





# Machines Climbing Pearl's Ladder of Causation

Devendra Singh Dhami  
Matej Zečević  
Adele Ribeiro



Machines Climbing Pearl's Ladder of Causation

# A Quick Introduction



- **Devendra Singh Dhami**
  - Independent Research Group Leader at TU Darmstadt and Hessian Center for Artificial Intelligence (hessian.AI)



- **Matej Zečević**
  - PhD student at the Artificial Intelligence and Machine Learning Lab at TU Darmstadt



- **Adèle Ribeiro**
  - Postdoctoral Researcher at Philipps-Universität Marburg

# 5 Thrilling Days

- Day 1: Introduction to Causality (Devendra)
- Day 2: Representation of Causal Knowledge and Causal Discovery (Adèle)
- Day 3: Causal Identification and Estimation (Matej)
- Day 4: Hands-on Session (Matej)
- Day 5: Current state of Research in Causality (Devendra and Adèle)



# The Ultimate Hitchhiker's Introduction to Causality for ML

Devendra Singh Dhami



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

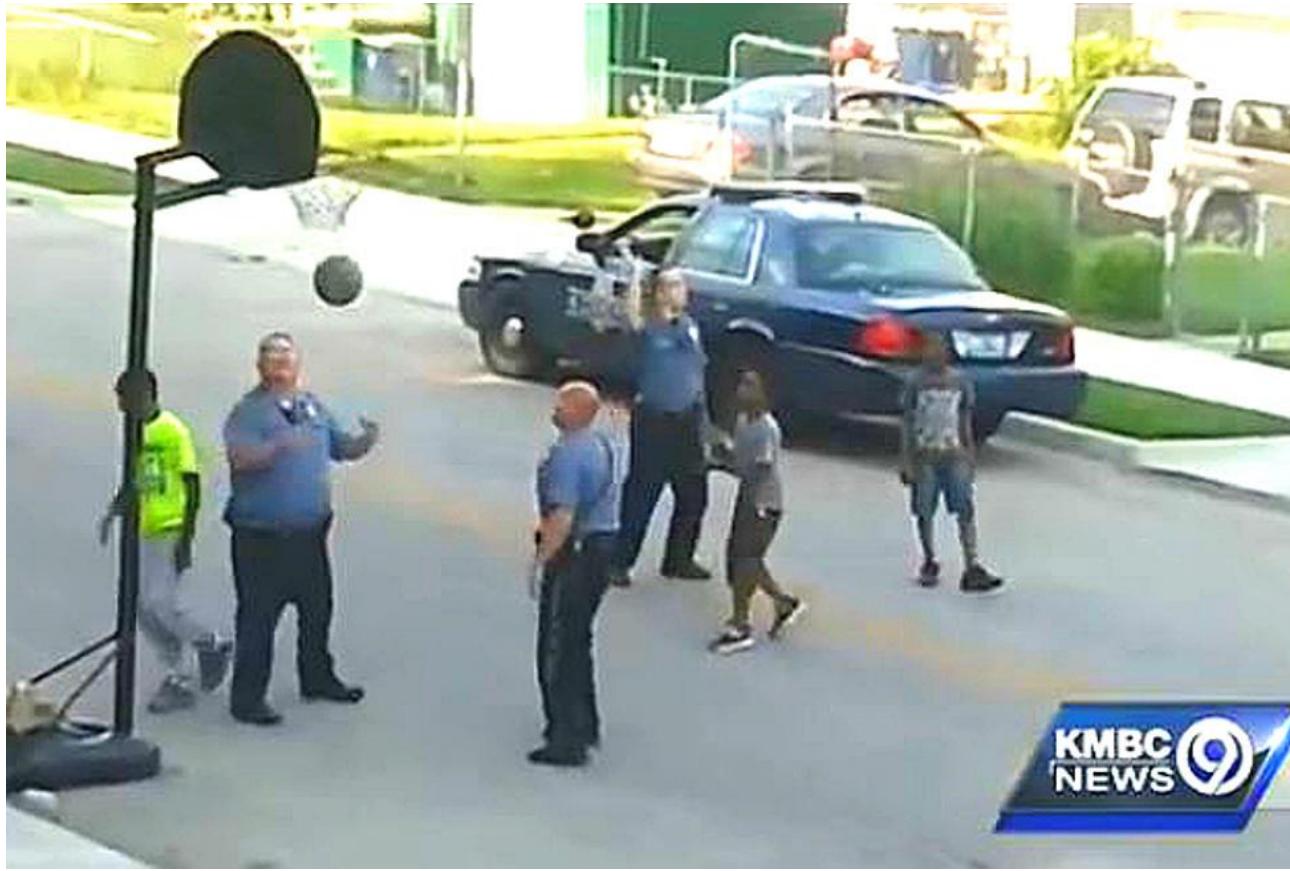


KMBC  
NEWS 9



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation



KMBC  
NEWS 9



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Causality is hard

$$1+1=2$$

$$2-1=1$$

# Causality is hard(er)



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

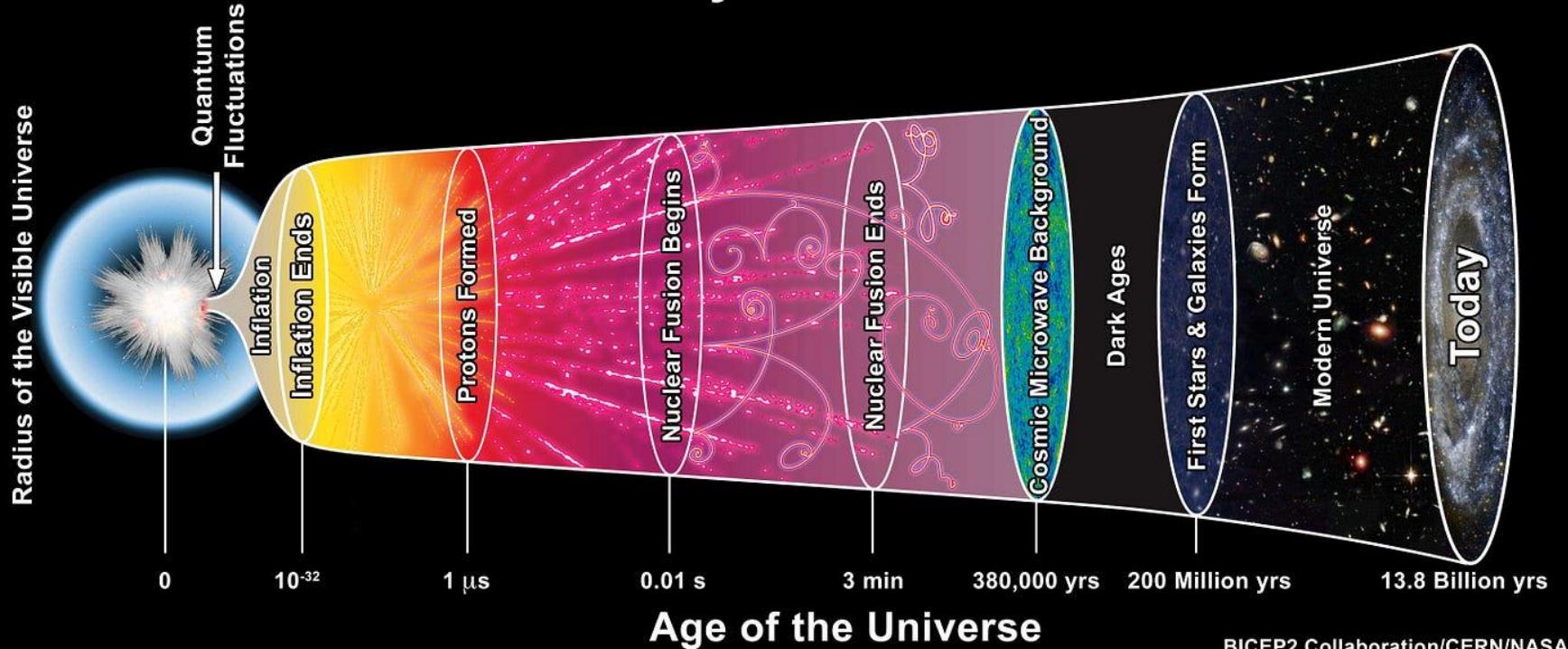
Machines Climbing Pearl's Ladder of Causation

# Causality is hard(er)



Machines Climbing Pearl's Ladder of Causation

# History of the Universe



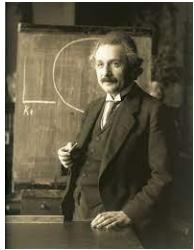
**Carl Sagan:** If you wish to make an apple pie from scratch, you must first invent the universe



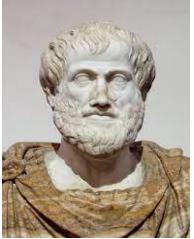
**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Causality is Omnipresent!

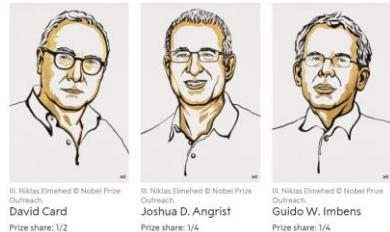


**Classical Physics:** An effect cannot occur from a cause that is not in the back (past) light cone of that event. Similarly, a cause cannot have an effect outside its front (future) light cone



**Philosophy:** The material cause of a being is its physical properties or makeup. The formal cause is the structure or direction of a being. The efficient cause is the thing or agent, which actually brings it about. And the final cause is the ultimate purpose for its being

The Sveriges Riksbank Prize in  
Economic Sciences in Memory of  
Alfred Nobel 2021



**Economics:** Develops explicit models of outcomes where the causes of effects are investigated and the mechanisms governing the choice of treatment are analyzed

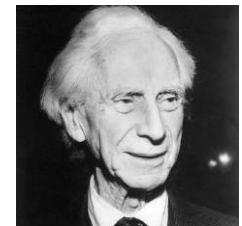
# Causality is Omnipresent?

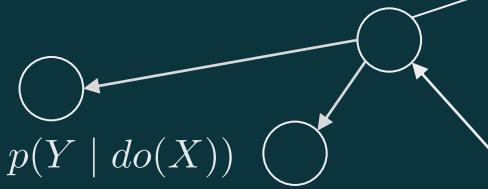
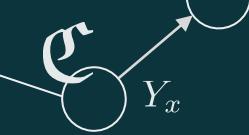
**David Hume:** Causation is a relationship between two impressions or ideas in the mind. Because causation is defined by experience, any cause-and-effect relationship could be incorrect because thoughts are subjective and therefore causality cannot be proven



**Barber's Paradox:** There is a village where the barber shaves all those and only those who do not shave themselves. Who shaves the barber?

A causal problem → Are the non-shavers actually non-shavers because they're shaved by the barber or whether they're shaved by the barber because they're non-shavers. If the first is true, the barber is a member of both groups: he shaves himself and is shaved by the barber

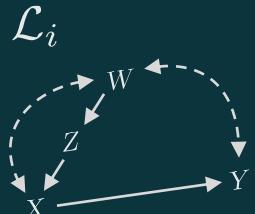




I

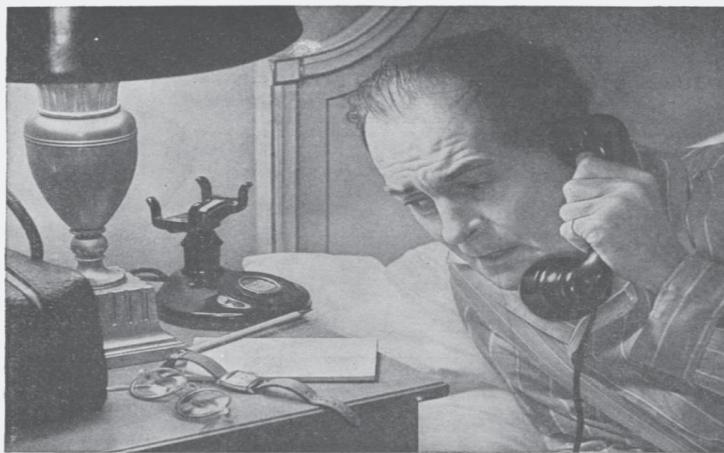
Why?

Machine Learning needs Causality?



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation



## "I'll Be Right Over!"

*... 24 hours a day your doctor  
is "on duty"... guarding  
health... protecting and  
prolonging life...*

• Plays... novels... motion pictures... have been written about the "man in white." But in his daily routine he lives more drama, and displays more devotion to the oath he has taken, than the most imaginative mind could ever invent. And he asks no special credit. When there's a job to do, he does it. A few winks of sleep... a few puffs of a cigarette... and he's back at that job again...



R. J. Reynolds Tobacco Company, Winston-Salem, N.C.

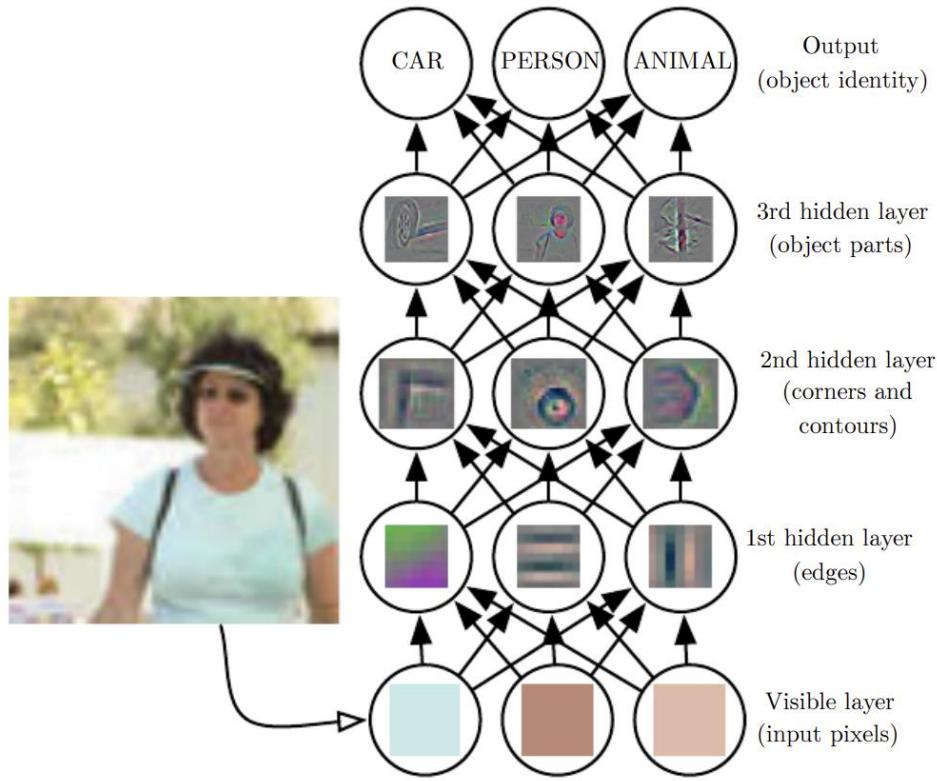
According to a  
recent independent  
nationwide survey:  
**More Doctors  
Smoke Camels  
than any other cigarette**



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Current Trend: Deep Learning



Goodfellow et al., Deep Learning, MIT Press, 2017

# Current Trend: Deep Learning

DE

an example causal graph for unexperienced audience



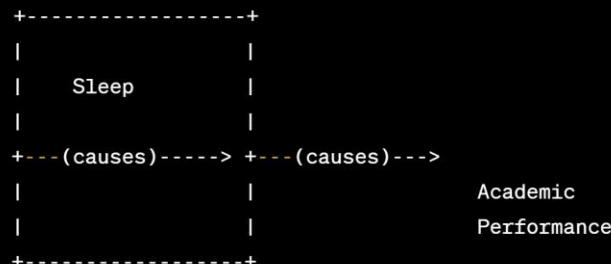
Sure! Let's create a simple causal graph with an example that is easy to understand for an inexperienced audience. Suppose we want to demonstrate the relationship between sleep and academic performance in students.

## Example Causal Graph: Sleep and Academic Performance

In this example, we have two variables: "Sleep" and "Academic Performance." We will use arrows to represent causal relationships between these variables.

SCSS

Copy code



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Current Trend: Deep Learning

Explanation:

1. Sleep: This variable represents the number of hours a student sleeps each night.
2. Academic Performance: This variable represents a student's academic performance, which can be measured by their grades or exam scores.

Causal Relationship:

The arrow pointing from "Sleep" to "Academic Performance" indicates a causal relationship.

In this example, we are hypothesizing that sleep directly affects academic performance.

More sleep may lead to better academic performance, while less sleep may result in poorer academic performance.

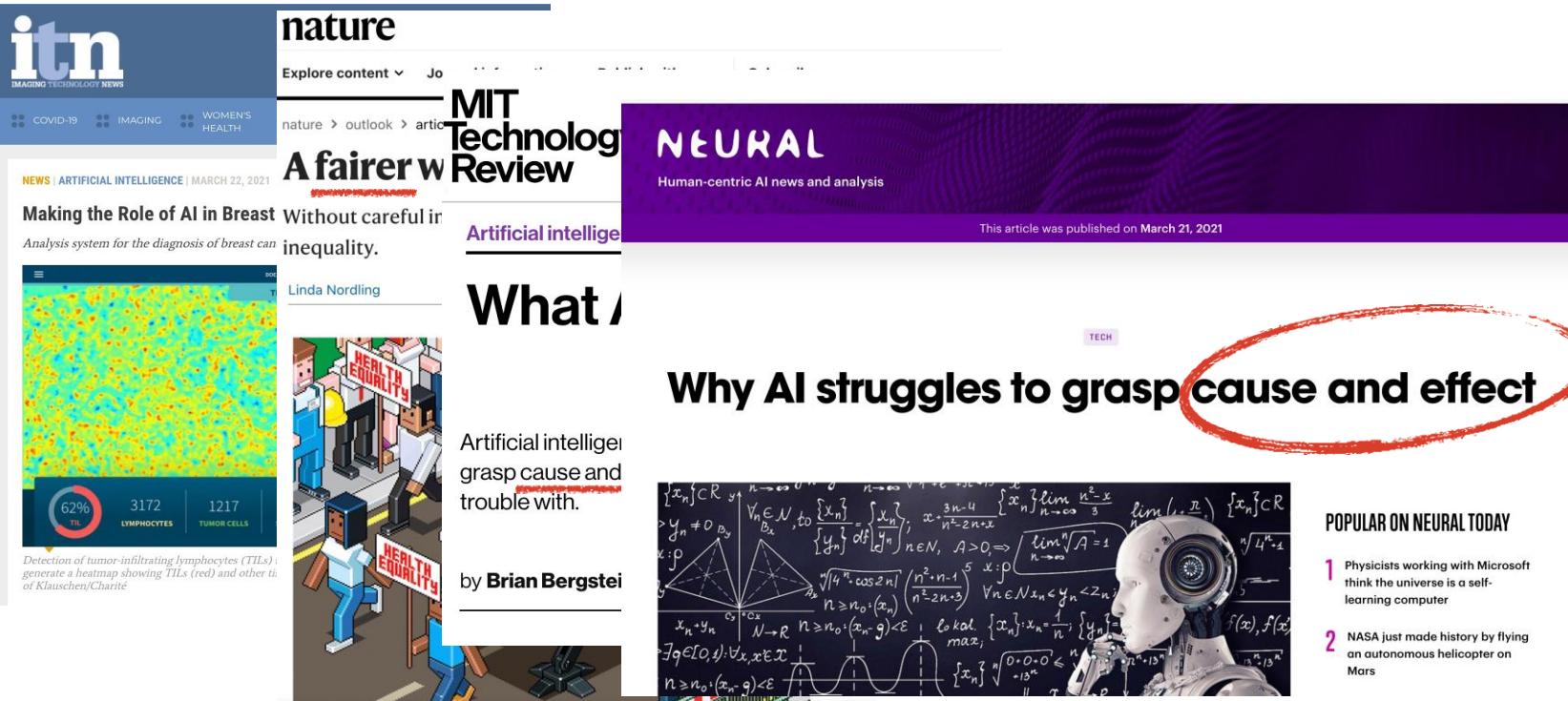
# Current Trend: Deep Learning



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Current Challenges in DS, AI, ML



## Machines Climbing Pearl's Ladder of Causation

# Current Challenges in DS, AI, ML

- Data-hungry & sample inefficiency
- Lack of interpretability & explainability
- Lack of robustness & generalizability
- Unfair & unethical decision-making

Lack of Causal Inference Capabilities



5

# What can we achieve with causality?

**Data Fusion:** provides language and theory to cohesively combine prior knowledge and data from multiple and heterogeneous studies.

**Effect identifiability:** can determine the effect of unrealized interventions rather than just predicting an outcome (i.e., can distinguish between association and causation)

**Generalizability:** allows the transportability of causal effects across different domains.

**Explainability:** provides a better understanding of the underlying mechanisms.

**Fairness:** captures and disentangles any mechanisms of discrimination that may be present, including direct, indirect-mediated, and indirect-confounded.

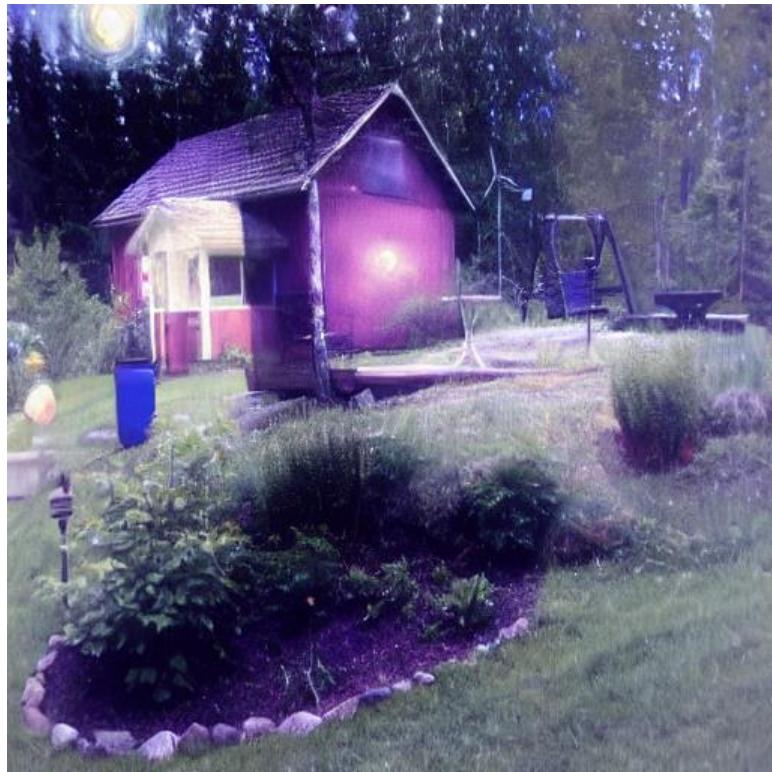


Prompt: Make it a dark **start-lit** night.



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation



Prompt: Make it a dark star-lit night.



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation



Prompt: Make it a dark **star-lit** night.

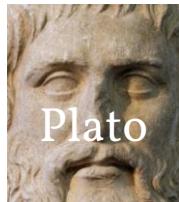


ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# What is Causality?

We might want to start here first..

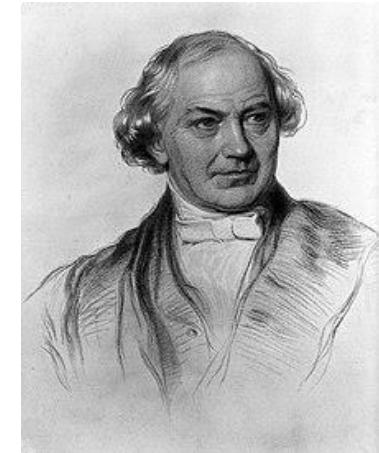


Probably, Plato was the first to state the principle of causality:

“Everything that becomes or changes must do so owing to some cause; for nothing can come to be without a cause.” - *Timaeus* 28a

# Axioms of Causality

- Nothing takes place without a cause.
- The magnitude of an effect is proportional to the magnitude of its cause.
- To every action there is an equal and opposed reaction.



Example: if a baseball is moving through the air, it must be moving this way because of a previous interaction with another object, such as being hit by a baseball bat

**Principle of Universal Causation:** Every phenomenon has a cause, which it invariably follows; and from this are derived other invariable sequences among the successive stages of the same effect, as well as between the effects resulting from causes which invariably succeed one another

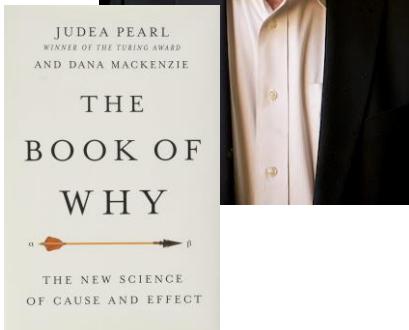
# Judea Pearl's opinion

Pioneer of Causality for AI, Turing awardee

“To Build Truly Intelligent Machines,  
Teach Them Cause and Effect”

“All the impressive achievements of deep learning  
amount to just curve fitting”

Judea Pearl in “The Book of Why”  
and in an interview with quanta magazine in 2018



Machines Climbing Pearl's Ladder of Causation

# Yoshua Bengio's opinion

Pioneer of Deep Learning, Turing awardee



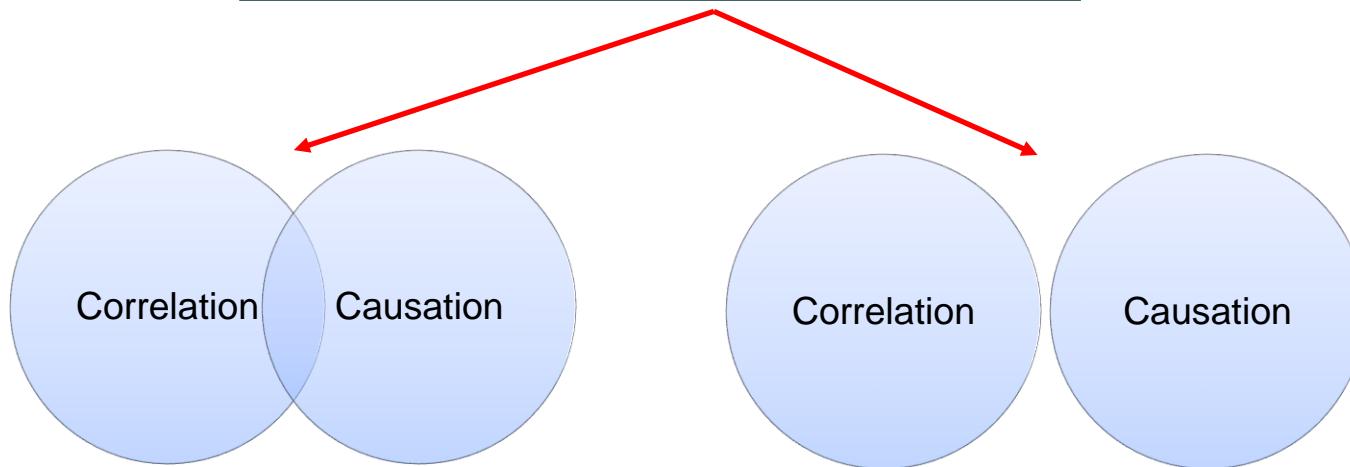
“..for conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing.”

“Causality is very important for the next steps of progress of machine learning,”

Yoshua Bengio in an interview with IEEE Spectrum, 2020

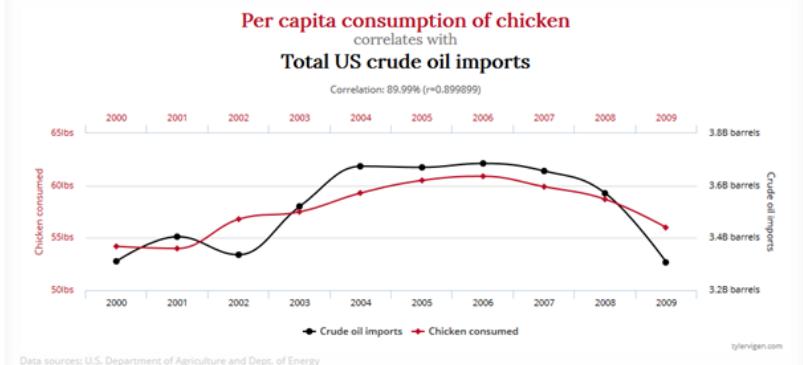
# We have All Heard the Phrase

***“Correlation does not imply causation”***

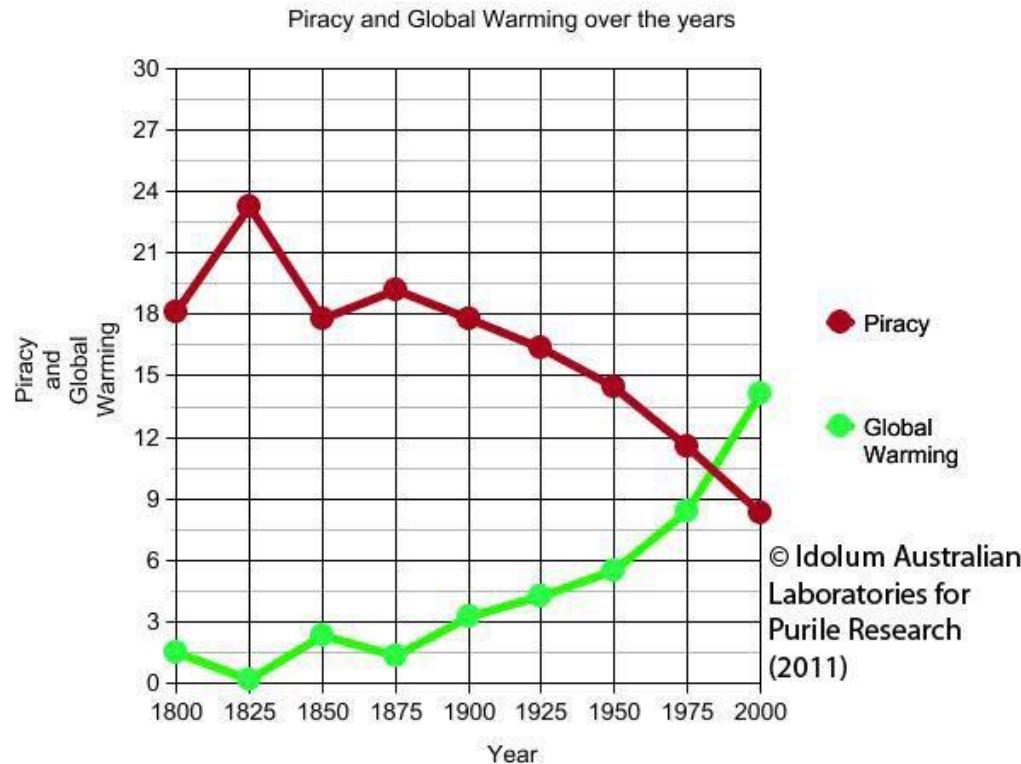


Pic Credit: Wikipedia

# Correlation ∩ Causality = Ø



# Correlation => Causality via a third factor



DE

what is the effect of global warming on pirate ships

edit



As of my last update in September 2021, there is no direct scientific evidence or research specifically focused on the effect of global warming on pirate ships. Global warming primarily refers to the long-term increase in Earth's average surface temperature due to human activities, such as the emission of greenhouse gases like carbon dioxide.



However, we can speculate on some potential indirect effects that global warming might have on pirate ships:

1. Rising sea levels: Global warming contributes to the melting of polar ice caps and glaciers, leading to rising sea levels. Higher sea levels can alter coastlines and navigational routes, making it challenging for pirate ships to access certain areas or ports.
2. Changes in weather patterns: Global warming can lead to shifts in weather patterns, including more intense storms, hurricanes, and cyclones. These extreme weather events can be dangerous for pirate ships at sea and may increase the risk of accidents or shipwrecks.
3. Ocean acidification: The increased concentration of carbon dioxide in the atmosphere not only causes global warming but also results in more CO<sub>2</sub> being absorbed by the oceans. This process leads to ocean acidification, which can have implications for marine life, including the plankton and other organisms that form the base of the marine food chain. Changes in marine ecosystems could indirectly impact pirate activity if their traditional hunting grounds are affected.
4. Temperature changes: While global warming generally leads to higher temperatures globally, localized effects can vary. Some regions may experience altered wind patterns and temperature shifts, which can influence the timing and strength of monsoons, trade winds, and other maritime conditions that pirates historically relied on for navigation.
5. Economic and geopolitical impacts: Global warming can disrupt economies and lead to resource scarcity in certain regions. These changes may influence trade routes and shift the focus of piracy to areas that were previously less affected.

It's essential to note that these are speculative effects, and any actual impact would likely depend on various factors such as the severity and pace of global warming, regional climate changes, and human responses to these challenges.



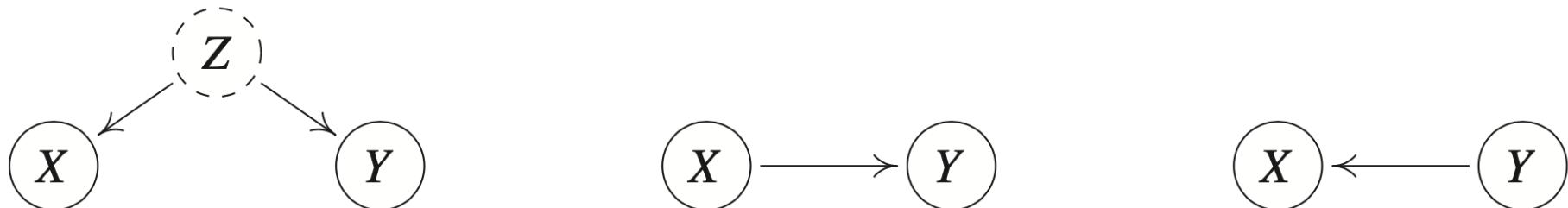
ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Reichenbach's Common Cause Principle

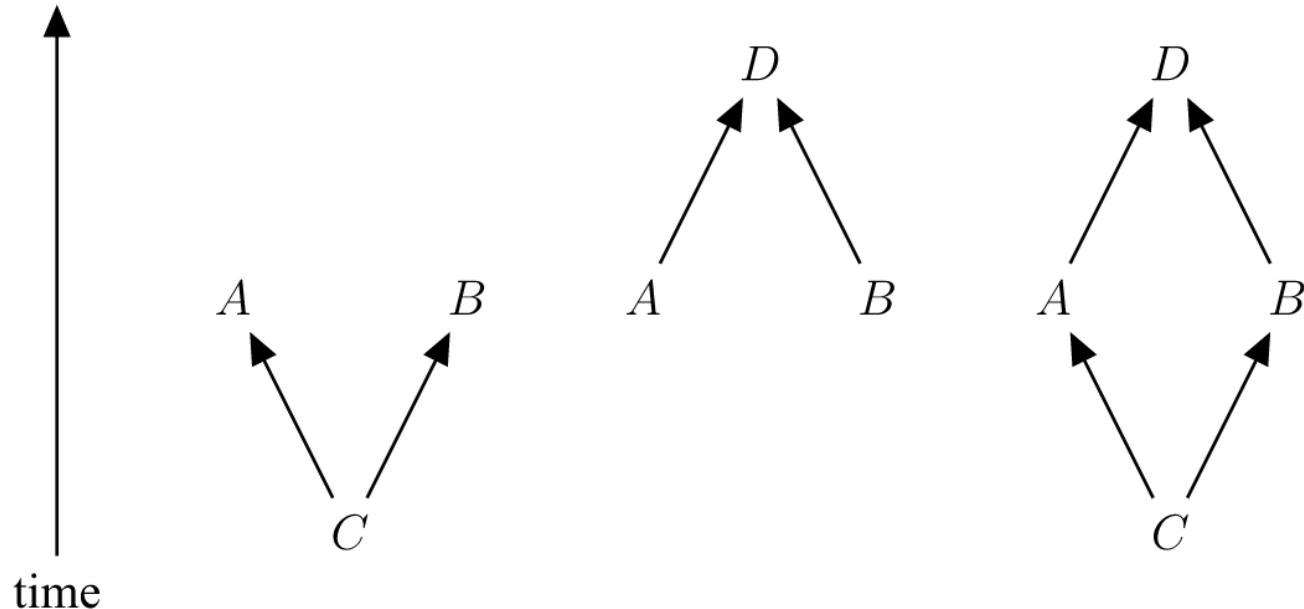
## Defining Confounders

**Principle 1.** *If two random variables  $X$  and  $Y$  are statistically dependent ( $X \neq Y$ ), then there exists a third variable  $Z$  that causally influences both. (As a special case,  $Z$  may coincide with either  $X$  or  $Y$ .) Furthermore, this variable  $Z$  screens  $X$  and  $Y$  from each other in the sense that given  $Z$ , they become independent,  $X \perp Y | Z$ .*



Reichenbach's *Direction of Time* (1956)

# Reichenbach's Common Cause Principle



Conjunctive fork:    a) open to the future    b) open to the past    c) closed fork

<https://plato.stanford.edu/entries/physics-Rpcc/>

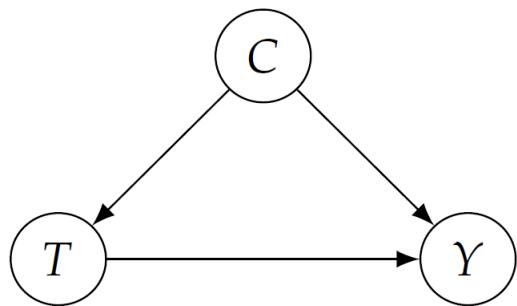
# Let's Illustrate Correlation does not imply causation: Simpson's Paradox

Treatment	Condition		
	Mild	Severe	Total
A	15% (210/1400)	30% (30/100)	<b>16%</b> (240/1500)
B	<b>10%</b> (5/50)	<b>20%</b> (100/500)	19% (105/550)

more effective treatment is completely dependent on the  
**causal structure** of the problem

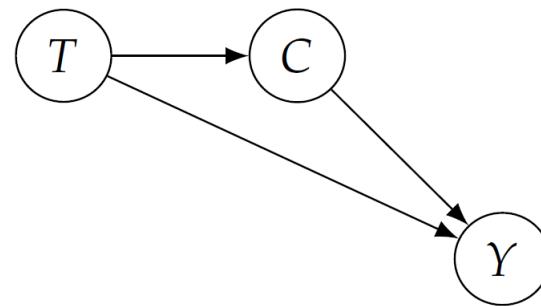
# Let's Illustrate Correlation does not imply causation : Simpson's Paradox

Scenario 1: Confounders



Treatment A preferable

Scenario 2: Treatment causes condition



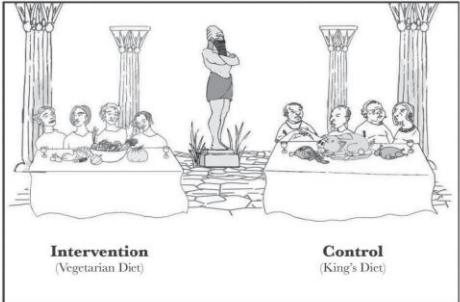
Treatment B preferable

# From Neyman-Rubin to Pearl

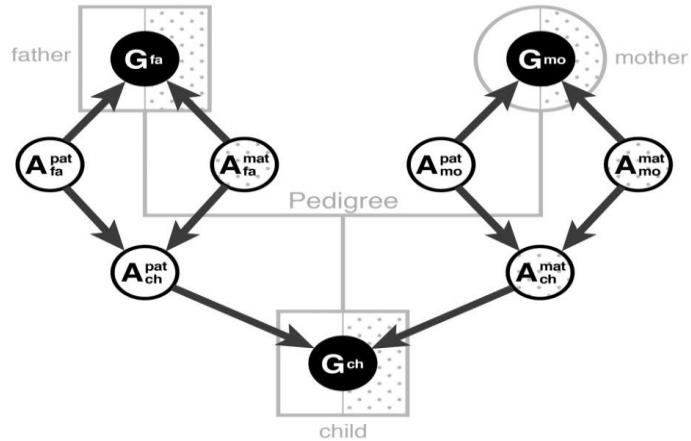
- Potential Outcome theory

to

Bayesian networks and DAG's



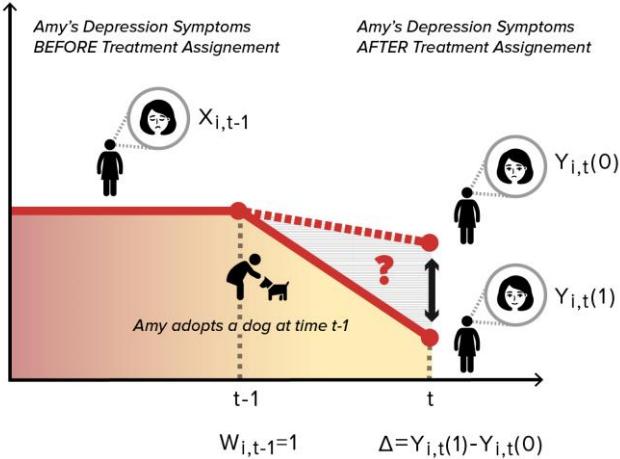
The biblical story of Daniel, often cited as the first controlled experiment. Daniel (third from left?) realized that a proper comparison of two diets could only be made when they were given to two groups of similar individuals, chosen in advance. King Nebuchadnezzar (rear) was impressed with the results. (Source: Drawing by Dakota Harr.)



Nodes of network:

- G** Genotype (observed in DNA test)
- A<sup>pat</sup>** Allele, paternal (unobservable)
- A<sup>mat</sup>** Allele, maternal (unobservable)

# Fundamental Problem of Causal Inference



Credit: Dominici et al., From Controlled to Undisciplined Data: Estimating Causal Effects in the Era of Data Science Using a Potential Outcome Framework

## Definition 1: Individual treatment effect

The individual treatment effect,  $\delta_i$ , equals  $Y_i^1 - Y_i^0$

## Definition 3: Switching equation

An individual's observed health outcomes,  $Y$ , is determined by treatment assignment,  $D_i$ , and corresponding potential outcomes:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$$
$$Y_i = \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases}$$

## Definition 2: Average treatment effect (ATE)

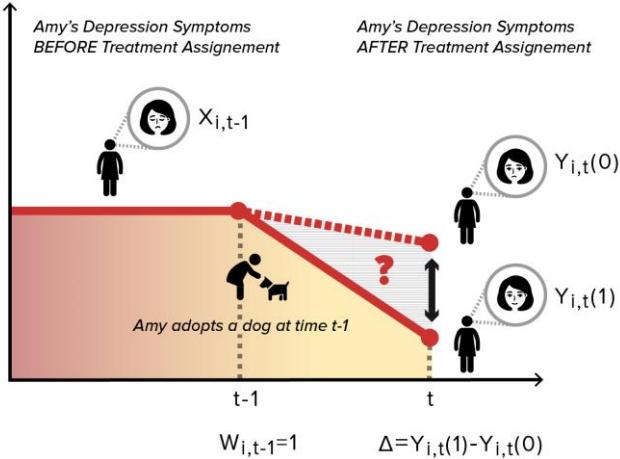
The average treatment effect is the population average of all  $i$  individual treatment effects

$$\begin{aligned} E[\delta_i] &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned}$$

What is the Apparent Problem?

Credit: Scott Cunningham

# Fundamental Problem of Causal Inference



Credit: Dominici et al., From Controlled to Undisciplined Data: Estimating Causal Effects in the Era of Data Science Using a Potential Outcome Framework

## Definition 1: Individual treatment effect

The individual treatment effect,  $\delta_i$ , equals  $Y_i^1 - Y_i^0$

## Definition 3: Switching equation

An individual's observed health outcomes,  $Y$ , is determined by treatment assignment,  $D_i$ , and corresponding potential outcomes:

$$Y_i = D_i Y_i^1 + (1 - D_i) Y_i^0$$
$$Y_i = \begin{cases} Y_i^1 & \text{if } D_i = 1 \\ Y_i^0 & \text{if } D_i = 0 \end{cases}$$

## Definition 2: Average treatment effect (ATE)

The average treatment effect is the population average of all  $i$  individual treatment effects

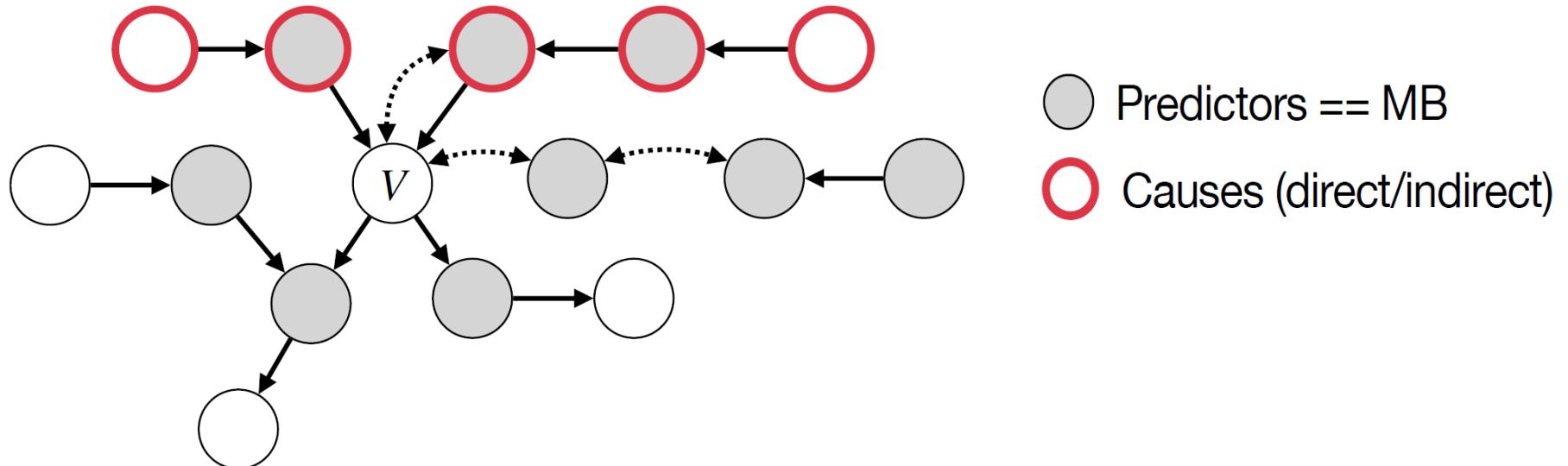
$$\begin{aligned} E[\delta_i] &= E[Y_i^1 - Y_i^0] \\ &= E[Y_i^1] - E[Y_i^0] \end{aligned}$$

## Definition 4: Fundamental problem of causal inference

It is impossible to observe both  $Y_i^1$  and  $Y_i^0$  for the same individual and so individual causal effects,  $\delta_i$ , are unknowable.

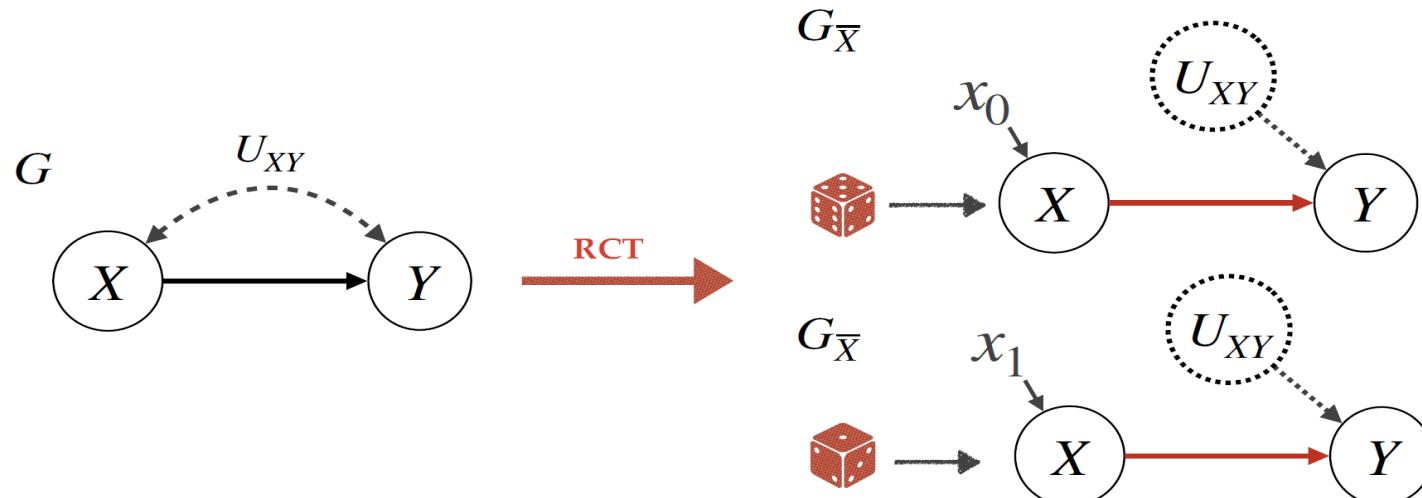
Credit: Scott Cunningham

# Graphically Explaining Causes and Predictors



# Randomized Experiments

Randomized Experiments / Control Trials (e.g. RCT) allow the identification of causal effects by leveraging randomization of the treatment assignment.



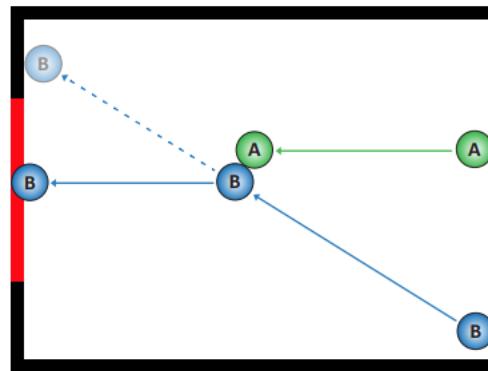
# Pearlian Causality

A success story

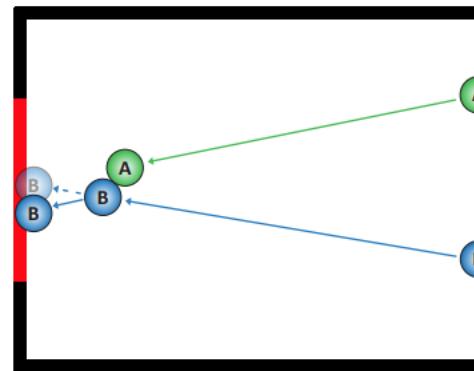
The formalization with most success in AI/ML so far.

Works in Cognitive Science also in support of the key ideas in the formalism i.e., **humans reason counterfactually**.

(a) A caused B to go into the gate.



(b) A didn't cause B to go into the gate.

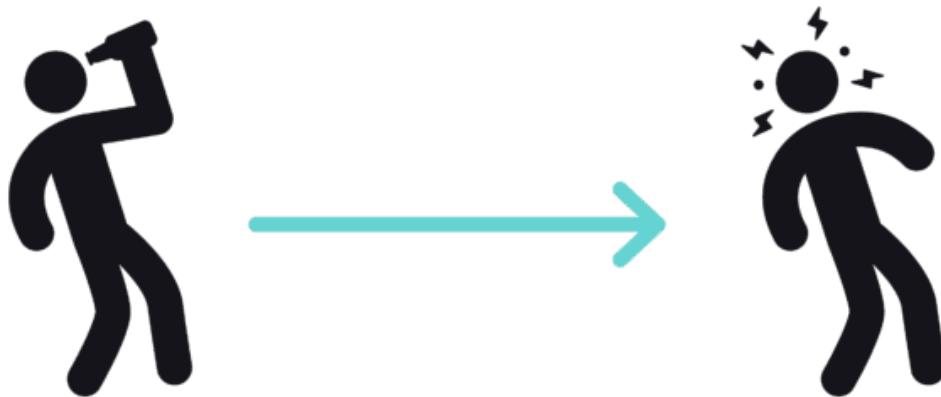


Gerstenberg. What would have happened? [...] PTRBAE 2022.

Causality allows us to talk about **modelling assumptions**

Causality allows us to consider not just the joint distribution but the **data generating process** which induces said distribution

# So, What is Causality?



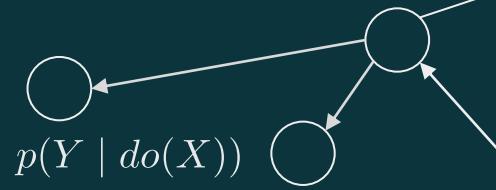
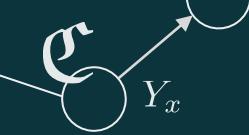
Bayesian network:

$$P(\text{ } \cdot \text{ } | \text{ } \cdot \text{ })$$

Causal AI model:

 causes 

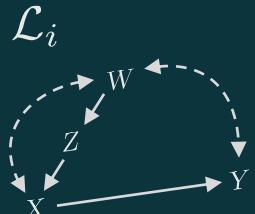
<https://www.causalens.com/white-paper/why-causal-ai-models-outrun-bayesian-networks/>



2

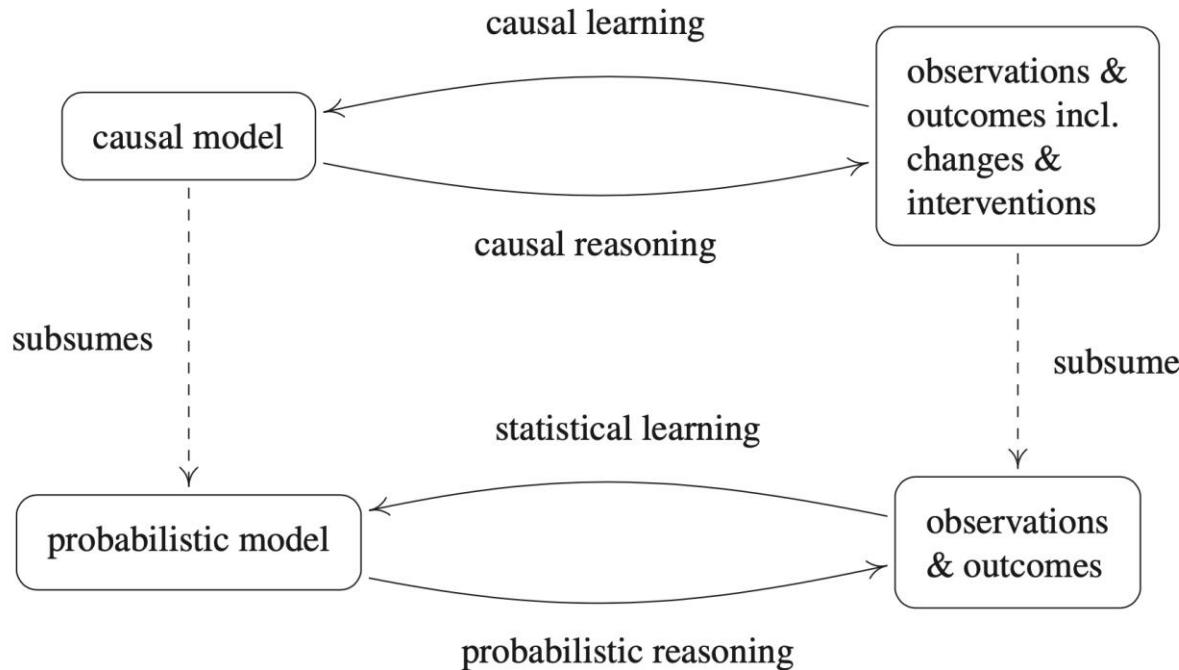
What?

Does Pearlian Causality look like?



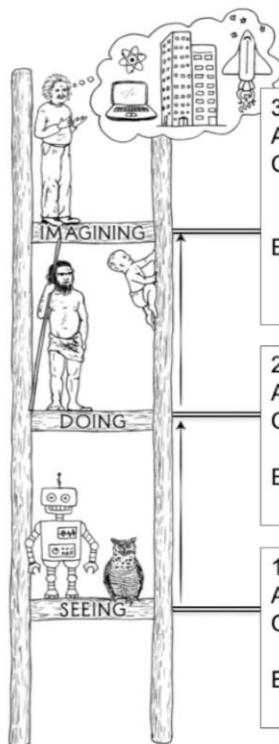
Machines Climbing Pearl's Ladder of Causation

# Causal versus Probabilistic Inference



Peters et al. 2017, Fig.1.1

# The Causal Hierarchy



## 3-LEVEL HIERARCHY

### 3. COUNTERFACTUALS

ACTIVITY: Imagining, Retrospection, Understanding

QUESTIONS: *What if I had done . . . ? Why?*

(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

EXAMPLES: Was it the aspirin that stopped my headache?  
Would Kennedy be alive if Oswald had not killed him? What if I had not smoked the last 2 years?

### 2. INTERVENTION

ACTIVITY: Doing, Intervening

QUESTIONS: *What if I do . . . ? How?*

(What would Y be if I do X?)

EXAMPLES: If I take aspirin, will my headache be cured?  
What if we ban cigarettes?

### 1. ASSOCIATION

ACTIVITY: Seeing, Observing

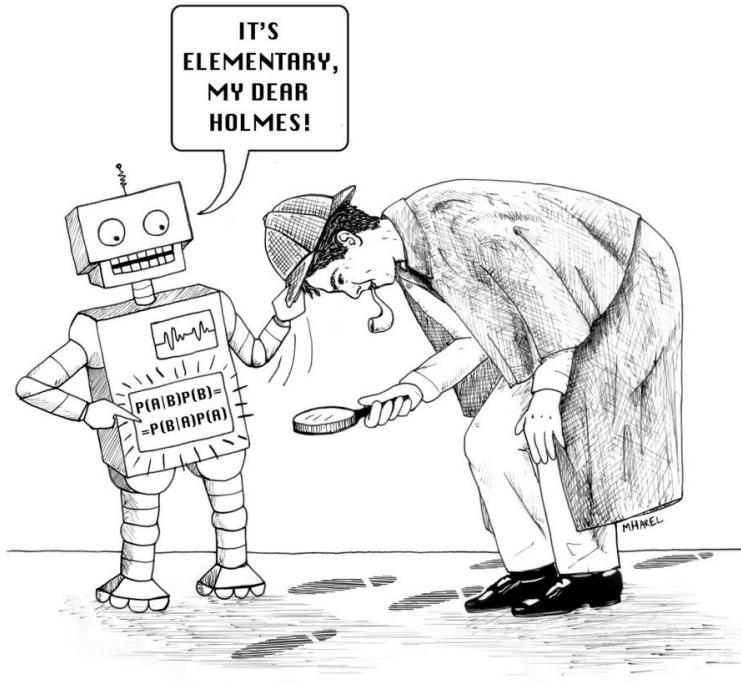
QUESTIONS: *What if I see . . . ?*

(How would seeing X change my belief in Y?)

EXAMPLES: What does a symptom tell me about a disease?  
What does a survey tell us about the election results?



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA



Sherlock Holmes meets his modern counterpart, a robot equipped with a Bayesian network. In different ways both are tackling the question of how to infer causes from observations. The formula on the computer screen is Bayes's rule. (Source:

Drawing by Maayan Harel.)



Machines Climbing Pearl's Ladder of Causation

# Pearl Causal Ladder

## Level I

### 1. ASSOCIATION

**ACTIVITY:** Seeing, Observing

**QUESTIONS:** *What if I see ...?*

(How are the variables related?)

How would seeing X change my belief in Y?)

**EXAMPLES:** What does a symptom tell me about a disease?  
What does a survey tell us about the  
election results?



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA



Machines Climbing Pearl's Ladder of Causation

# Pearl Causal Ladder

## Level II

### 2. INTERVENTION

**ACTIVITY:** Doing, Intervening

**QUESTIONS:** *What if I do ...? How?*

(What would Y be if I do X?  
How can I make Y happen?)

**EXAMPLES:** If I take aspirin, will my headache be cured?  
What if we ban cigarettes?

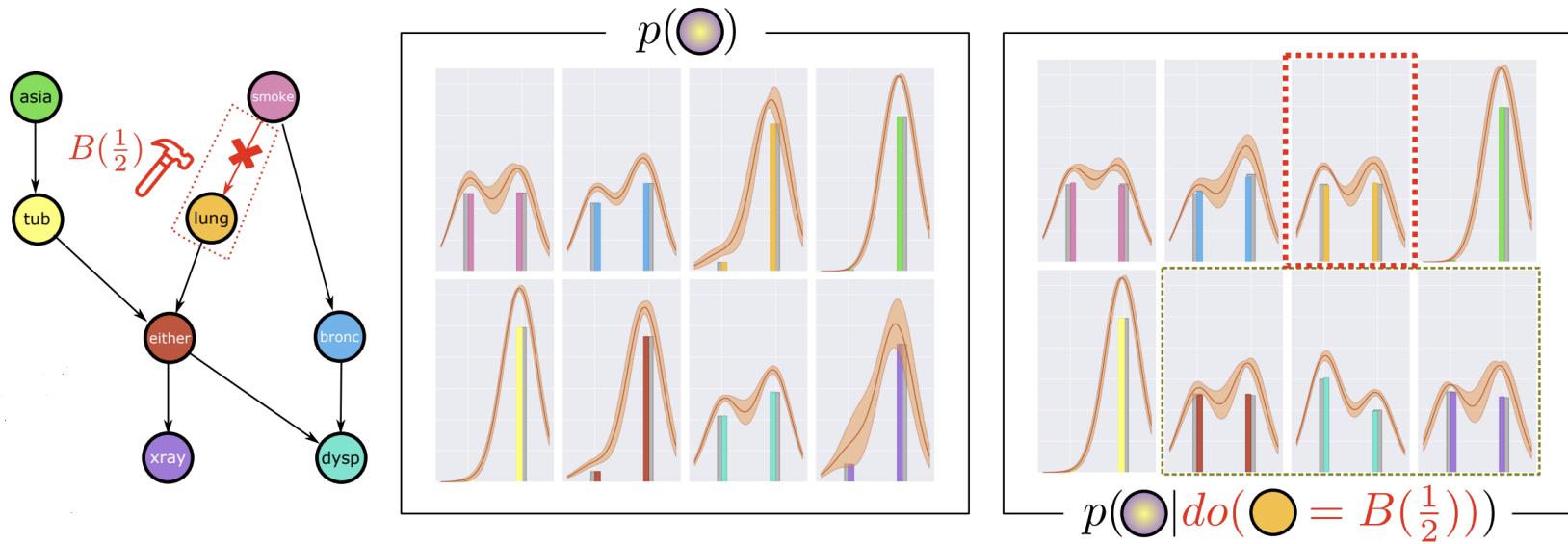


**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

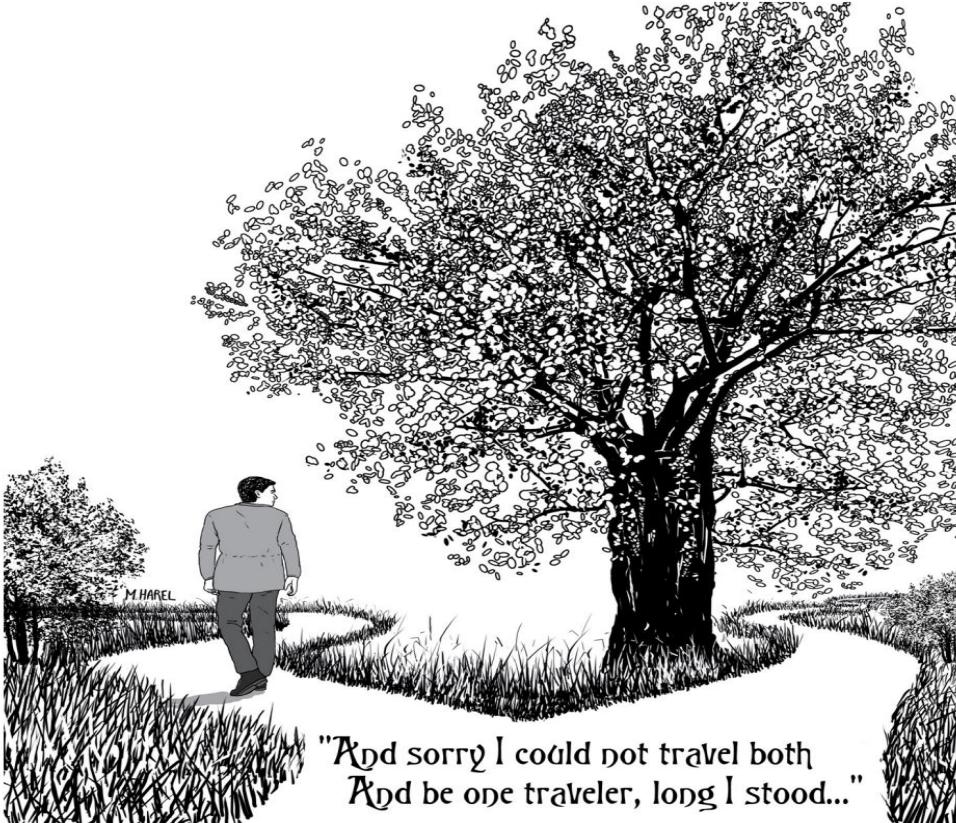
Machines Climbing Pearl's Ladder of Causation

# Pearl Causal Ladder

## Level II



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA



"And sorry I could not travel both  
And be one traveler, long I stood..."

Robert Frost's famous lines show a poet's acute insight into counterfactuals. We cannot travel both roads, and yet our brains are equipped to judge what would have happened if we had taken the other path. Armed with this judgment, Frost ends the poem pleased with his choice, realizing that it "made all the difference." (Source:

Drawing by Maayan Harel.)



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Pearl Causal Ladder

## Level III

### 3. COUNTERFACTUALS

**ACTIVITY:** Imagining, Retrospection, Understanding

**QUESTIONS:** *What if I had done ...? Why?*

(Was it X that caused Y? What if X had not occurred? What if I had acted differently?)

**EXAMPLES:** Was it the aspirin that stopped my headache?  
Would Kennedy be alive if Oswald had not killed him? What if I had not smoked for the last 2 years?



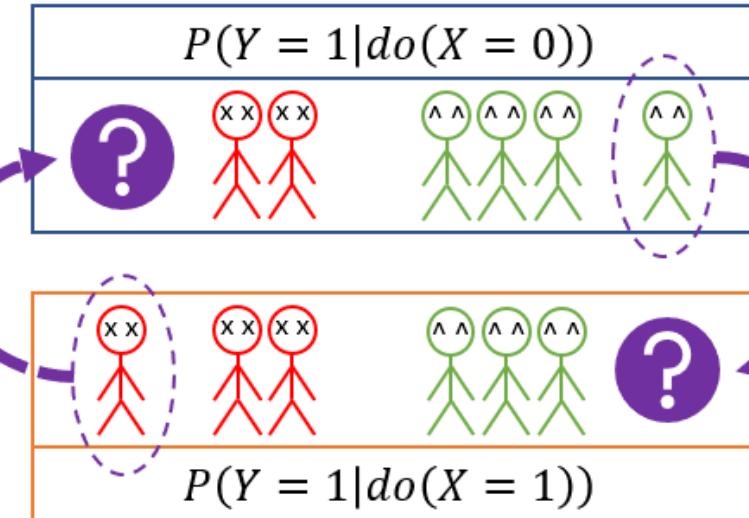
**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

# Pearl Causal Ladder

Level III

Would they have recovered *had we treated differently?*

Was this treatment *necessary for their recovery?*



Credit: <https://forns.lmu.build/classes/spring-2019/cmsi-498/lecture-5T.html>

# Pearl's Hierarchy

## Ladder of Causation

Layer	Task / Language	Typical Question	Examples
3. Imagining	Counterfactual $P(y_x   x', y')$	Structural Causal Model	What if I had acted differently? Was it the aspirin that stopped my headache?
2. Doing	Interventional $P(y   \text{do}(x), c)$	ML- Reinforcement (Causal Bayes Net)	What if I do X? What would Y be if I intervene on X? Will my headache be cured if I take aspirin?
1. Seeing	Associational $P(y   x)$	ML- (Un)Supervised (Decision trees, Deep nets, ...)	What if I see? How would seeing X change my belief in Y? What does a symptom tell us about the disease?

\* Book of Why & On Pearl's Hierarchy and the Foundations of Causal Inference, E. Bareinboim, J. Correa, D. Ibeling, T. Icard, in press. <https://causalai.net/r60.pdf>



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

# Structural Causal Model (SCM)

## Definition

A structural causal model  $\mathcal{M}$  (or data generating model) is a tuple  $\langle \mathbf{V}, \mathbf{U}, \mathcal{F}, P_{\mathbf{U}} \rangle$ , where

$\mathbf{V}$  are endogenous variables

$\mathbf{U}$  are exogenous variables

$\mathcal{F}$  are functions determining  $\mathbf{V}$  i.e.,  $v_i = f_i(\mathbf{pa}_i, \mathbf{u}_i)$

$P_{\mathbf{U}}$  is the probability distribution over  $\mathbf{U}$ .

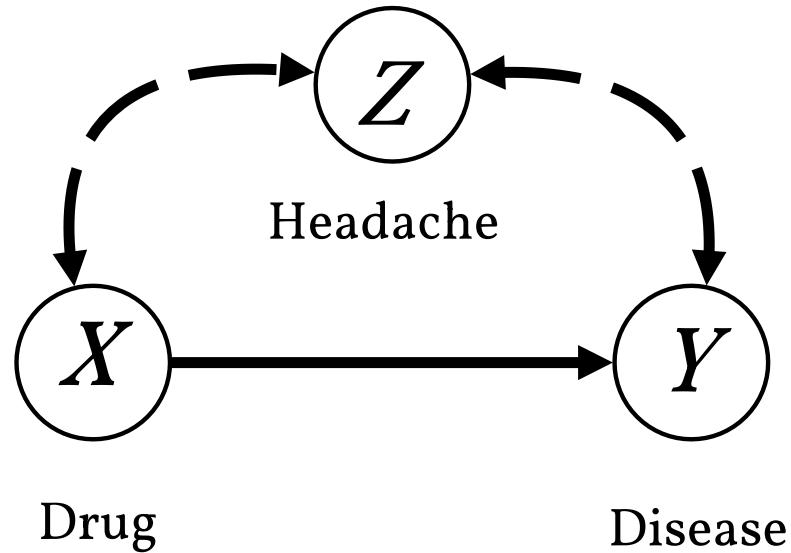
Assumption:  $\mathcal{M}$  is recursive i.e., there are no feedback (cyclic) mechanisms

# The Causal Graph

An induced property of the SCM

latent

$$\mathcal{F} = \begin{cases} X = f_X(U_X, U_{XZ}) \\ Y = f_Y(X, U_Y, U_{YZ}) \\ Z = f_Z(U_Z) \end{cases}$$



# The Causal Graph + Data

Structural Causal Model

$$A = U(0, 100)$$

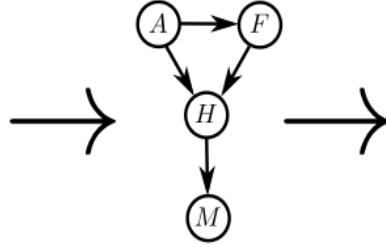
$$F = \frac{1}{2}A + \mathcal{N}(10, 10)$$

$$H = \frac{1}{100}(100 - A^2) + \frac{1}{2}F + \mathcal{N}(40, 30)$$

$$M = \frac{1}{2}H + \mathcal{N}(20, 10)$$

(A)ge  
(F)oood Habits  
(H)ealth  
(M)obility

*G*



*D*

A	62	...	18	24	...	21	...
F	48		60	20		32	
H	34		90	40		64	
M	39		75	37		66	

A	62	...	18	24	...	21	...
F	48		60	20		32	
H	+		-	+		-	
M	+			+			

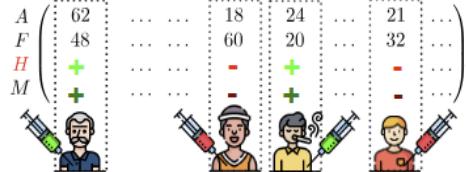
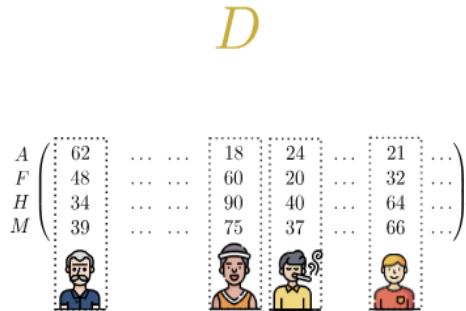
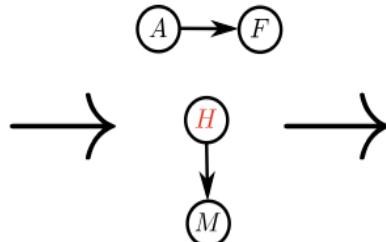


$$A = U(0, 100)$$

$$F = \frac{1}{2}A + \mathcal{N}(10, 10)$$

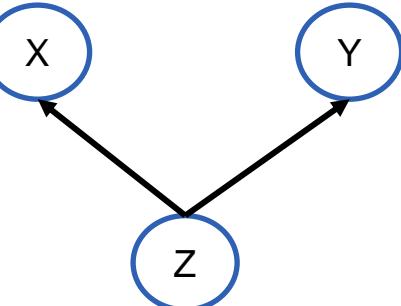
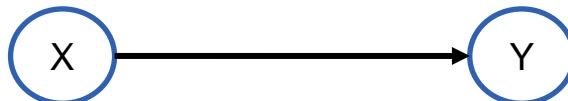
$$H = U(0, 100)$$

$$M = \frac{1}{2}H + \mathcal{N}(20, 10)$$

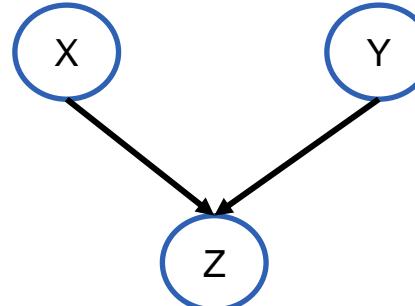


# The 3 Basic Graphs: “gifts from the gods”

Chain



Fork



Collider

# Graphical Representation of an SCM

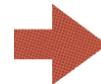
Structural Causal Model (SCM)

Observational

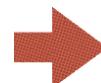
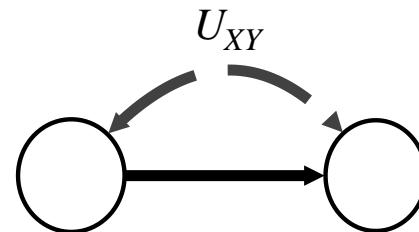
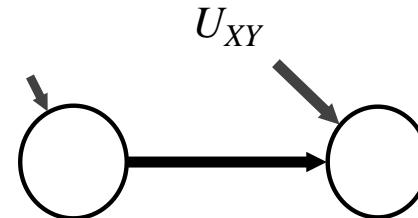
$$\mathcal{M} = \begin{cases} V = \{X, Y\} \\ U = \{U_{XY}, U_X, U_Y\} \\ \mathcal{F} = \{ X = f_X(U_X, U_{XY}), Y = f_Y(X, U_Y, U_{XY}) \} \\ P(U) \end{cases}$$

Interventional

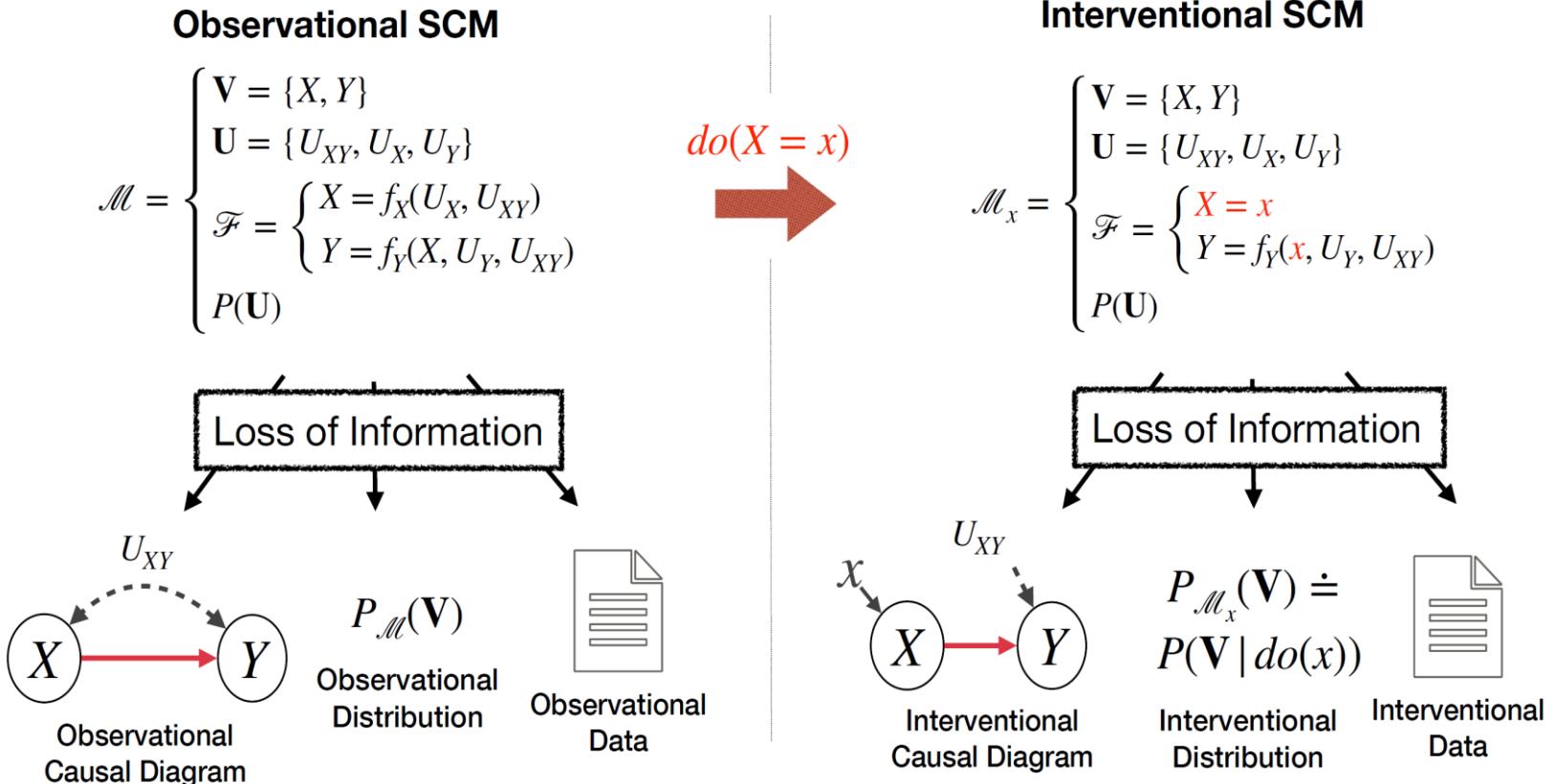
$$do(X = x)$$
$$\mathcal{M} = \begin{cases} V = \{X, Y\} \\ U = \{U_{XY}, U_X, U_Y\} \\ \mathcal{F} = \{ X = x, Y = f_Y(x, U_Y, U_{XY}) \} \\ P(U) \end{cases}$$

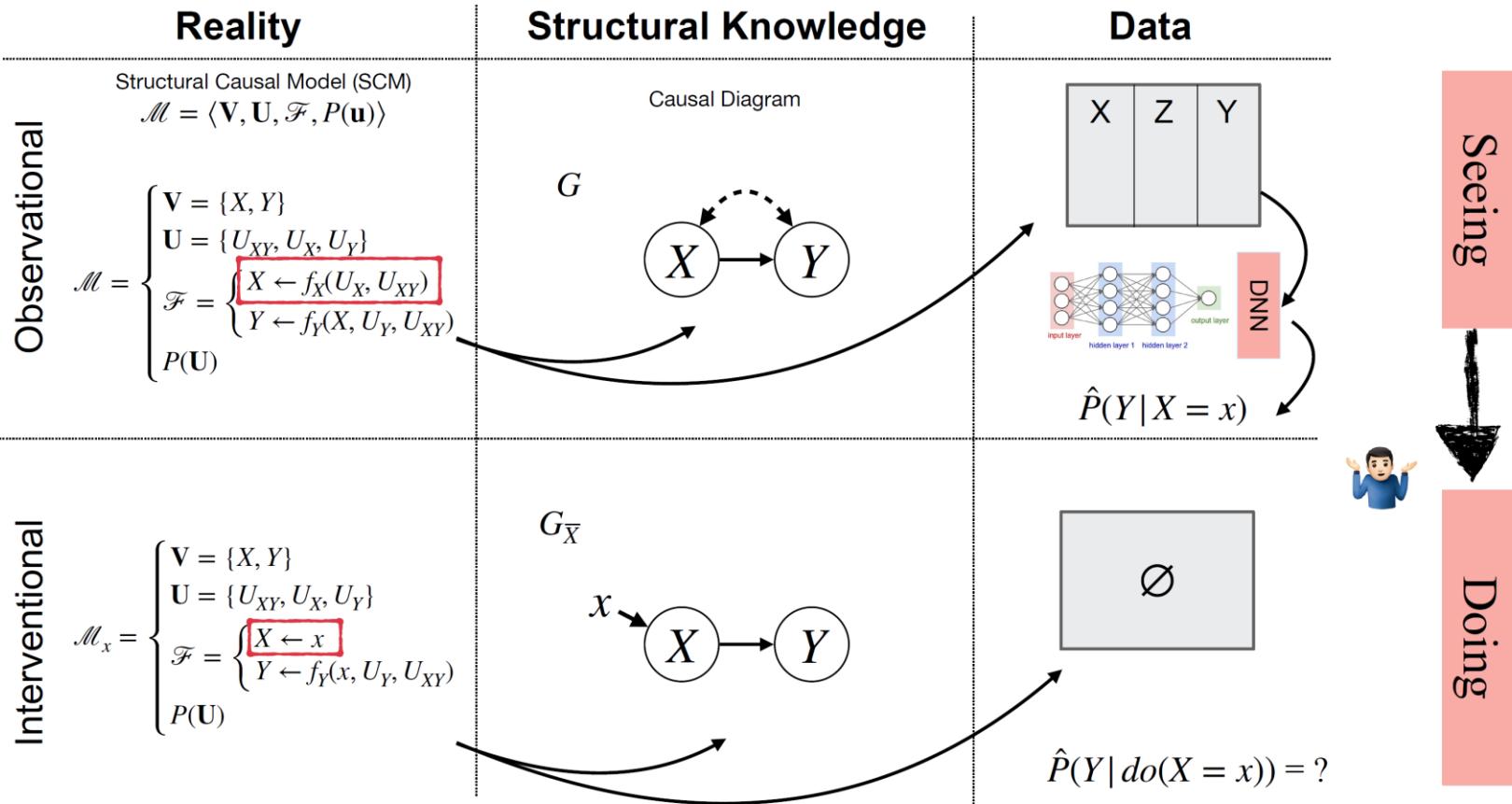


Graphical Causal Model (Causal Diagram)


$$do(X = x)$$


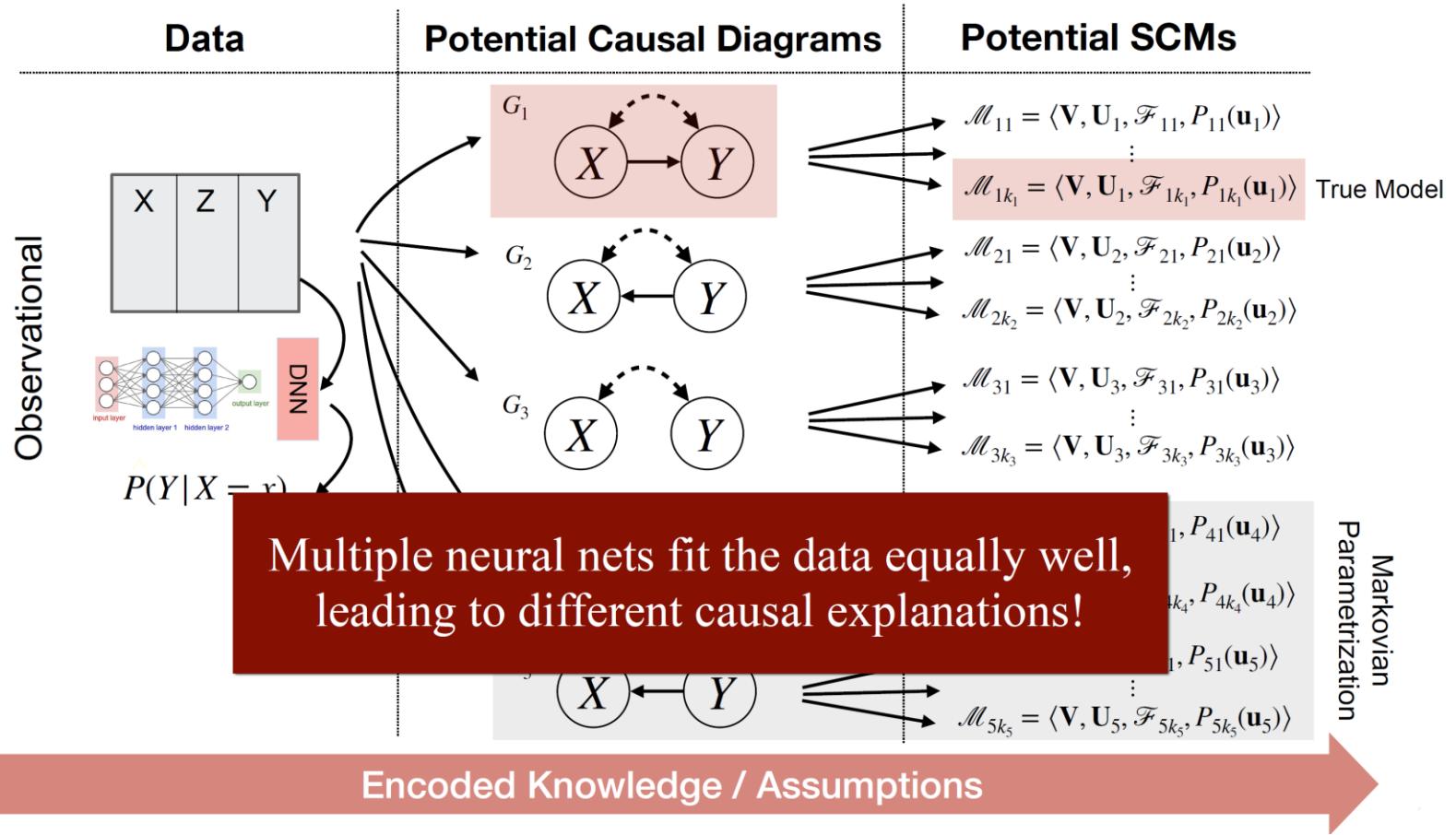
# What is induced by the SCM?

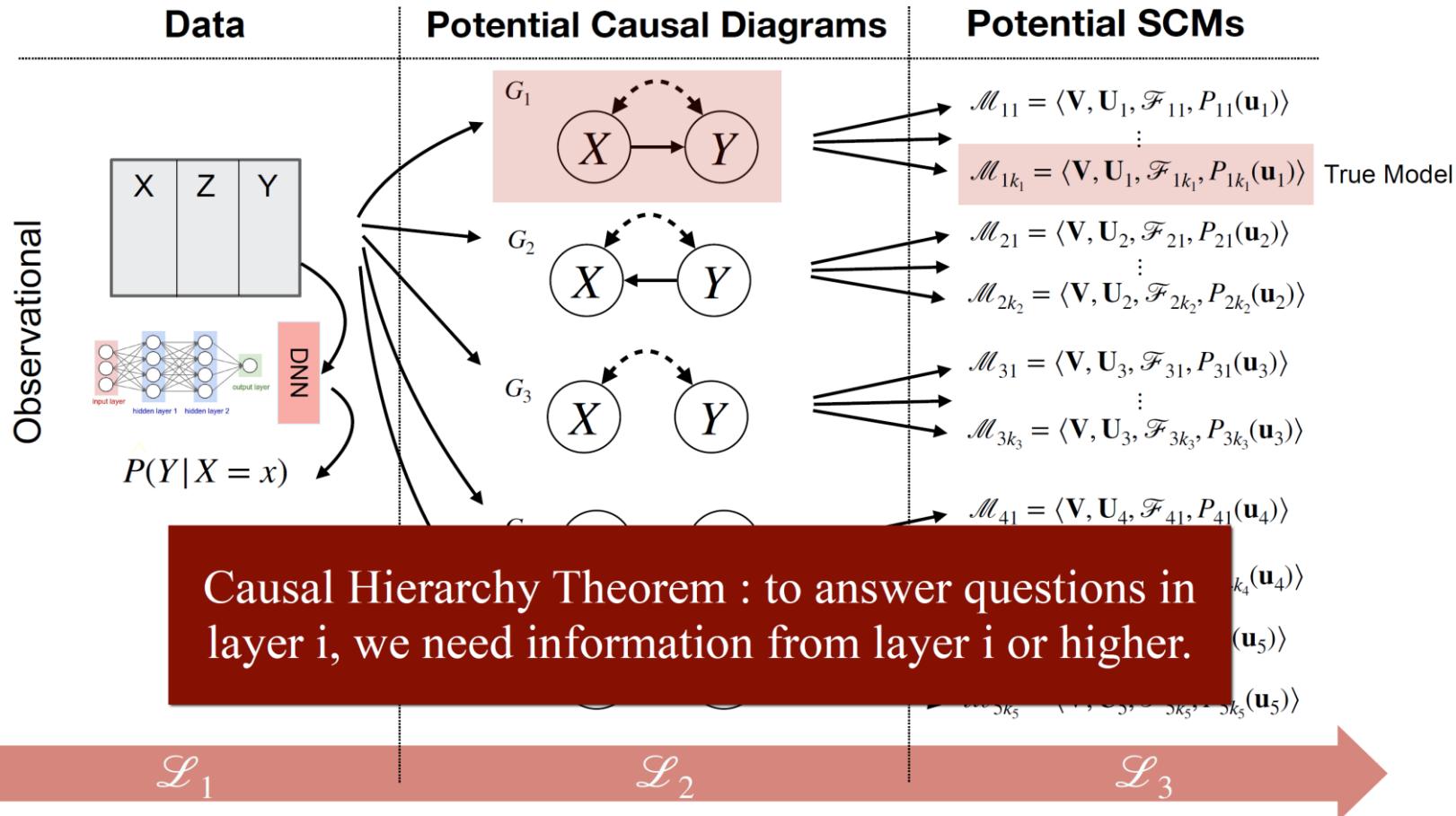




33



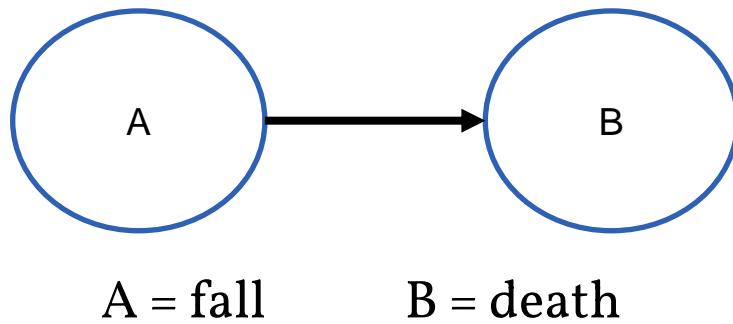




ESSAI & ACAI 2023

LJUBLJANA, SLOVENIA

# Causality vs Conditional Probability



- But death is possible even without fall
- All the people who fall get wings and it ensures survival

$$P(B = 0 | A = 1) = 1 > P(B = 0 | A = 0)$$



# Counterfactuals

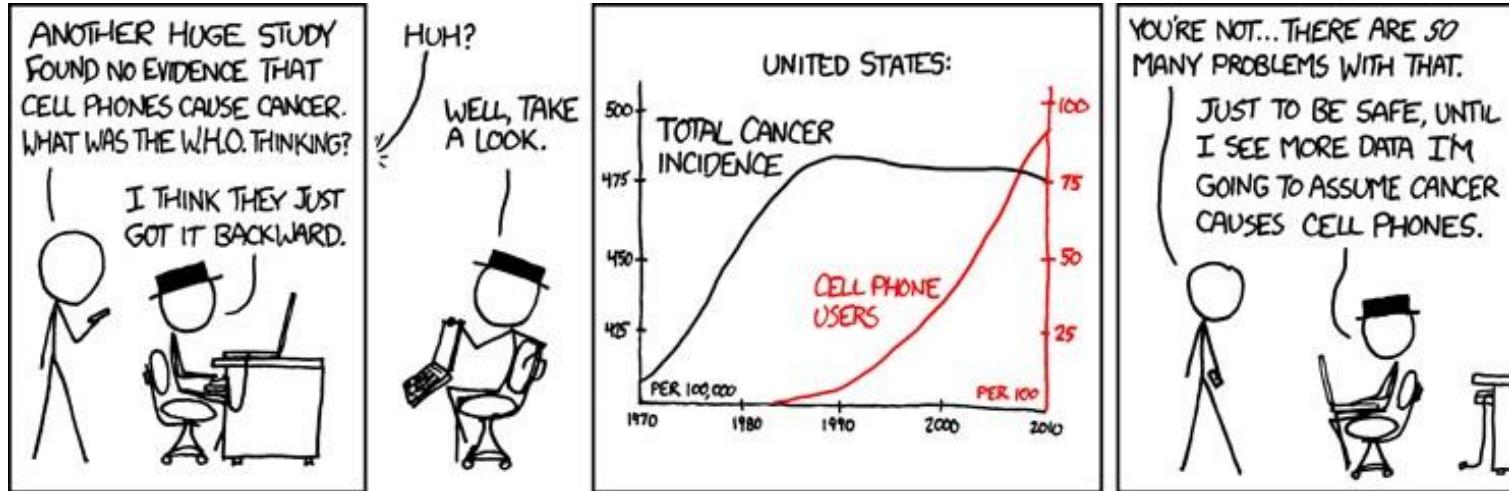
A 3-step procedure

Abduction : Update belief in exogenous variables given evidence

Action : Change equations accordingly,  $do(X = x)$

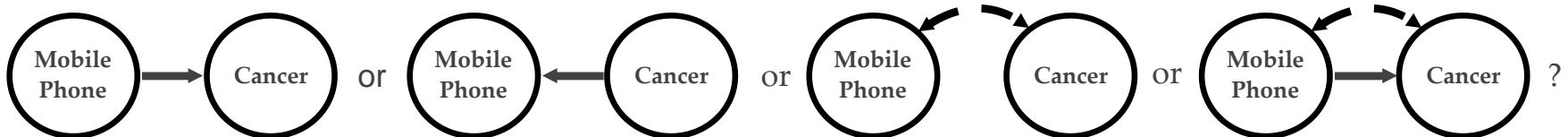
Prediction : Look at variable of interest  $P(Y = y)$

# Association vs Causation

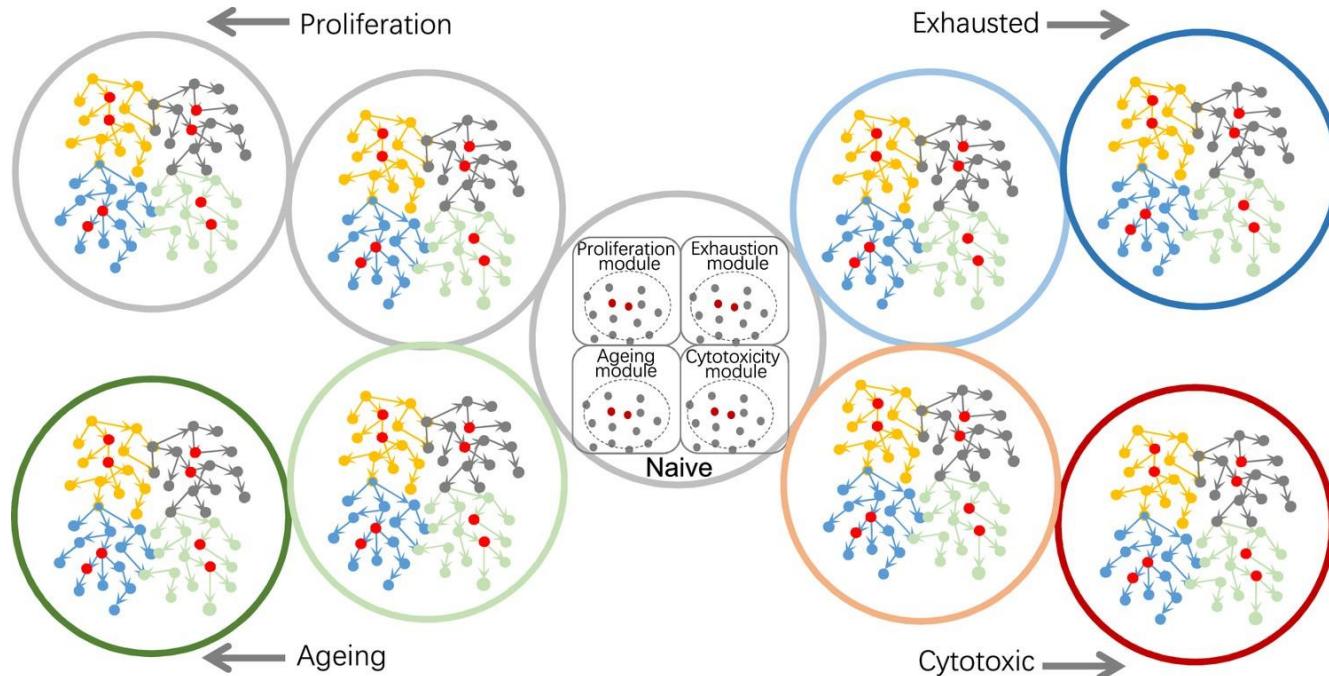


<https://xkcd.com/925/> - Creative Commons Attribution-NonCommercial 2.5 License.

Will we be able to decide the true relationship just by “seeing” more data?

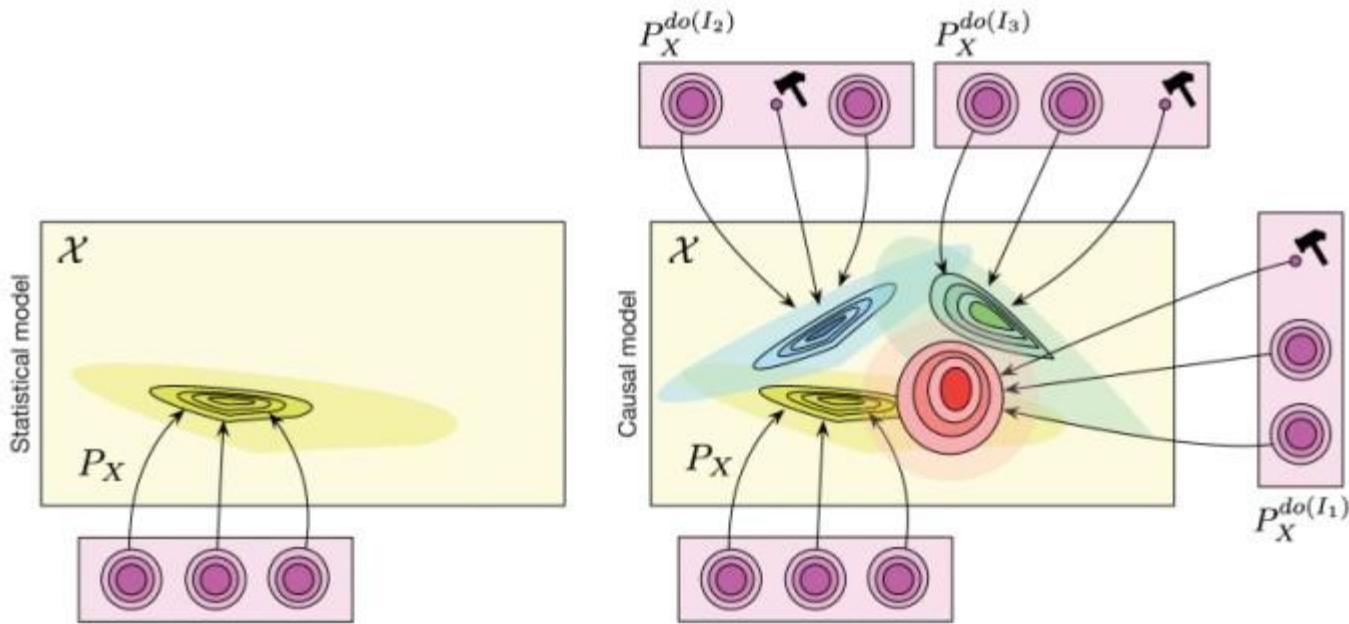


# Problems to be solved

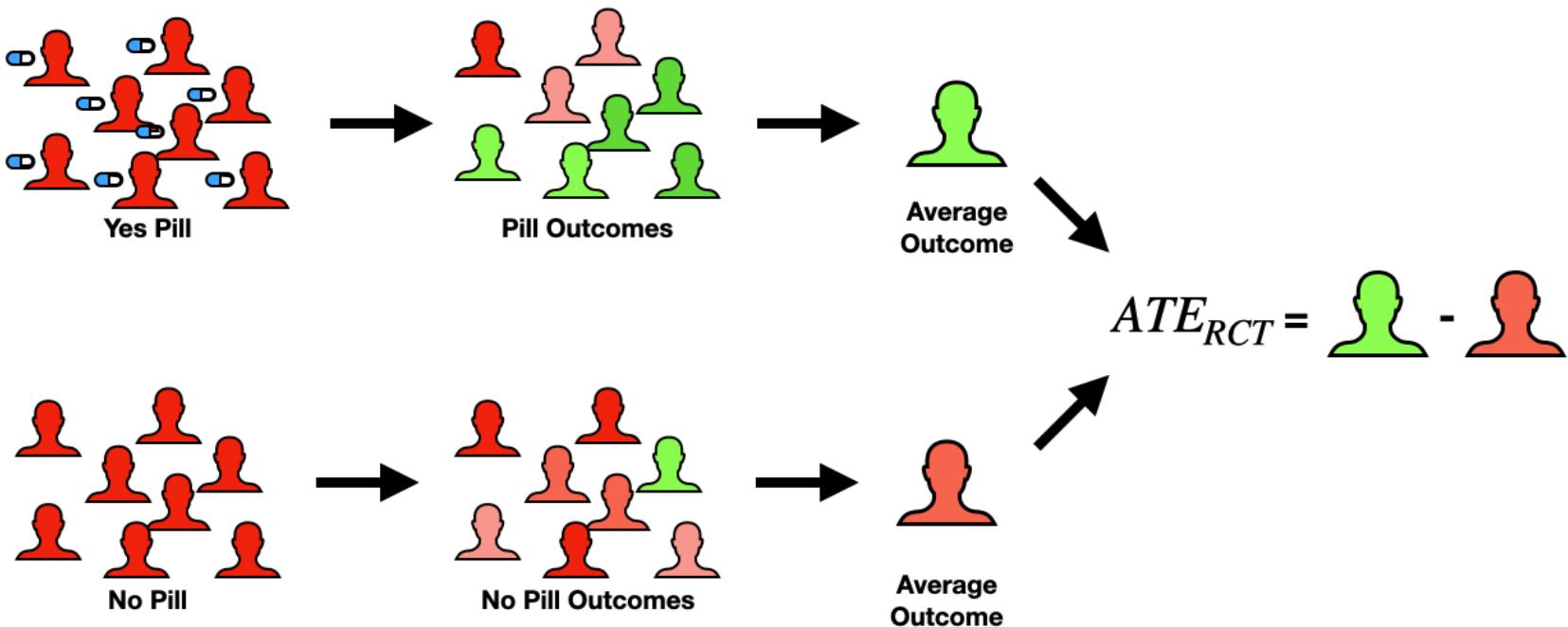


Wen et al., Applying causal discovery to single-cell analyses using CausalCell, eLife 2023

# Problems to be solved



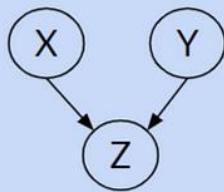
# Problems to be solved



<https://towardsdatascience.com/causal-effects-f30f962ffff3>

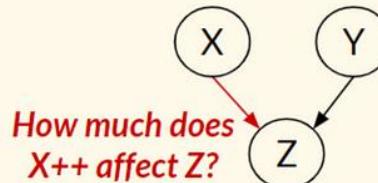
# Problems to be solved

Assume the variables and causal graph.



## Causal Effect Estimation

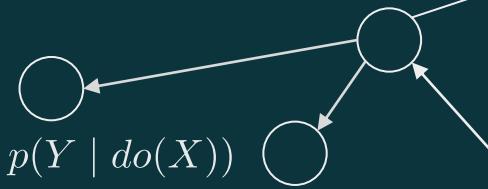
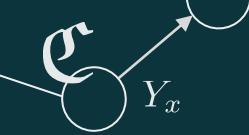
3. By how much?



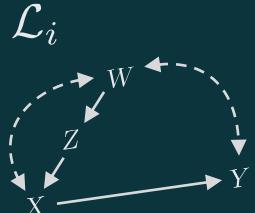
4. Full understanding

$$Z = f(X, Y, \sigma)$$

Human civilization  
advances

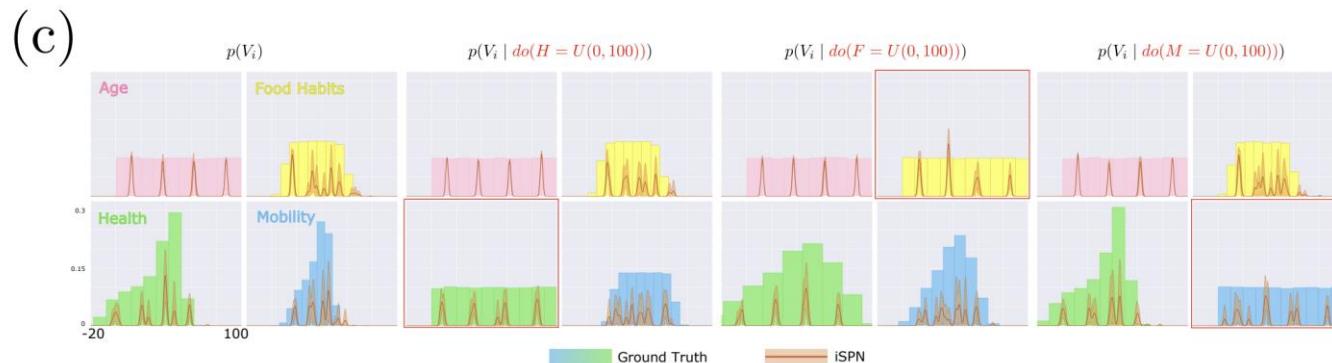
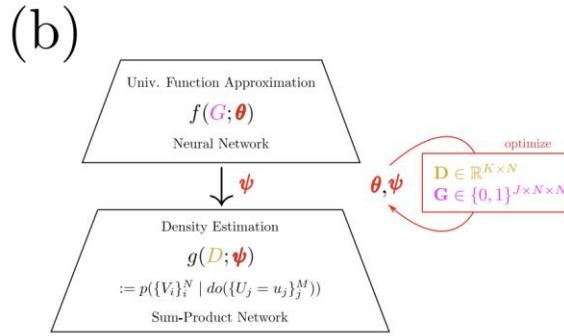
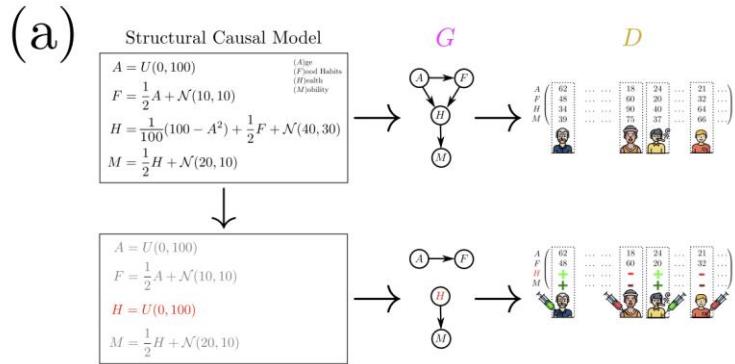


# 3 | Machine Learning for Causality: A Flavour

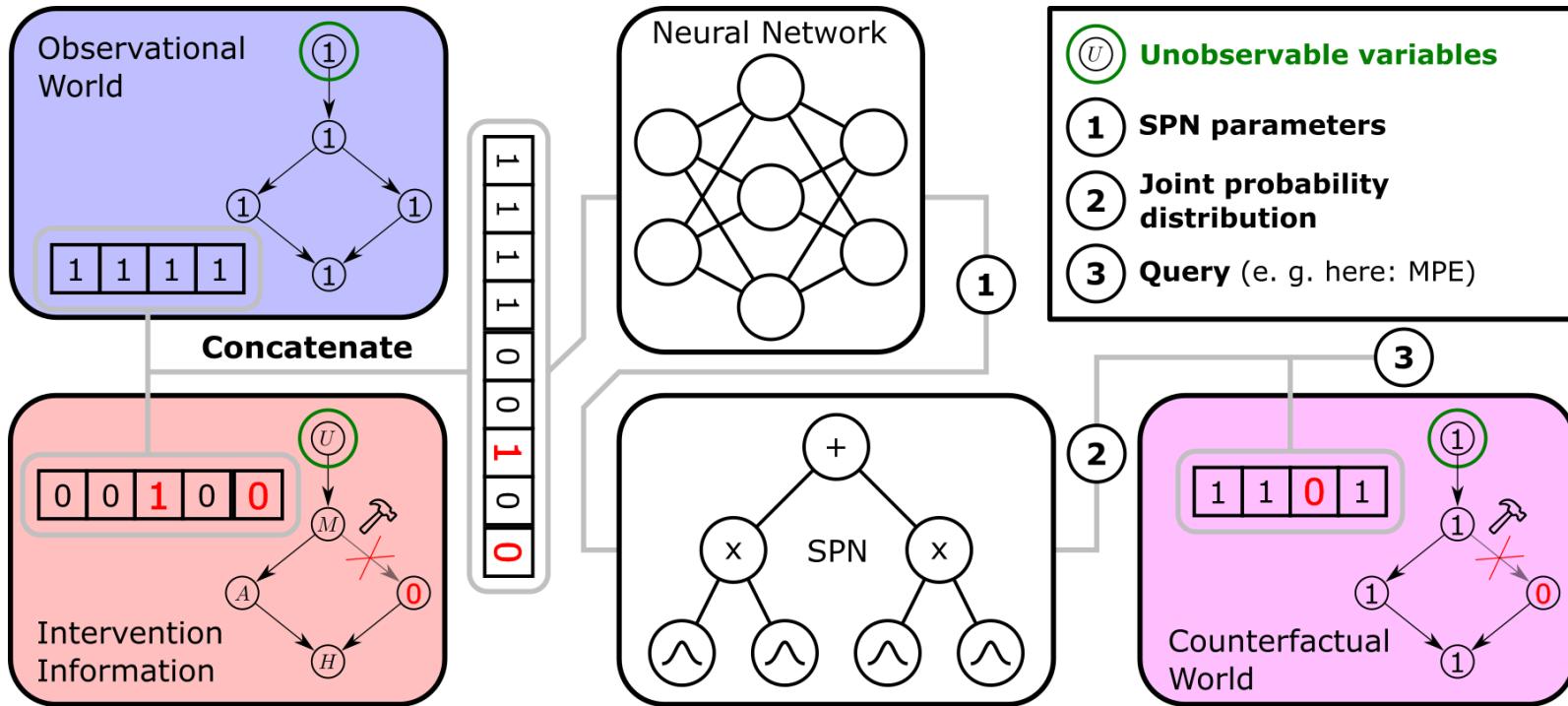


Machines Climbing Pearl's Ladder of Causation

# Probabilistic Circuits + Causality

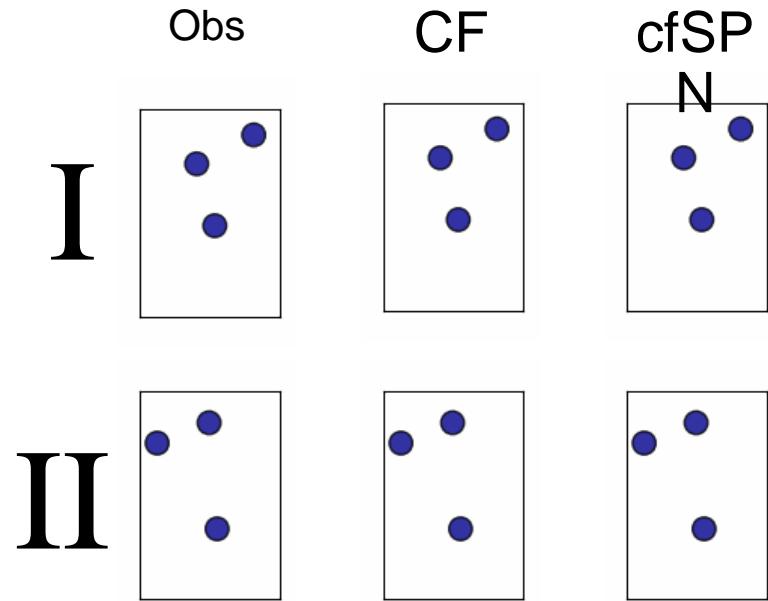


# Counterfactual Sum-Product Networks



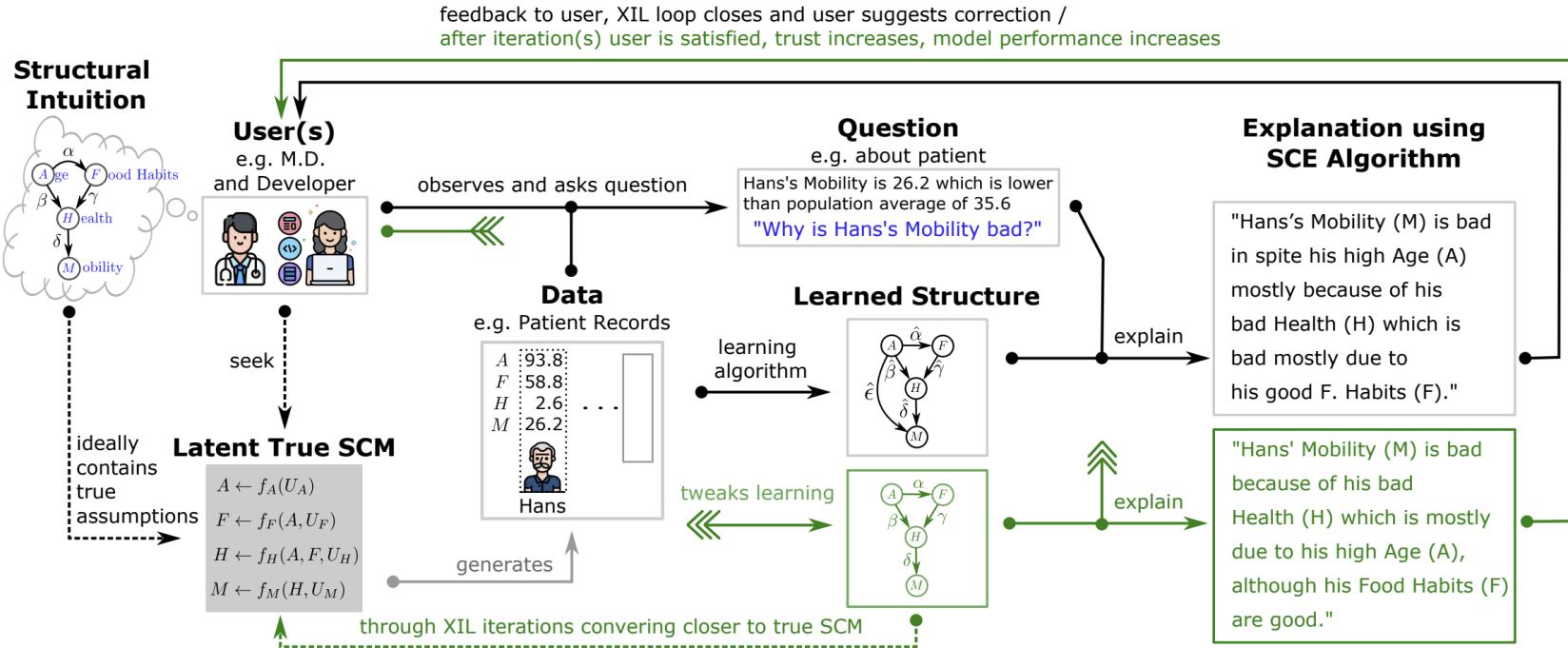
# Experiment: Particle Collision

- A difficult problem: Particle simulation with gravity and collisions
- Goal: cfSPN prediction should match true counterfactual simulation (CF)
- I: Move the bottom particle to the right after some timesteps
- II: Change the velocity of the top particle to slightly upwards at the start



Based on the code repository for particle simulation:  
[https://github.com/ineporozhnii/particles\\_in\\_a\\_box](https://github.com/ineporozhnii/particles_in_a_box)

# Logic + Causality = Explanations



# Free Code Libraries

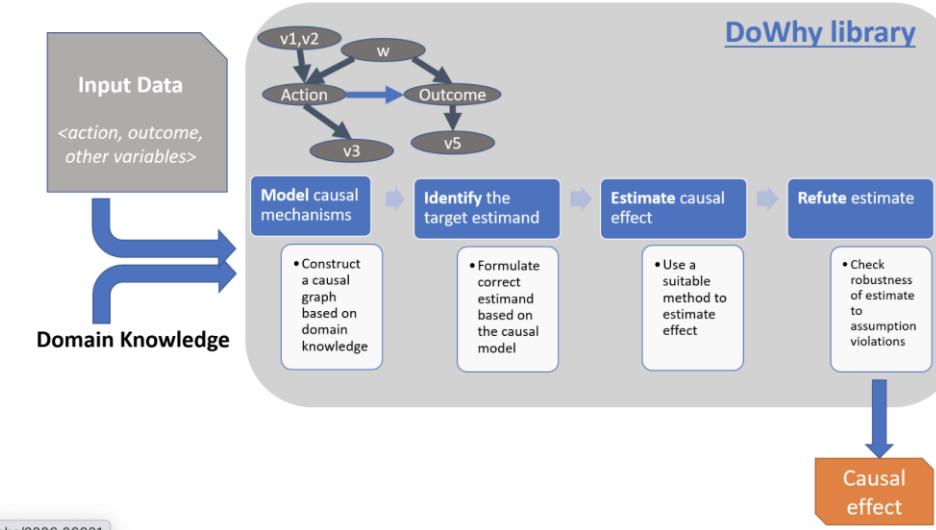
Do it for you

## DoWhy | An end-to-end library for causal inference

Introducing DoWhy and the 4 steps of causal inference | Microsoft Research Blog | Video Tutorial | Arxiv Paper  
| Arxiv Paper (GCM-extension) | Slides

Read the [docs](#) | Try it online!  [launch binder](#)

Case Studies using DoWhy: [Hotel booking cancellations](#) | [Effect of customer loyalty programs](#) | [Optimizing article headlines](#) | [Effect of home visits on infant health \(IHDP\)](#) | [Causes of customer churn/attrition](#)

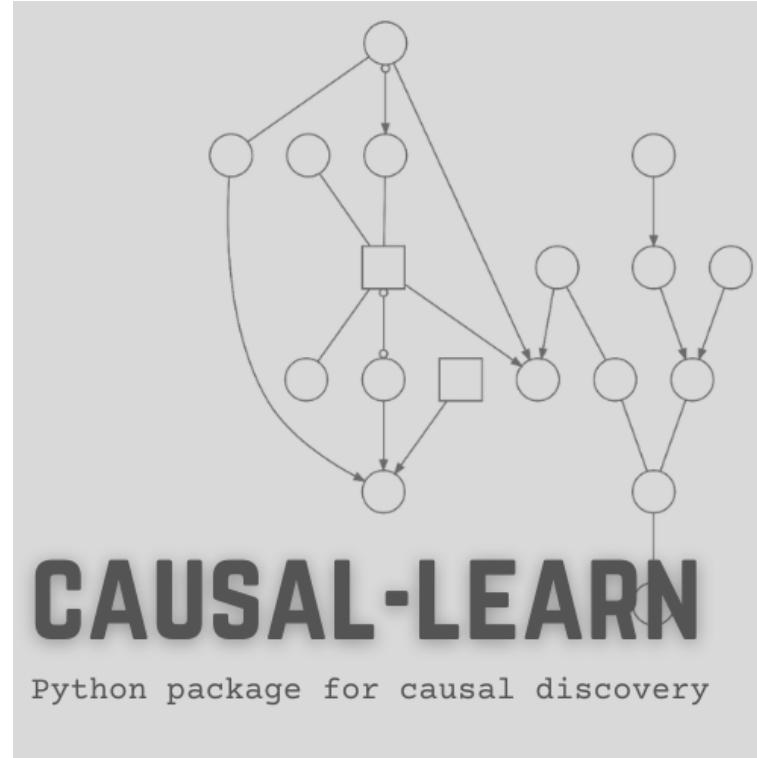


[mabs/2206.06821](#)

DoWhy, <https://github.com/py-why/dowhy>

# Free Code Libraries

Do it for you



DoWhy, <https://github.com/py-why/causal-learn>



# Pointers to Causal Inference References

- ✓  Judea Pearl, “**Causality**”, Cambridge University Press, 2009.
- ✓  Peters et al., “**Elements of Causal Inference**”, MIT Press, 2017.

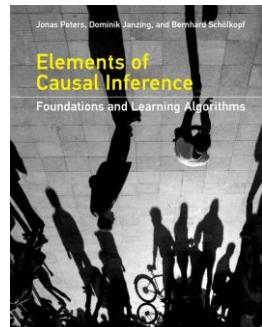
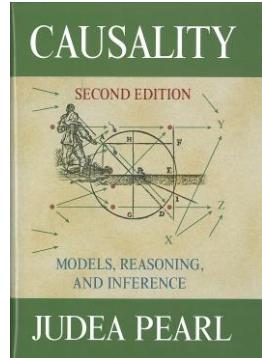
- Elias Bareinboim Lecture “**Causal Data Science**”, 2019.

<https://www.youtube.com/watch?v=dUsokjG4DHC>

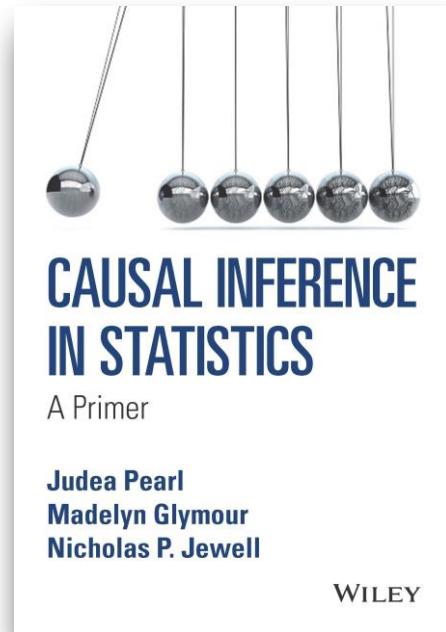
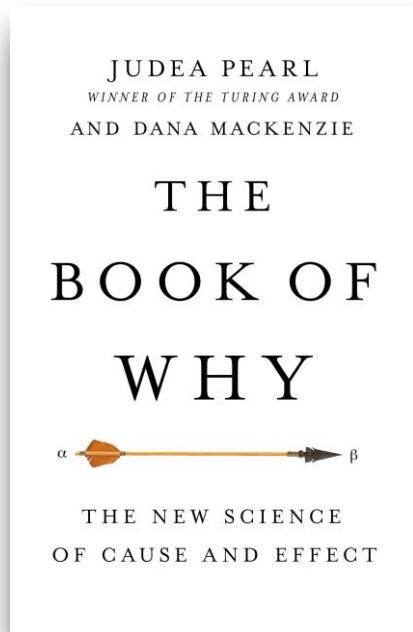
- ✓  Brady Neal’s Free Online Course “**Introduction to Causal Inference**”, 2020.
- <https://www.bradyneal.com/causal-inference-course>

- Jonas Peters Lecture Series “**Causality**”, 2017.

<https://www.youtube.com/watch?v=zvrcyqcN9Wo>



# Causality Theory by Judea Pearl



ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Every Week with Paper Authors → Discuss LIVE



**530 members  
on Slack**

Join the community  
via  
[discuss.causality.link](https://discuss.causality.link)



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

Past Sessions: [Password: Causality, Direct Access Link]

- Session 01.03.2023 | **Deep Counterfactual Estimation with Categorical Background Variables** | Discussant: Edward De Brouwer
- Session 22.02.2023 | **Information-Theoretic Causal Discovery and Intervention Detection over Multiple Environments** | Discussant: Osman Ali Mian
- Session 08.02.2023 | **CLEAR: Generative Counterfactual Explanations on Graphs** | Discussants: Jing Ma, Ruocheng Guo
- Session 01.02.2023 | **Causal Transformer for Estimating Counterfactual Outcomes** | Discussant: Valentyn Melnychuk
- Session 25.01.2023 | **Abstracting Causal Models** | Discussant: Sander Beckers
- Session 18.01.2023 | **Desiderata for Representation Learning: A Causal Perspective** | Discussant: Yixin Wang
- Session 11.01.2023 | **Causal Feature Selection via Orthogonal Search** | Discussant: Ashkan Soleymani
- Session 14.11.2022 | **Rewind 2022** | Final session of 2022 to simply rewind on what we experienced throughout the year
- Session 07.12.2022 | **Causal Inference Through the Structural Causal Marginal Problem** | Discussant: Luigi Gresele
- Session 30.11.2022 | **Selecting Data Augmentation for Simulating Interventions** | Discussant: Maximilian Ilse
- Session 23.11.2022 | **On Disentangled Representations Learned from Correlated Data** | Discussant: Frederik Träuble
- Session 16.11.2022 | **Causal Curiosity: RL Agents Discovering Self-supervised Experiments for Causal Repr. Learning** | Discussant: Sumedh Sontakke
- Session 09.11.2022 | **Causal Machine Learning: A Survey and Open Problems** | Discussants: Jean Kaddour, Aengus Lynch
- Session 02.11.2022 | **A Critical Look at the Consistency of Causal Estimation with Deep Latent Variable Models** | Discussant: Severi Rissanen
- Session 26.10.2022 | **Nonlinear Invariant Risk Minimization: A Causal Approach** | Discussant: Chaochao Lu
- Session 19.10.2022 | **CausalVAE: Disentangled Representation Learning via Neural Structural Causal Models** | Discussant: Mengyue Yang
- Session 12.10.2022 | **Weakly Supervised Causal Representation Learning** | Discussant: Johann Brehmer
- Session 05.10.2022 | **Towards Causal Representation Learning** | Discussant: Anirudh Goyal
- Session 21.09.2022 | **Selection Collider Bias in Large Language Models** | Discussant: Emily McMilin
- Session 14.09.2022 | **The Causal-Neural Connection: Expressiveness, Learnability, and Inference** | Discussants: Kai-Zhan Lee, Kevin Xia
- Session 07.09.2022 | **Self-Supervised Learning with Data Augmentations Provably Isolates Content from Style** | Discussant: Julius von Kügelgen

35+ Sessions  
Completed  
and  
All Recorded



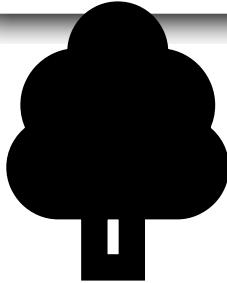
ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

Want to learn more  
*about the genealogy*  
of causality research?

# Genealogy of Causality

Access via [genealogy.causality.link](http://genealogy.causality.link)



Name	Institution	Supervisor	Location	Previous Positions
<b>UCLA</b>				
Judea Pearl	UCLA	?	US	Rutgers, Technion, New
Wesley Salmon	UCLA	Hans Reichenbach	US	?
Hans Reichenbach	UCLA	Paul Hensel, Max Noeth	US	Berlin, Istanbul, Erlange
<b>John Hopkins</b>				
Ilya Shpitser	John Hopkins		US	UCLA, Judea Pearl
<b>Oregon State University</b>				
Karthika Mohan	Oregon State University	Judea Pearl	US	
<b>CMU</b>				
Kun Zhang	CMU		Pittsburgh, US	MPI Tübingen
Clark Glymour	CMU	Wesley Salmon	Pittsburgh, US	
Peter Spirtes	CMU		Pittsburgh, US	
<b>ETH Zürich</b>				
Peter Bühlmann	ETH		Zürich	?
Marloes Maathuis	ETH		Zürich	?
Nicolai Meinshausen	ETH			
<b>LMU Munich</b>				
Stephan Hartmann	LMU		Munich, Germany	
<b>MPI Tübingen</b>				
Bernhard Schölkopf	MPI Tübingen	Vladimir Vapnik	Tübingen, Germ	TU Berlin
Ulrike von Luxburg	MPI Tübingen		Tübingen, Germany	
Michel Besserve				

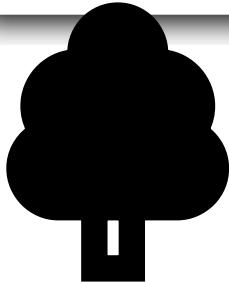


ESSAI & ACAI 2023  
LJUBLJANA, SLOVENIA

Machines Climbing Pearl's Ladder of Causation

# Genealogy of Causality

Access via [genealogy.causality.link](http://genealogy.causality.link)



**We Want You! To Extend This!**

The slide features a green header bar with the text "Name", "Institution", "Supervisor", "Location", and "Previous Positions". Below the header is a table with several rows of data, each consisting of a green header cell and a white data cell. The data cells contain various names and titles, such as "UCLA", "John Doe", "Jane Smith", "Software Engineer", "New York City", "Manager of Data Science", etc.



**ESSAI & ACAI 2023**  
LJUBLJANA, SLOVENIA

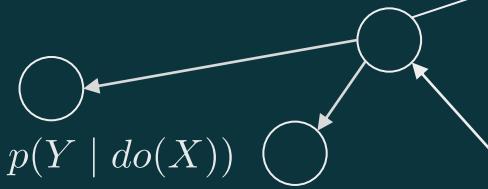
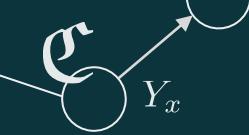
Machines Climbing Pearl's Ladder of Causation

Join the community  
via [discuss.causality.link](https://discuss.causality.link)



Access the genealogy  
via [genealogy.causality.link](https://genealogy.causality.link)

Genealogy of Causality				
Name	UCLA	US	Position	Previous Positions
Judea Pearl				Rutgers, Technion, New
Wesley Salmon	UCLA	Hans Reichenbach	US	?
Hans Reichenbach	UCLA	Paul Hensel, Max Noeth	US	Berlin, Istanbul, Erlange
John Hopkins				
Ilya Shpitser	John Hopkins		US	UCLA, Judea Pearl
Oregon State University				
Karthika Mohan	Oregon State University	Judea Pearl	US	
CMU				
Kun Zhang	CMU		Pittsburgh, US	MPI Tübingen
Clark Glymour	CMU	Wesley Salmon	Pittsburgh, US	
Peter Spirtes	CMU		Pittsburgh, US	
ETH Zürich				
Peter Bühlmann	ETH		Zürich	?
Marloes Maathuis	ETH		Zürich	?
Nicolai Meinshausen	ETH			
LMU Munich				
Stephan Hartmann	LMU		Munich, Germany	
MPI Tübingen				
Bernhard Schölkopf	MPI Tübingen	Vladimir Vapnik	Tübingen, Germ	TU Berlin
Ulrike von Luxburg	MPI Tübingen		Tübingen,	Germany
Michel Besserve				



After having seen all this, we realize..

*“As X-rays are to the surgeon, graphs are for causation.”*

-Judea Pearl in Causality (2009)

