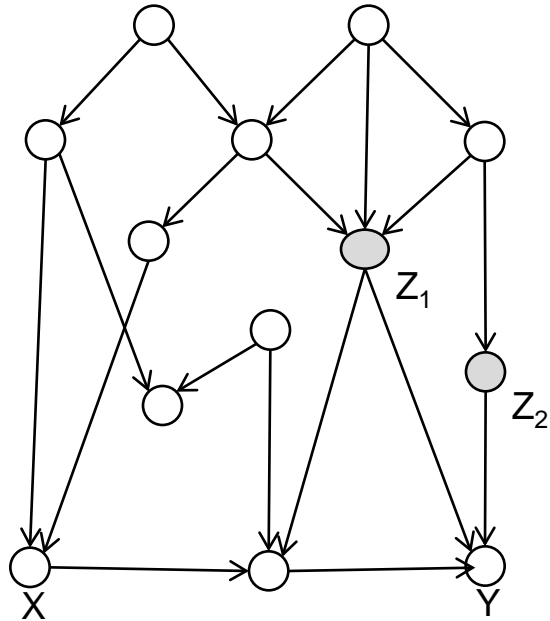
In 1975, UC-Berkley was investigated for gender bias in graduate admissions.

Overall data showed a higher rate of admission for male applicants, but, broken down into departments, data showed a slight bias in favour of admitting female applicants.

Explanation: Female applicants tend to apply for more competitive departments than males.

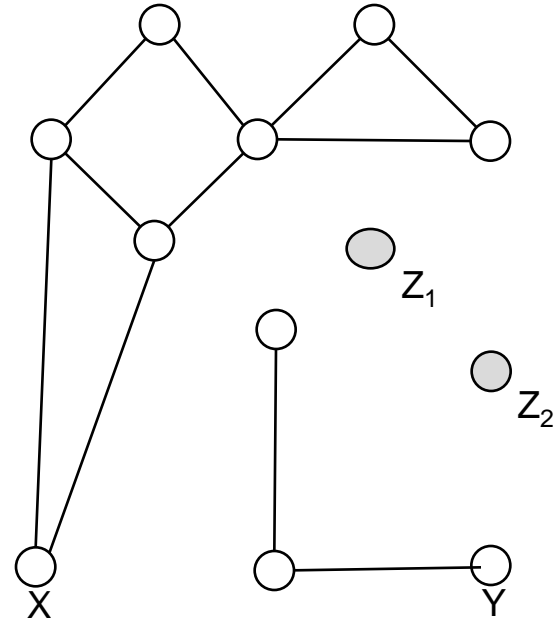
# A graphical solution

Which parameters to consider to determine the effect of X on Y?



Are  $Z_1$  and  $Z_2$  sufficient measures?

After  
a  
series  
of  
graph  
operations



Yes, if X is disconnected from Y in the resulting graph.

## Conclusion

Need a language  
for talking about causal relations  
and  
graphical representations  
to visualize the relations and do the reasoning

Thank you