UNIVERSITÄT ZU LÜBECK

Automated Planning and Acting

Intervention and Causal Planning

Institute of Information Systems

Dr. Mattis Hartwig

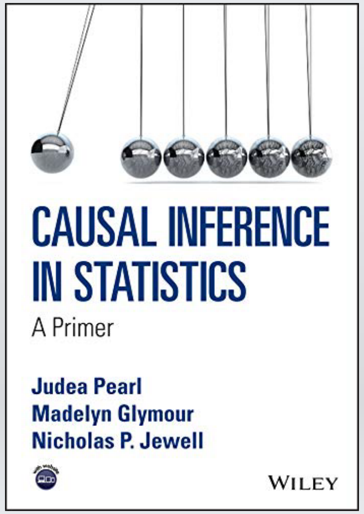
Content


Kat Maddox
@ctrlshifti
Is it a coincidence? I don't think so. Stop testing your software
Bugs found
Time spent testing software
14:06 · 20 Sep 20 · Twitter Web App
1,717 Retweets 141 Quote Tweets
Posted in r/ProgrammerHumor
reddit

Acknowledgements

Inspired by Slides from Prof. Dr. Ralf Möller and Dr. Özgür Özçep

Based on “Causal Inference in Statistics: A Primer”.

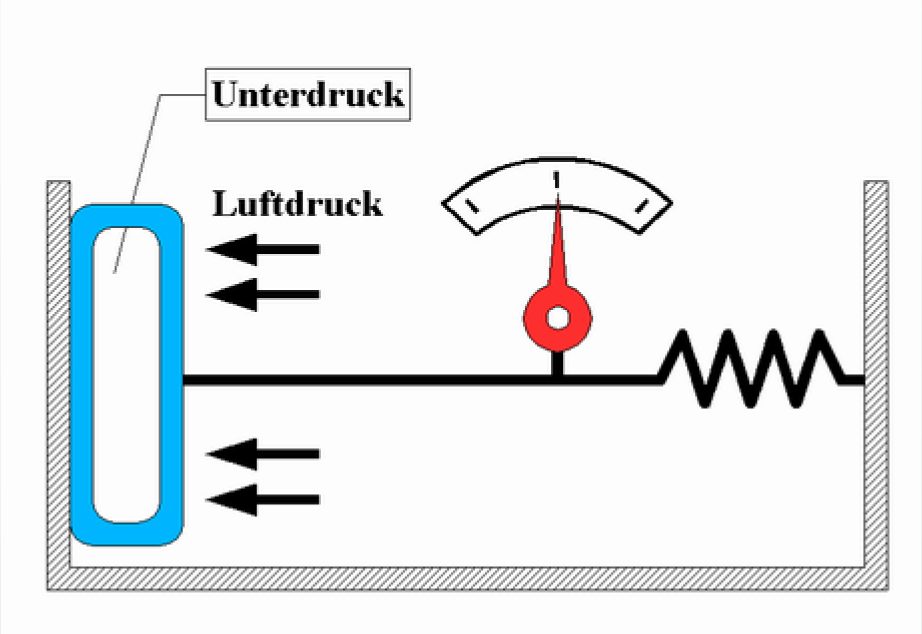


Typically, we are not only interested in observing data but we also want to intervene

- Important aim for given data: Where to intervene in order to achieve desired effects.
- Example interventions
  - Should we stop smoking?
  - What are the best methods to decrease wild-fires?

- Difference in measuring the atmospheric pressure with a barometer vs. forcing the needle to a specific measurement

Barometer



Randomized Controlled Experiment

- Randomized controlled experiment gold standard
  - Aim: Answer question whether a change in RV X has indeed an effect on some target RV Y
  - If outcome of experiment is yes, X is a RV to intervene upon
  - Test condition: all variables different from X are static (fixed) or vary fully randomly.
- Problem: Cannot always set up such an experiment
  - Example: cannot control weather in order to test variables influencing wildfire
- Instead: use observational data & causal model

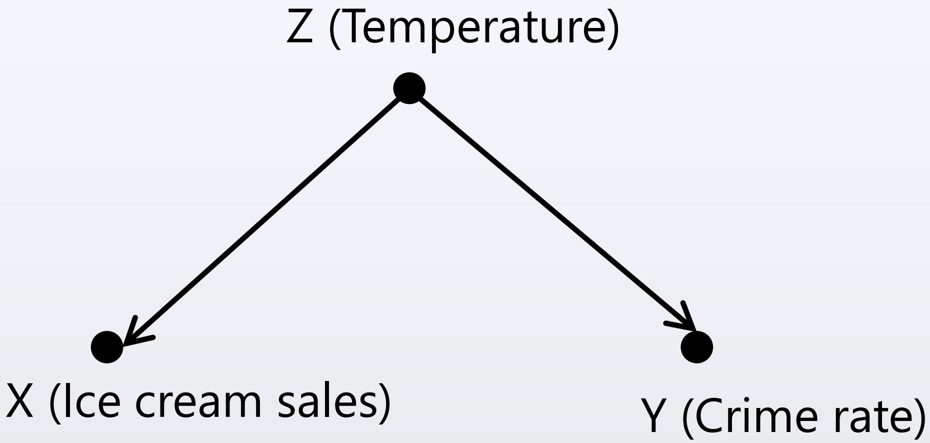
Intervention vs. Conditioning

- Intervention denoted by do(Y = y)
  - P(Z = z | do(Y = y)) =
  - probability of event Z = z on intervening upon Y by setting Y = y
    Intervention changes the data generation mechanism

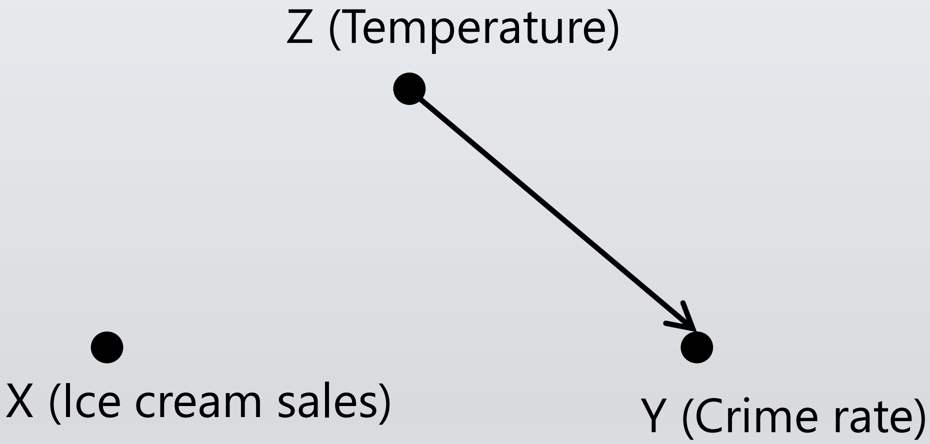
- In contrast observation
  - P(Z = z | Y = y) =
  - probability of event Z = z when knowing that Y = y
    Conditioning only filters on the data

Intervention changes the graph structure

- Observing high ice cream sales tells us something about the crime rate



- Intervention on ice cream sales does not change the crime rate
- The edge from X to its parents when using do(X) needs to be removed



Recap: Simpsons Paradox

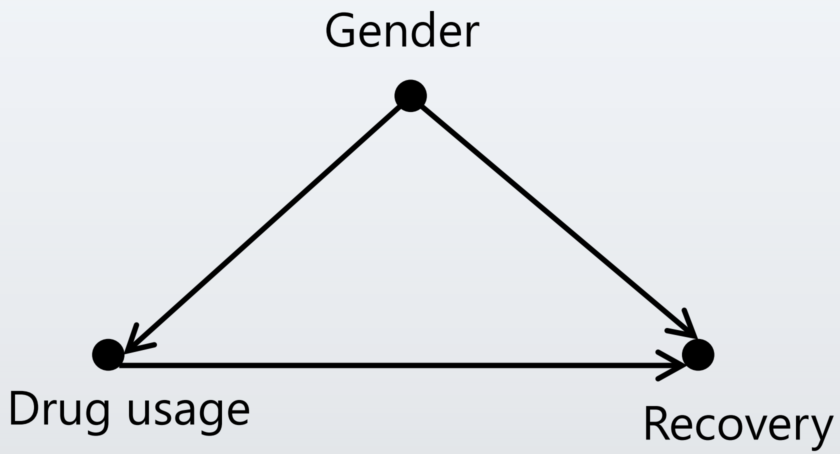
- Record recovery rates of 700 patients given access to a drug

	Recovery rate with drug	**Recovery rate without drug**
Men	81/87 (93%)	234/270 (87%)
Women	192/263 (73%)	55/80 (69%)
Combined	273/350 (78%)	289/350 (83%)

- Paradox:
  - For men, taking drugs has benefit
  - For women, taking drugs has benefit, too.
  - But: for all persons taking drugs has no benefit

Recap: Resolving the Paradox Formally

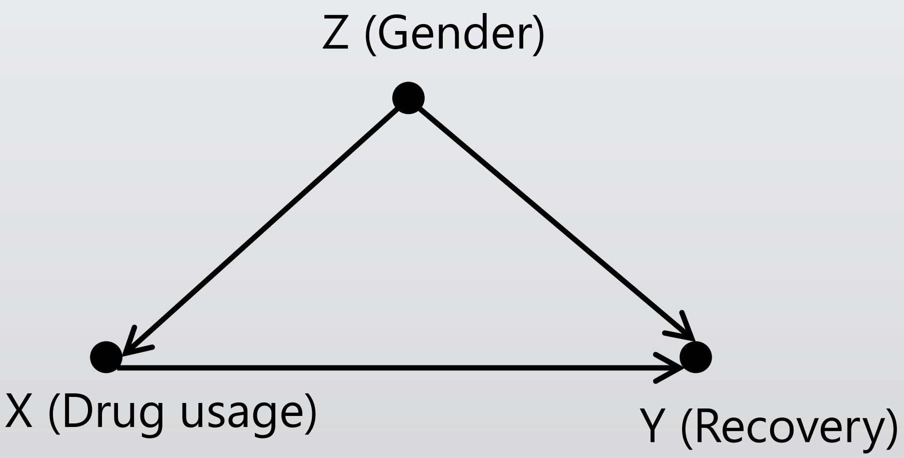
- We have to understand the causal mechanisms that lead to the data in order to resolve the paradox



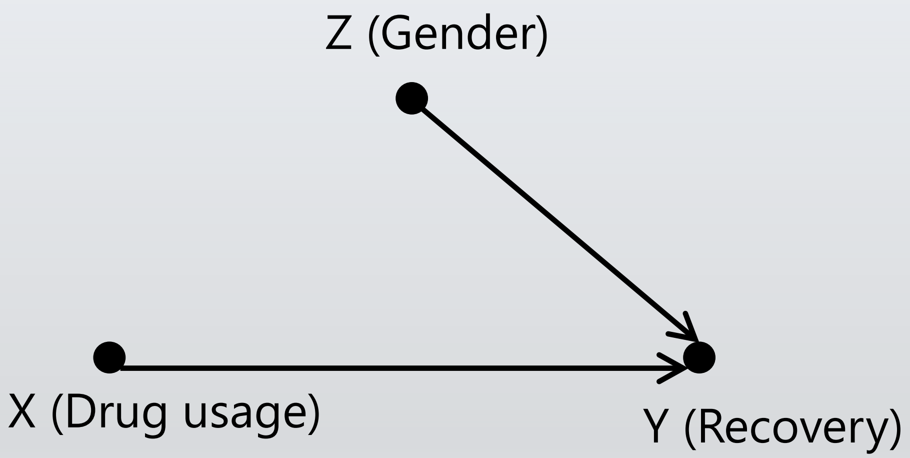
- Drug usage and recovery have common cause
- Gender is a confounder

Average Causal Effect (ACE)

- We would like to find out how effective the drug is in the population
- A hypothetical intervention would uniformly distribute the drug to the entire population and compare the recovery rate under complementary intervention
- First intervention denoted by do(X=1) second intervention denoted as do(X=0)
- Task is to compute: $P(Y = 1|do(X = 1)) - P(Y = 1|X = 0)$
- $P(Y = y|do(X = x)) = P_m(Y = y|X = x)$



Original Model P



Manipulated Model P_m

Adjustment formula

- $P(Z = z)$ is invariant under the intervention
- $P(Y = y|Z = z, X = x)$ is invariant, because the process by which Y responds to X and Z remains the same regardless of whether X changes spontaneously or by manipulation
- Therefore
  - $P_m(Y = y|Z = z, X = x) = P(Y = y|Z = z, X = x)$ and $P_m(Z = z) = P(Z = z)$
- Also, we know that Z and X are d-separated in the modified model and therefore
  - $P_m(Z = z|X = x) = P_m(Z = z) = P(Z = z)$
- Putting it together
- $$\begin{aligned} P(Y = y|do(X = x)) &= P_m(Y = y|X = x) \\ &= \sum_z P_m(Y = y|Z = z, X = x)\, P_m(Z = z|X = x) \\ &= \sum_z P_m(Y = y|Z = z, X = x)\, P_m(Z = z) \\ &= \sum_z P(Y = y|Z = z, X = x)\, P(Z = z) \end{aligned}$$

Adjustment Formula

Definition

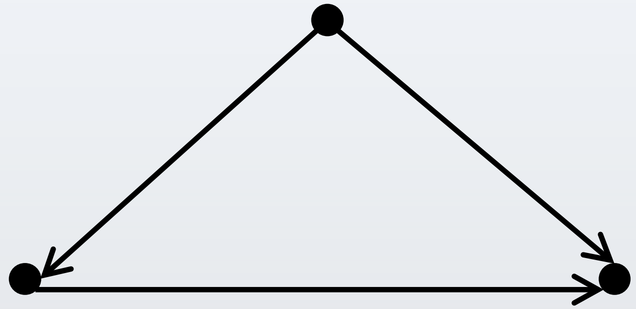
The adjustment formula (for single parent Z of X) for the calculation of the GCE is given by

$$P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, Z=z)\, P(Z = z)$$

- Wording: „Adjusting for Z" or „controlling Z"

Example of Drug and Recovery

$$
\begin{aligned}
P(Y=1 \mid do(X=1)) &= P(Y=1 \mid X=1, Z=1)P(Z=1) + P(Y=1 \mid X=1, Z=0)P(Z=0) \\
&= 0.93(87+270)/700 + 0.73(263+80)/700 \\
&= 0.832 \\
P(Y=1 \mid do(X=0)) &= 0.7818
\end{aligned}
$$



$$ACE = 0.832 - 0.7818 = 0.0502 > 0$$

One has to segregate the data w.r.t. Z (adjust for Z)

	Recovery rate with drug	**Recovery rate without drug**
Men	81/87 (93%)	234/270 (87%)
Women	192/263 (73%)	55/80 (69%)
Combined	273/350 (78%)	289/350 (83%)

Causal Effect Rule

- Pa(X) = parents of X
- z = instantiation of all parent variables of X

Rule (Calculation of causal effect)

$$P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, Pa(X) = z)\, P(Pa(X) = z)$$

Rule (Calculation of causal effect (alternative))

$$P(Y = y \mid do(X = x)) = \sum_z P(Y = y, X = x, Pa(X) = z) / P(X = x \mid Pa(X) = z)$$

Backdoor Criterion (Motivation)

- Intervention on X requires adjusting parents of X
- But sometimes those variables are not measurable (though perhaps represented in graph)

- Need more general criterion
  to identify adjustment variables
  1. Block all spurious paths between X and Y
  2. Leave all directed paths from X to Y unperturbed
  3. Do not create new spurious paths

Backdoor Criterion (Formulation)

- Can adjust for Z satisfying backdoor criterion
- $P(Y = y \mid do(X = x)) = \sum_z P(Y = y \mid X = x, Z = z)P(Z=z)$

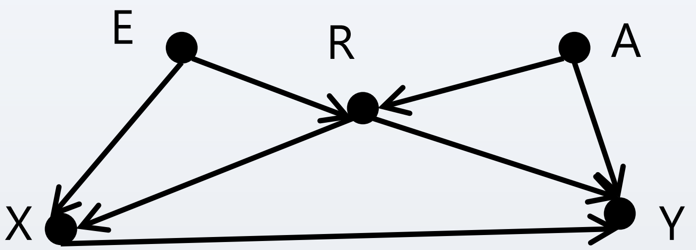
Definition

Set of variables Z satisfies backdoor criterion relative to a pair (X,Y) of variables iff

1. No node in Z is a descendant of X and
2. Z blocks every path between X and Y that contains an arrow into X

Quiz

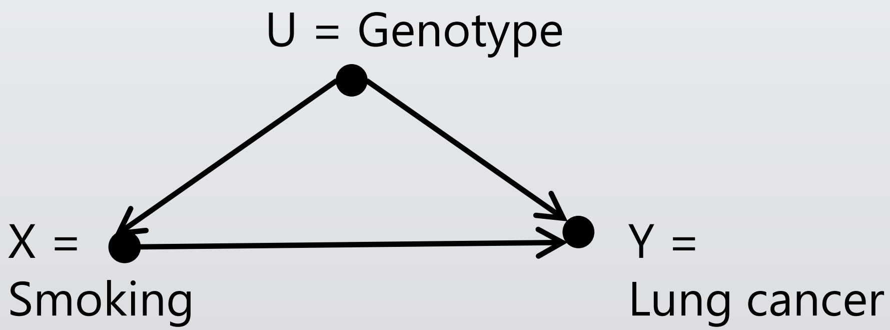
- Sometimes also need to condition on colliders
- There are four backdoor paths from X to Y
  - X ← E → R → Y
  - X ← E → R ← A → Y
  - X ← R → Y
  - X ← R ← A → Y
- What are potential blocking sets:

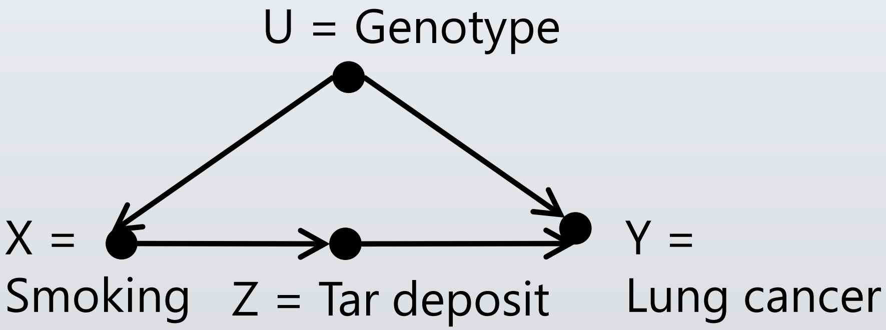


18

Front-door Criterion Motivation

- The do-Operator can be applied to scenarios that do not satisfy the backdoor criterion
- Consider the following example
- We would like to know P(Y=y|do(X=x))
- It is not possible to know which portion of the observed correlation between X and Y is spurious
- An intermediate variable can help us together with the front-door criterion





Example: Smoking Lobby

	Tar (400)		No tar (400)		All subjects (800)	
	Smokers (380)	Nonsmokers (20)	Smokers (20)	Nonsmokers (380)	Smokers (400)	Nonsmokers (400)
No cancer	323 (85%)	1 (5%)	18 (90%)	38 (10%)	341 (85%)	39 (9.75%)
Cancer	57 (15%)	19 (95%)	2 (10%)	342 (90%)	59 (15%)	361 (92.25%)

Tobacco industry argues:

- 15% of smoker w/ cancer < 92.25% nonsmoker w/ cancer
- Tar: 15% smoker w/ cancer < 95% nonsmoker w/ cancer
- Non tar: 10% smoker w/ cancer < 90% nonsmoker w/ cancer

Antismoking lobby argues:

- Choosing to smoke increases chances of tar deposit (95% = 380/400)
- Effect of tar deposit: look separately at smokers vs. Non-smokers
- Smokers: 10 % cancer 15 % cancer
- Nonsmokers: 90 % cancer 95 % cancer

Who is right?

Front-door criterion

- Idea: Separate the effects X on Y into X on Z and Z on Y
- Both individual effects can be assed
- X on Z: Easy, since there is not backdoor path from X to Z (adjustment on empty set)
  - P(Z=z|do(X=x))=P(Z=z|X=x)

- Z on Y: backdoor path Z←X←U→Y can be blocked by conditioning on X
  - $P(Y = y|do(Z = z)) = \sum_x P(Y = y|Z = z, X = x)$

- Now we chain the effects.
  - $P(Y = y|do(X = x)) = \sum_z P(Y = y|do(Z = z))P(Z = z|do(X = x))$
  - $P(Y = y|do(X = x)) = \sum_z \sum_{x'} P(Y = y|Z = z, X = x')P(X = x')P(Z = z|X = x)$

Front-door Criterion (Formulation & Theorem)

UNIVERSITÄT ZU LÜBECK

Definition

Set of variables Z satisfies front-door criterion w.r.t. pair of variables (X,Y) iff

1. Z intercepts all directed paths from X to Y
2. Every backdoor path from X to Z is blocked (by collider)
3. All Z-Y backdoor paths are blocked by X

Theorem (Front-door adjustment)

If Z fulfills front-door criterion w.r.t. (X,Y) and $P(x,z) > 0$

then $P(y|do(x)) = \sum_z P(z|x) \sum_{x'} P(y|z, x')P(x')$

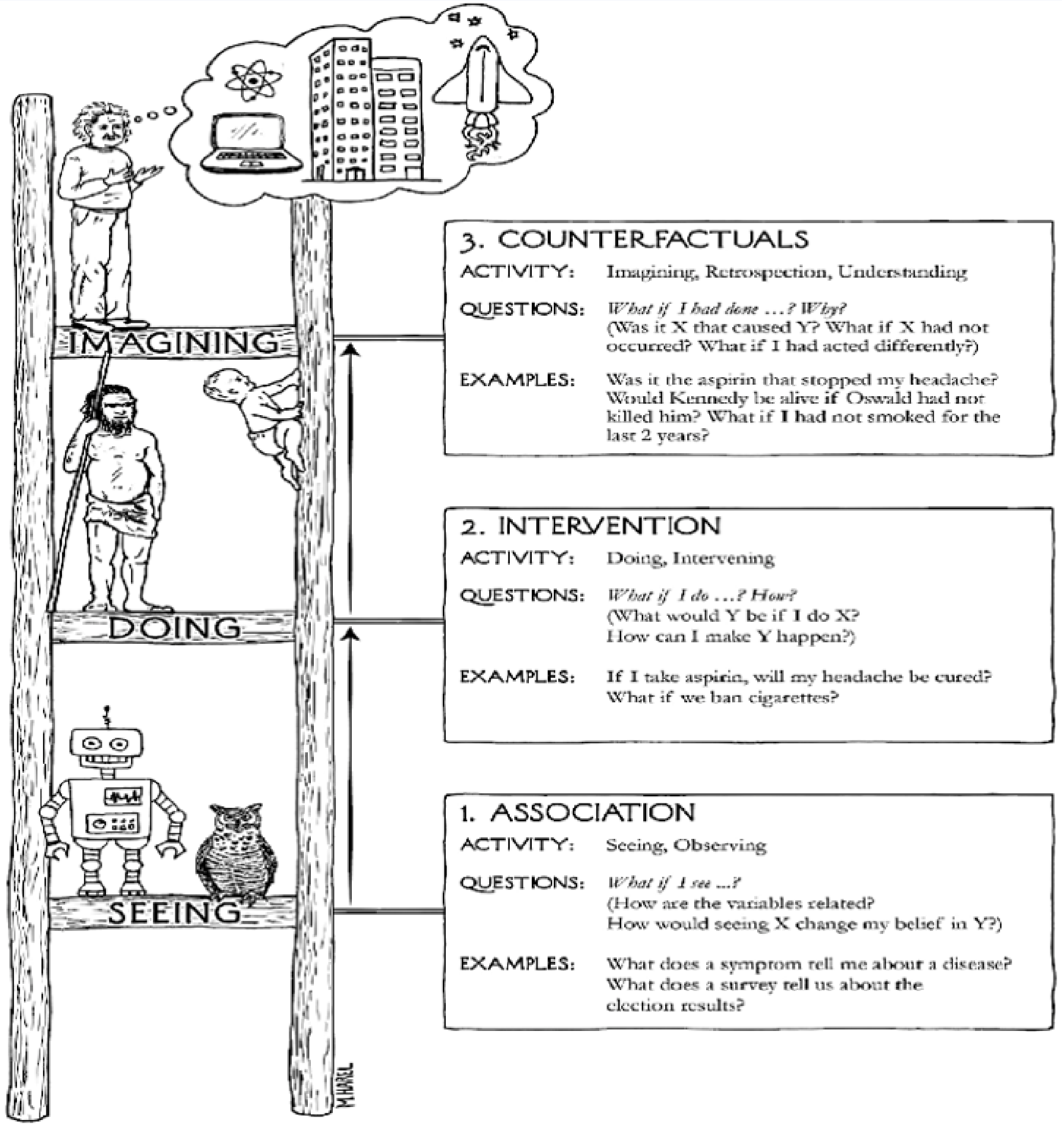
Pearl's Causal Hierarchy

The Ladder of Causality

"Actual" Causality

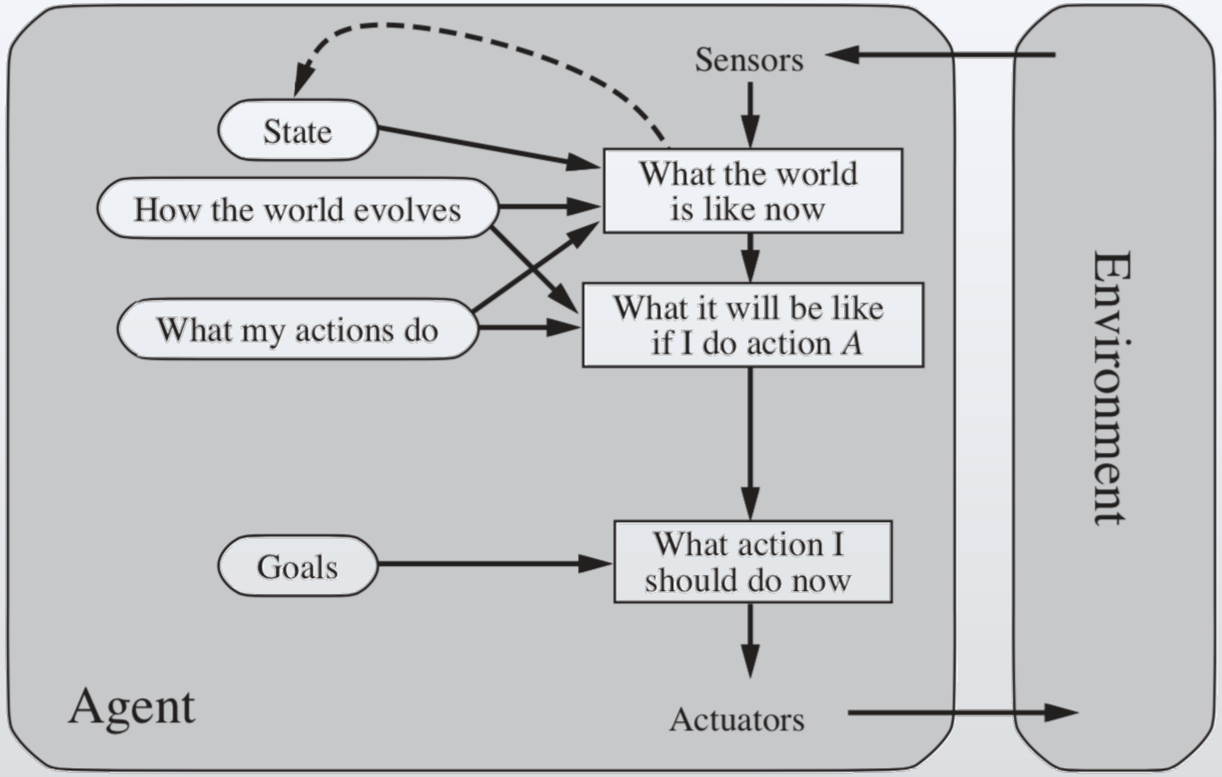
"Causality-in-mean"

Statistics



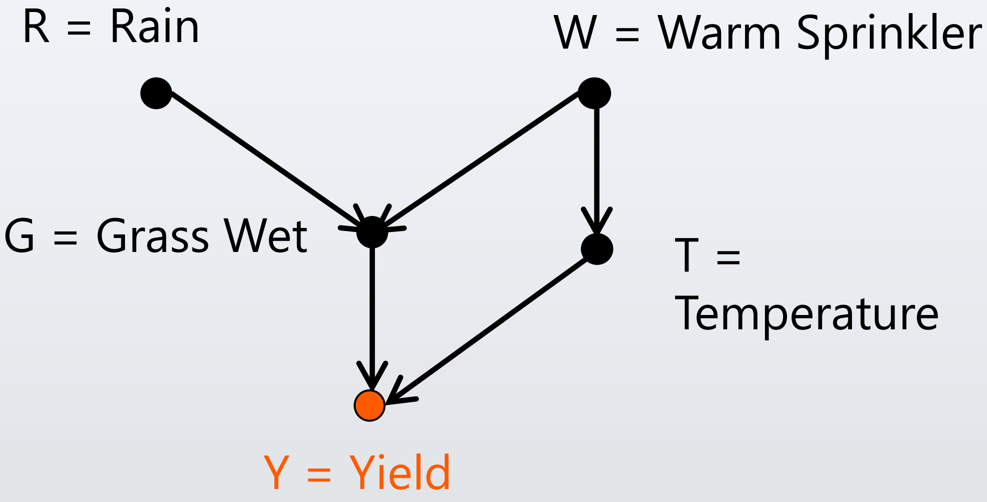
Causality – an agent perspective

- Models play an important role within an agent
- We have encountered cases where the agent is given a model of the environment or where the agent learns a model (Adaptive Dynamic Programming)
- When agent can freely act without limitations (e.g., trying out computer games) we are in a situation of unlimited random control group experiments (gold standard)
- When agent acts in the real world trying out things has consequences



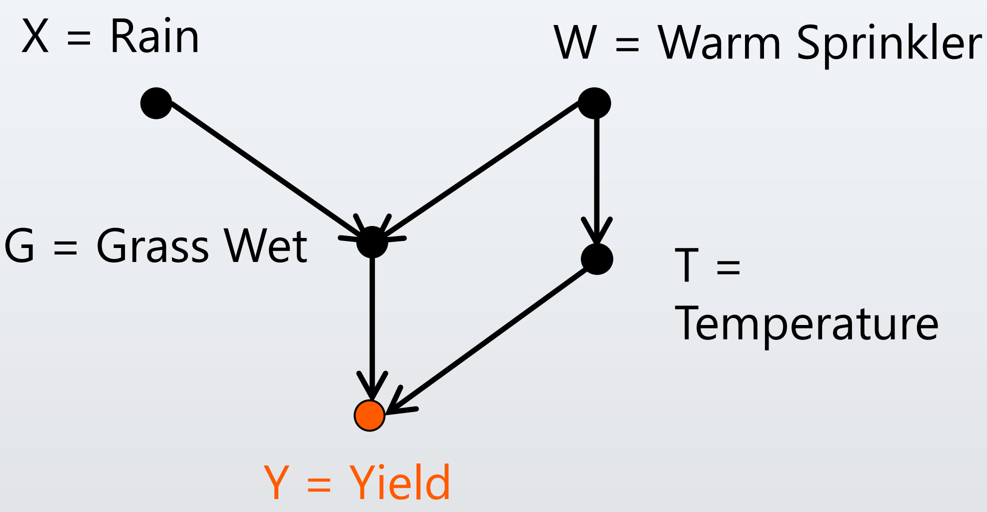
Formalizing a maximum utility action selection in a BN

- We have $X_i \in \boldsymbol{X}$ variables in the BN
- $D(X_i)$ is the domain of the variables
- State space S = $\bigotimes_{X_i \in \boldsymbol{X}} D(X_i)$ is the set of all possible combinations of values that the variables in the network can take
- Action space A consisting of applicable do operations
- Reward function R(s) that can contain individual rewards for all possible s ∈ S but an also only focus on individual variables (e.g., R(y))
- The agent is selecting an action that is maximizing the expected utility
- $\underset{a}{\text{argmax}}\, P(S|\text{a})\text{R(s)}$



Setting up an example

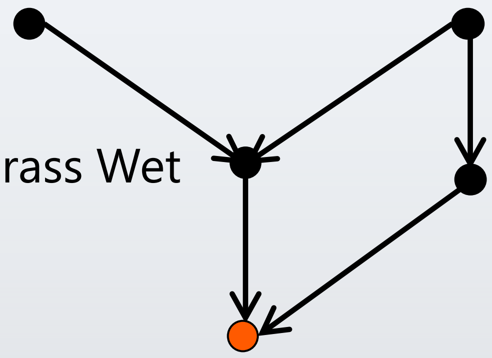
- D(R) = {yes, no}, D(W) = {yes, no}, D(T) = {low, med, high}, D(G) = {yes, no}, D(Y) = {low, med, high, exceptional}



Quiz

- D(R) = {yes,no}, D(W) = {yes,no}, D(T) = {low, med, high}, D(G) = {yes,no}, D(Y) = {low,med,high,exceptional}

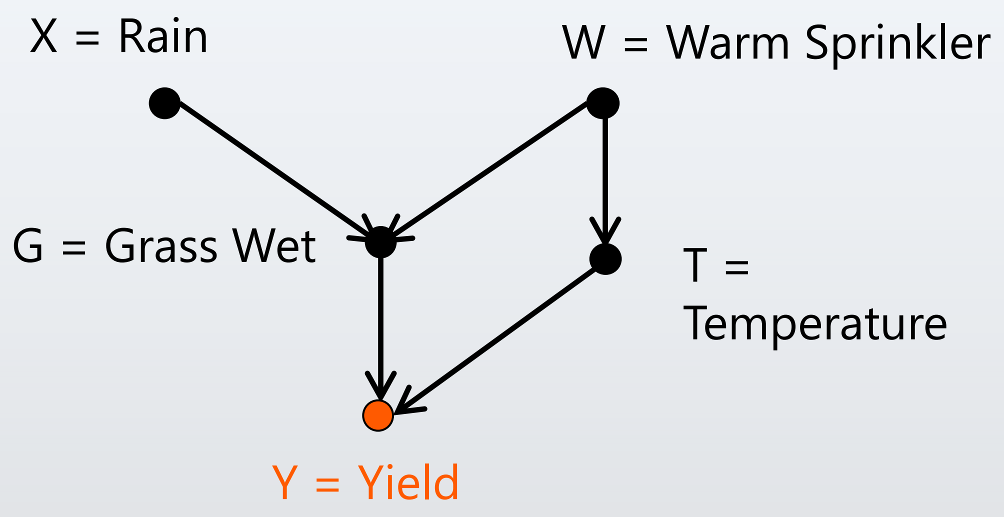
How many possible states has the state space?



Setting up an example

- D(R) = {yes, no}, D(W) = {yes, no}, D(T) = {low, med, high}, D(G) = {yes, no}, D(Y) = {low, med, high, exceptional}
- State space has 96 states
- R(S)→R(Y) only dependent on Y

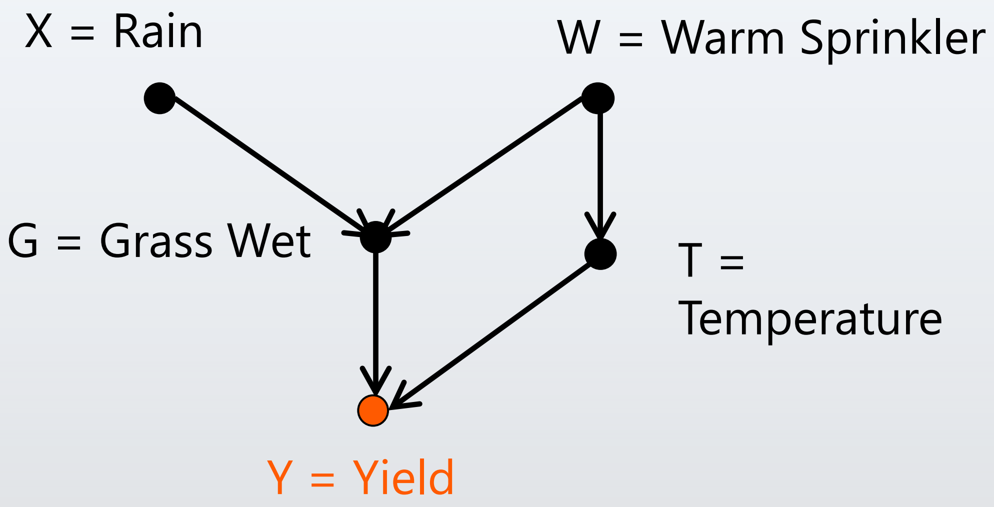
- Agent has two actions do(W=yes), do(G=yes) and can only perform one
- What is the best action to get maximum utility



Setting up an example

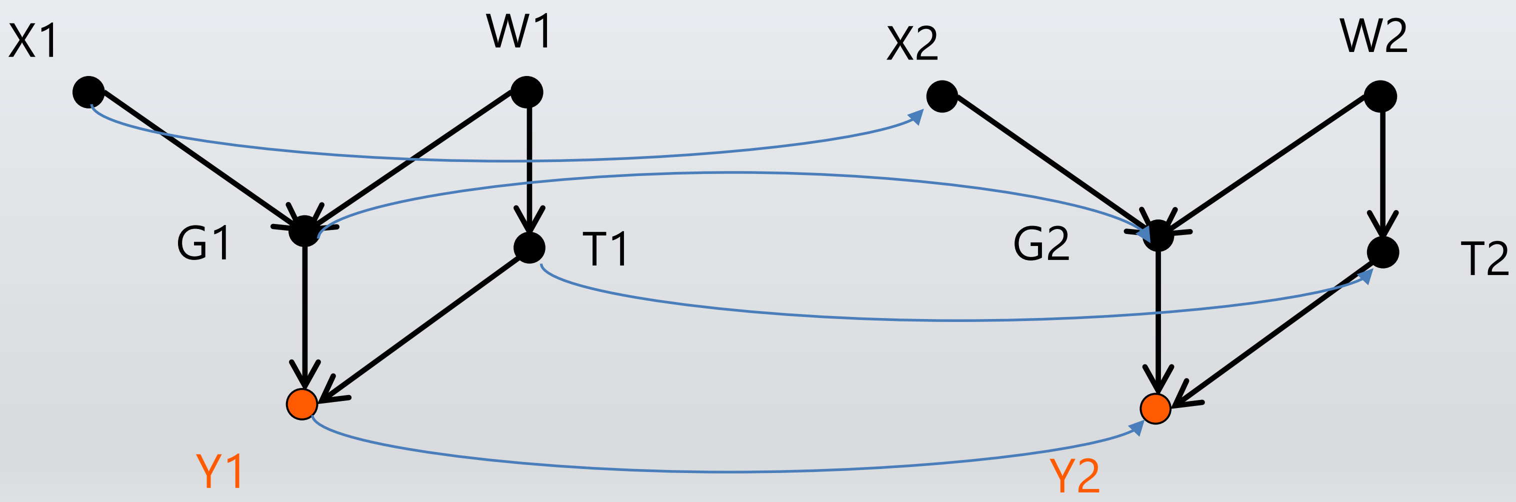
- Utility of action do(G=yes)
  →P(Y|do(G=yes))R(Y)
- Adjustment formula needed to block path to warm sprinkler
- P(Y|do(G=yes)) =
  $\sum_w P(Y = y|W = w, G = yes)\, P(W = w)$

- Utility of action do(W=yes)
  → P(Y|do(W=yes))R(Y)
- No adjustment needed



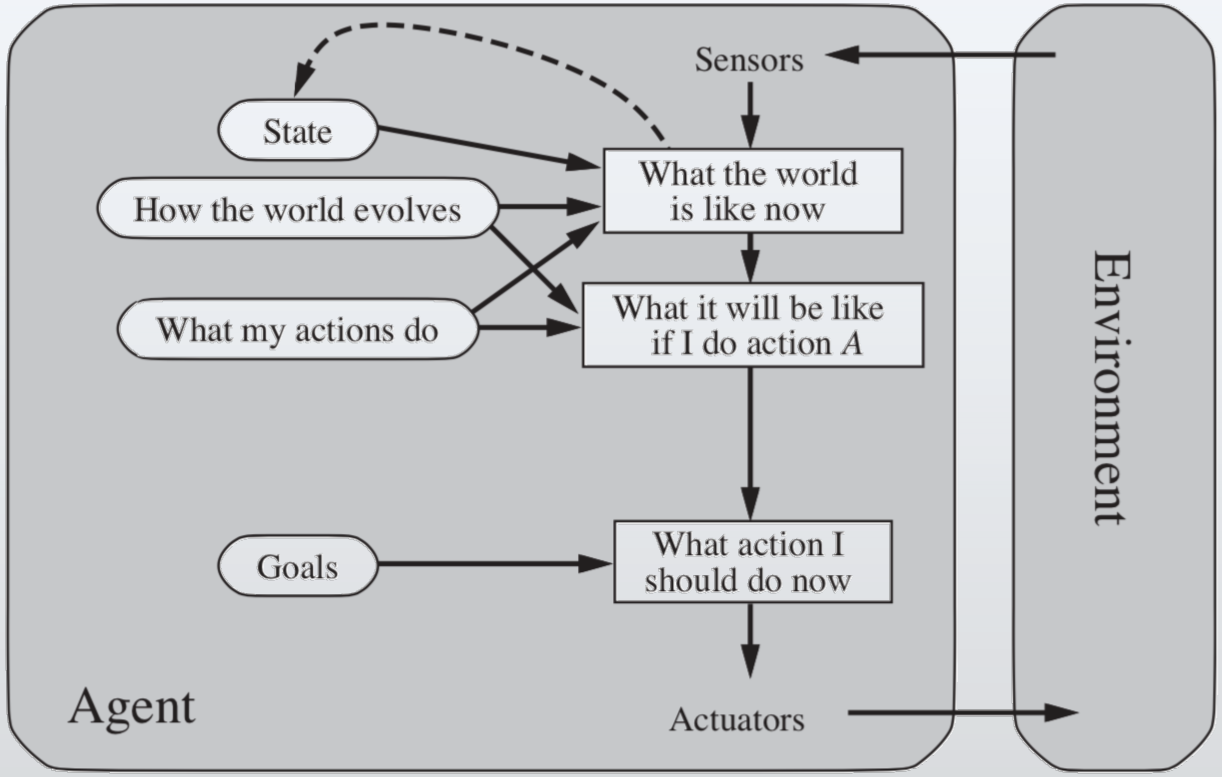
Of course, the example can be converted to a dynamic planning and acting cycle

- Actions over time → Plan/Conditional Plan
- Temporal combination of rewards (additive vs. discounted) as we know from previous lectures
- Planning/Acting Horizon
- Optionally also observable evidence over time

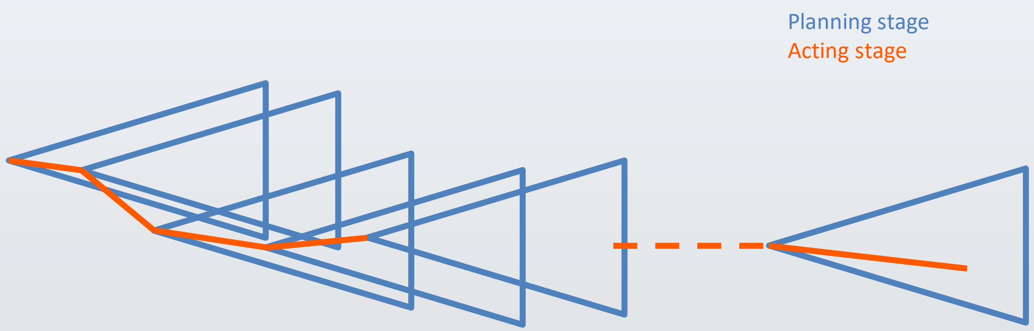


Now we are back to the beginning

Agent



Planning and Acting



Further interesting topics

- **Causal reinforcement learning**
  Lu, C., Schölkopf, B., & Hernández-Lobato, J. M. (2018). Deconfounding reinforcement learning in observational settings. *arXiv preprint arXiv:1812.10576*.

  Gasse, M., Grasset, D., Gaudron, G., & Oudeyer, P. Y. (2021). Causal reinforcement learning using observational and interventional data. *arXiv preprint arXiv:2106.14421*.

- **Causal inverse reinforcement learning**
  Ruan, K., Zhang, J., Di, S., & Bareinboim, E. (2022). Causal Imitation Learning Via Inverse Reinforcement Learning.

  Zhang, J., Kumor, D., & Bareinboim, E. (2020). Causal imitation learning with unobserved confounders. *Advances in neural information processing systems*, *33*, 12263-12274.

Content

1. Planning and Acting with **Deterministic** Models
2. Planning and Acting with **Refinement** Methods
3. Planning and Acting with **Temporal** Models
4. Planning and Acting with **Nondeterministic** Models
5. **Standard** Decision Making
6. Planning and Acting with **Probabilistic** Models
7. **Advanced** Decision Making
8. **Human-aware** Planning
9. Intro to **Causality**
10. Intervention and **Causal Planning**