Causal Networks

with slides by J. Pearl, E. Rosen, M Uflacker,
J. Huegle, C. Schmidt, S.Tyrvainen

Causality is central notion in science, decision-taking and daily life.

How to reason formally about cause and effect?

Question: How do you define cause and effect?

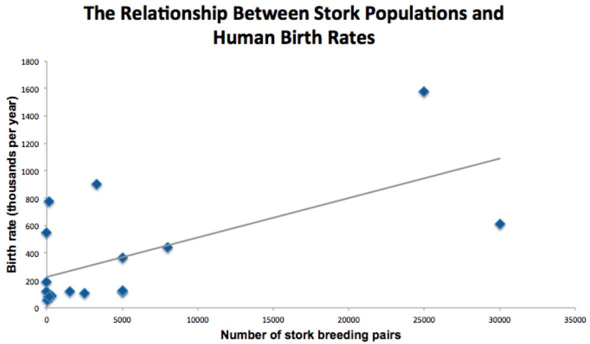
The subject of causality has a long history in philosophy.

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

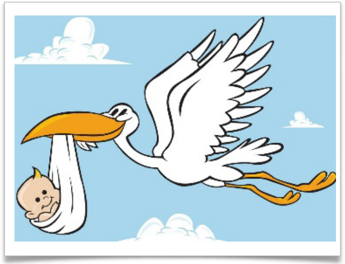
David Hume, A Treatise of Human Nature (1738)

Do storks deliver babies?

Country	Area (km^2)	Storks (pairs)	Humans (10^6)	Birth rate (10^3/yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576

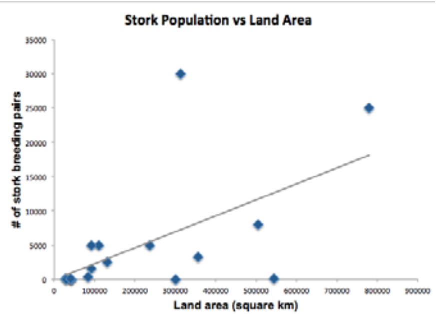


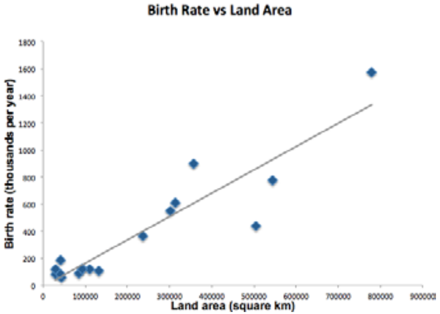
"Highly **statistically significant** degree of correlation between stork populations and birth rates" (or in technical terms, $p = 0.008$)



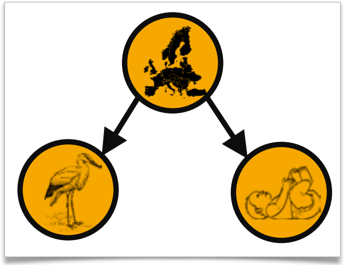
Do storks deliver babies?

Country	Area (km^2)	Storks (pairs)	Humans (10^6)	Birth rate (10^3/yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576

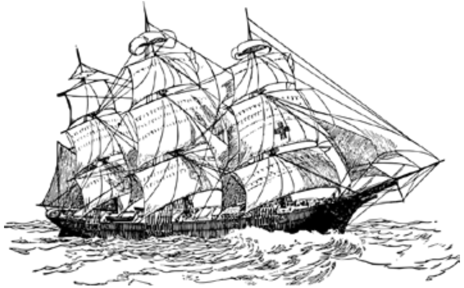




But a **simple variable that affects both** the birth rate and the stork population is the size of each country.

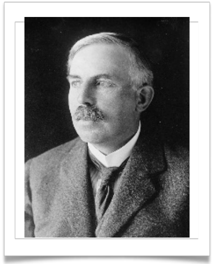


Challenge in Statistics: Draw conclusions from data



- James Lind (1716-1794): **How to treat scurvy?**
  - Scurvy results from a lack of vitamin C
  - 12 scorbutic sailor treated with different acids, e.g. vinegar, cider, lemon
  - Only the condition of the sailor treated by lemon improved

- *"If your experiment needs statistics, you ought to have done a better experiment."* Ernest Rutherford (1871-1937)



But: What if you cannot do a *randomized experiment* or receive ambiguous results?

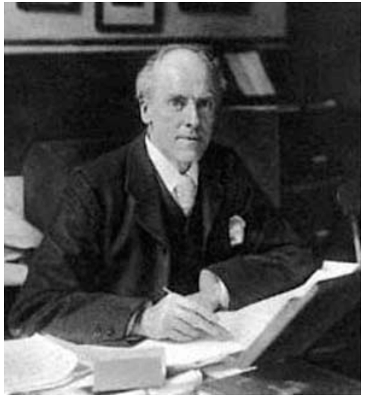
Use statistical tests to validate your hypothesis

Check whether it is statistically significant that

$$P(recovery \mid lemons) > P(recovery \mid no\ lemons)$$

Or in other words:

“Is there a dependence between recovery and the treatment with lemons?”



“Beyond such discarded fundamentals as `matter' and `force' lies still another fetish amidst the inscrutable arcana of even modern science, namely, the category of cause and effect.”

Karl Pearson (1857-1936)

Correlation does not imply causation.

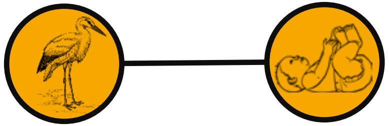
Since then, many statisticians tried to avoid causal reasoning

- *“Considerations of causality should be treated as they have always been in statistics: preferably not at all.”* (Terry Speed, 1990)
- *“It would be very healthy if more researchers abandon thinking of and using terms such as cause and effect.”* (Bengt Muthen, 1987)

But dependence says us something about causation:

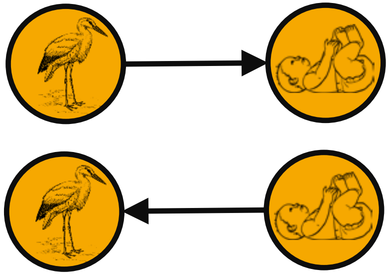
“Common Cause Principle”
Hans Reichenbach (1891-1953)

If there is a statistical dependence between variables X and Y, e.g.,

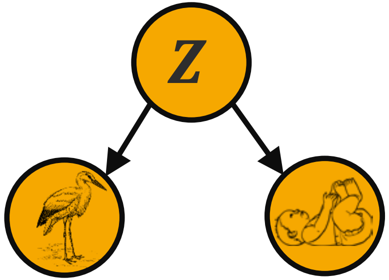


then either

- X causally influences Y (or vise versa), e.g.,



- or there exists Z causally influencing both, e.g.,

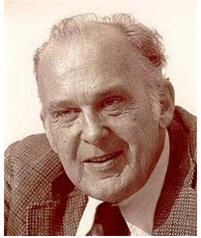


- The modeling of the underlying structures provides a language to encode causal relationships – the basis of a causality theory.
- Causality theory helps to decide when, and how, causation can be inferred from domain knowledge and data.

Some people who contributed to causality theories:

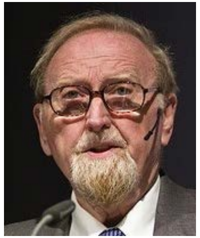
Donald Rubin (*1943)

Judea Pearl (*1936)



Donald Campbell (1916-1996)

Dawid Philip (*1946)



Clive Granger (1934-2009)

“[…] all approaches to causation are variants or abstractions of […] structural theory […].” Judea Pearl

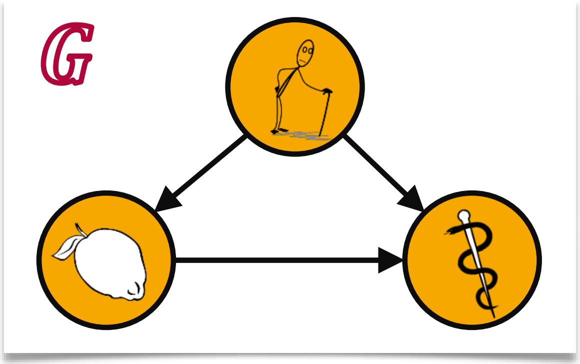
"Causality, although widely used, does not seem to be well-defined" (Lindley and Novick, 1981)

Problem: Probability theory has an associational, and not a causal nature.

To see this:

- Recap the scurvy experiment
- Assume that the data is generated by model $\mathbb{G}$.
  - The **recovery** of the scurvy is causally influenced by the **treatment with lemons**.
  - But now, both the **recovery** of scurvy as well as the **treatment with lemons** are causally influenced by the **age of the sailors**.
- The question remains:

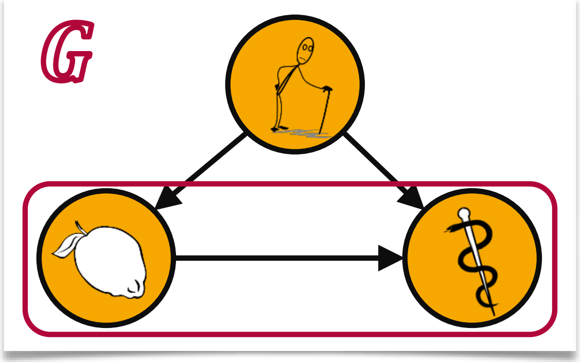
Should we treat scurvy with lemons?



- We run an experiment w.r.t. the model $\mathbb{G}$,
  i.e., we favor old sailors for treatment with lemons.

- The observed data of all sailors:

Combined	Recovery	No Recovery	Total	**Recovery Rate**
No lemons	20	20	40	50 %
Lemons	16	24	40	40 %
Total	36	44	80	



- Hence, we see that
  $$P(\boldsymbol{recovery}|\boldsymbol{lemons}) < P(\boldsymbol{recovery}|\boldsymbol{no\ lemons})$$

➡ Should we treat scurvy with lemons?

- The observed data of old sailors:

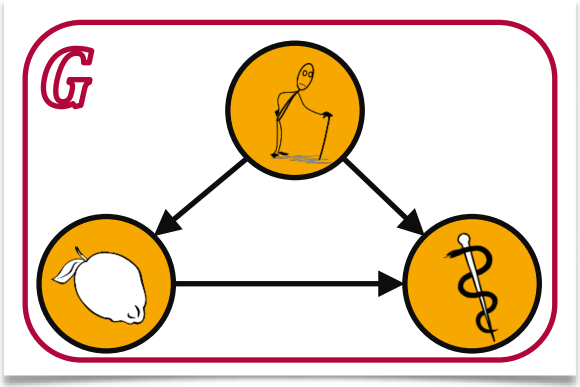
Old	Recovery	No Recovery	Total	Recovery Rate
No lemons	2	8	10	20 %
lemons	9	21	30	30 %
Total	11	29	40	

➡ $P(recovery|lemons, old) > P(recovery|no\ lemons, old)$

- The observed data of young sailors:

Young	Recovery	No Recovery	Total	Recovery Rate
No lemons	18	12	30	60 %
Lemons	7	3	10	70 %
Total	25	15	40	

➡ $P(recovery|lemons, young) > P(recovery|no\ lemons, young)$

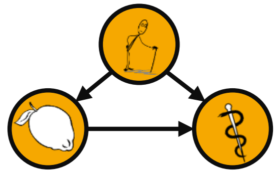


➡ Should we treat scurvy with lemons?

- This reversal of the association between two variables after considering the third variable is called **Simpson's Paradox**.

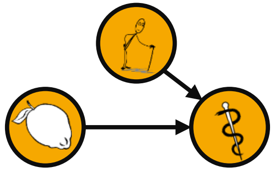
➡ How to resolute the paradox and find an answer?

Observational Regime



vs.

Interventional Regime



- In an interventional regime, all influences stemming from "natural causes" of the exposure variable are removed (e.g., see randomized experiments).
- Pearl extends probability calculus by introducing a new operator for describing interventions, the **do-operator**.

Example:

$P(lung\ cancer|smoke)$
Probability somebody gets lung cancer, **given that he smokes**.

$P(lung\ cancer|do(smoke))$
Probability somebody gets lung cancer, **if we force the person to smoke**.

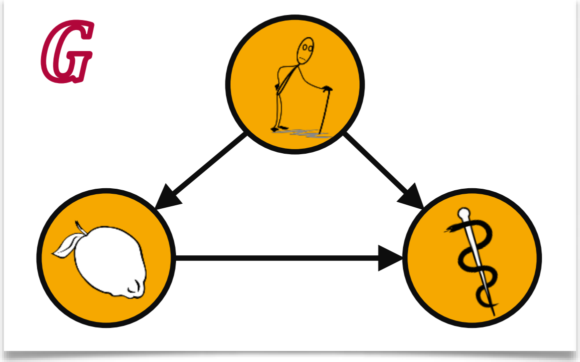
Resolution of the Simpson's paradox

- Simpson's paradox is only paradoxical if we misinterpret

  $P(recovery|lemons)$ as $P(recovery|do(lemons))$

- We should treat scurvy with lemons if

  $P(recovery|do(lemons)) > P(recovery|do(no\ lemons))$



Derivation of the do-operator

- If *identifiable*,

  $P(\cdot\,|do(\cdot))$ can be calculated from G and observational Data

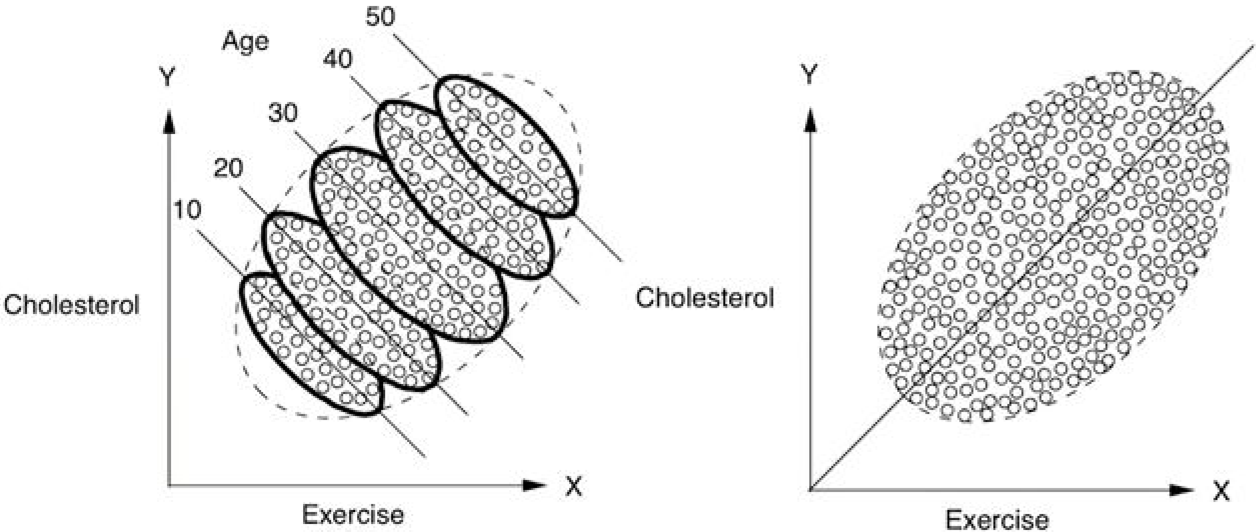
- In our example, we have

$$P(recovery|do(lemons)) = \sum_{age} P(age)\,P(recovery|age, lemons) = 0.5$$

$$P(recovery|do(no\ lemons)) = \sum_{age} P(age)\,P(recovery|age, no\ lemons) = 0.4$$

➡ We should treat scurvy with lemons!

WHY DATA CAN BE DUMB



Exercise is helpful in every age group but harmful for a typical person.

CAUSATION AS A PROGRAMMER'S NIGHTMARE

Input:

1. “If the grass is wet, then it rained”
2. “if we break this bottle, the grass will get wet”

Output:

“If we break this bottle, then it rained”

Bayesian Network for a Simple Conversation

Q1: If the season is dry, and the pavement is slippery, did it rain?

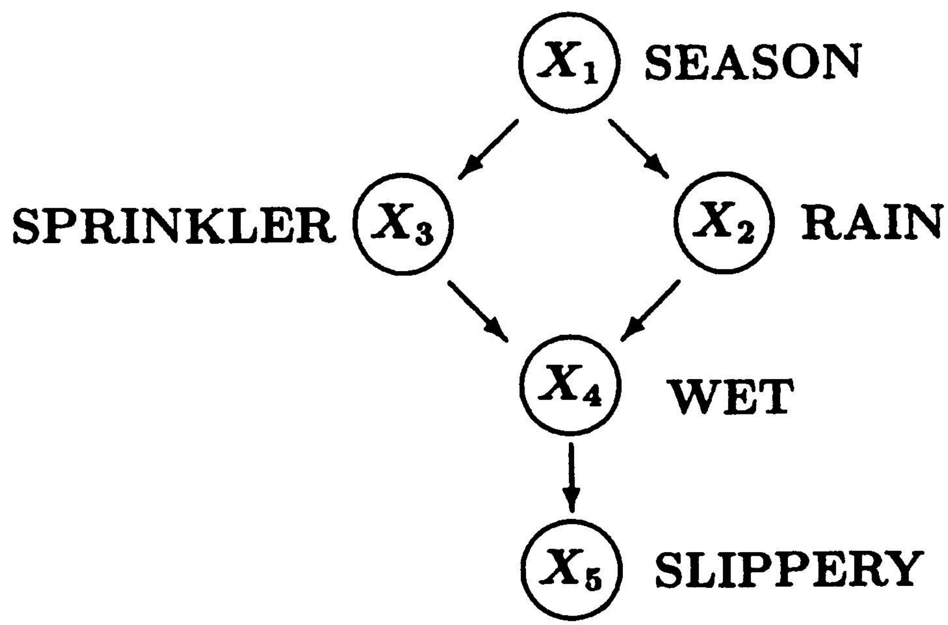
A1: Unlikely, it is more likely the sprinkler was ON.

Q2: But what if we SEE that the sprinkler is OFF?

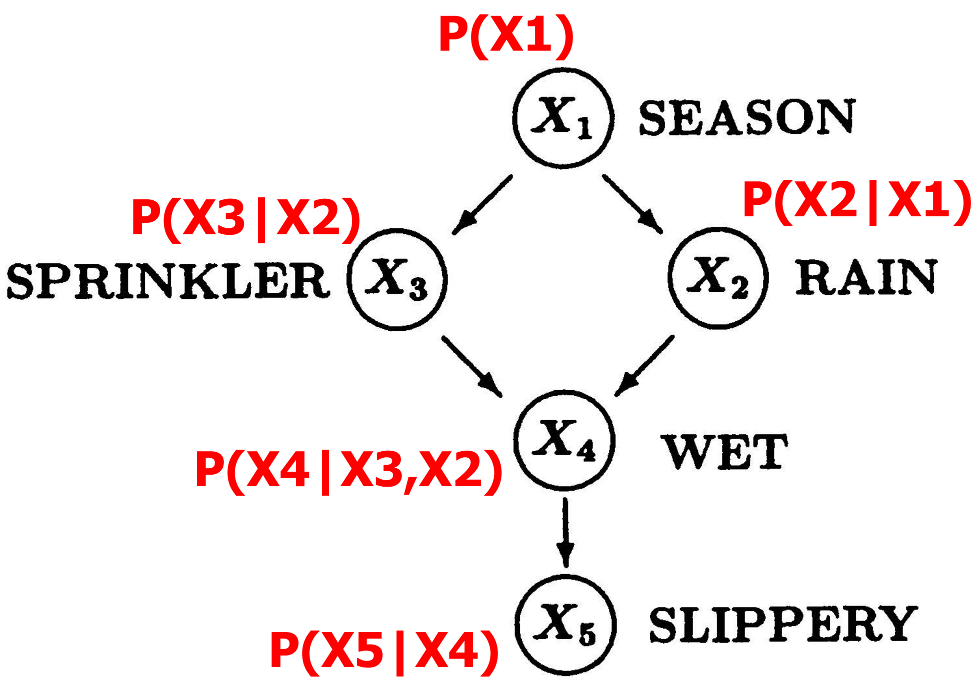
A2: Then it is more likely that it rained

.

The Story



Bayesian Network for a Simple Conversation



CPD:

X3	X2	wet=0,	wet=1
0	0	0.9	0.1
0	1	0.1	0.9
1	0	0.2	0.8
1	1	0	1

P(X1,X2,X3,X4,X5))= *P(X1) P(X2|X3) P(X3|X1) P(X4|X3,X2) P(X5|X4)*

Conditional Independencies ➡ Efficient Representation

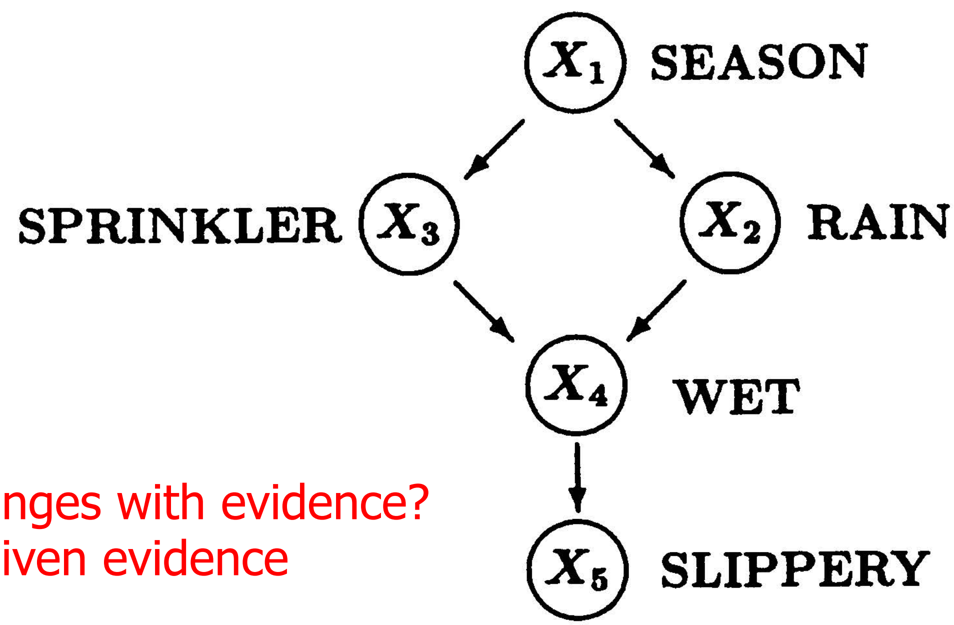
Bayesian Network for a Simple Conversation

Q1: If the season is dry, and the pavement is slippery, did it rain?

Q2: But what if we SEE that the sprinkler is OFF?

- **Belief updating**: how probability changes with evidence? What is more likely? Rain or not rain given evidence

The Story



$$P(X1,X2,X3,X4,X5)) = P(X1)\ P(X2|X3)\ P(X3|X1)\ P(X4|X3,X2)\ P(X5|X4)$$

Q1: Pr(rain=on | Slippery=yes, season= summer)?

Q2: Pr(rain=on | Slippery=off, season=winter)?

Bayesian Network for a Causal Conversation

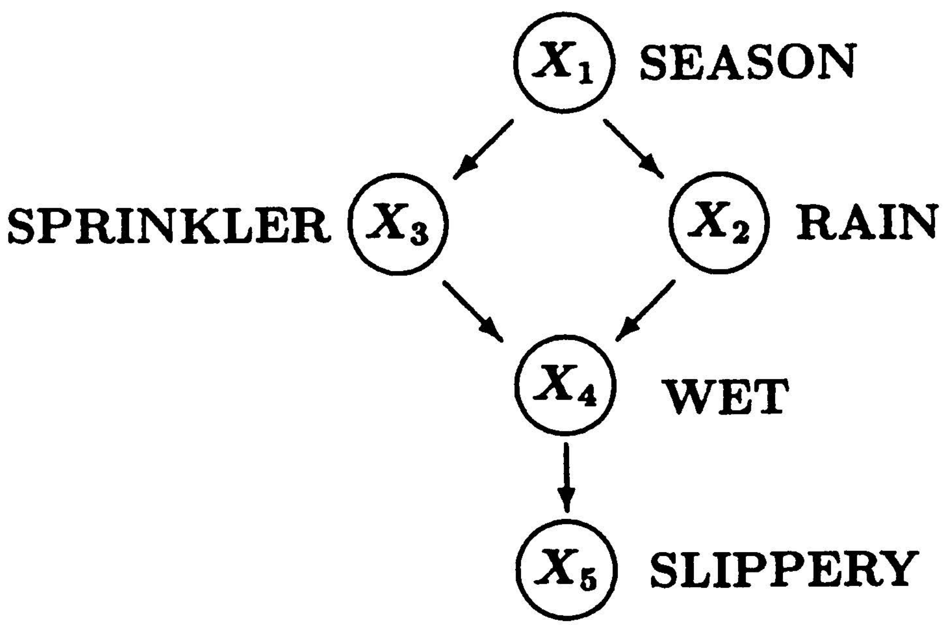
Q2: But what if we SEE that the sprinkler is OFF?

A2: Then it is more likely that it rained

Q3: Do you mean that if we actually turn the sprinkler OFF, the rain will be more likely?

A3: No, the likelihood of rain would remain the same

The Story



.

Observing (sprinkler=on) $\neq$ **Doing (sprinkler=on)**

NEEDED: ALGEBRA OF DOING

Available: algebra of seeing
e.g., What is the chance it rained if we see the grass wet?

$P\,(rain \mid wet) = ?$ $\{= P(wet|rain)\,\dfrac{P(rain)}{P(wet)}\}$

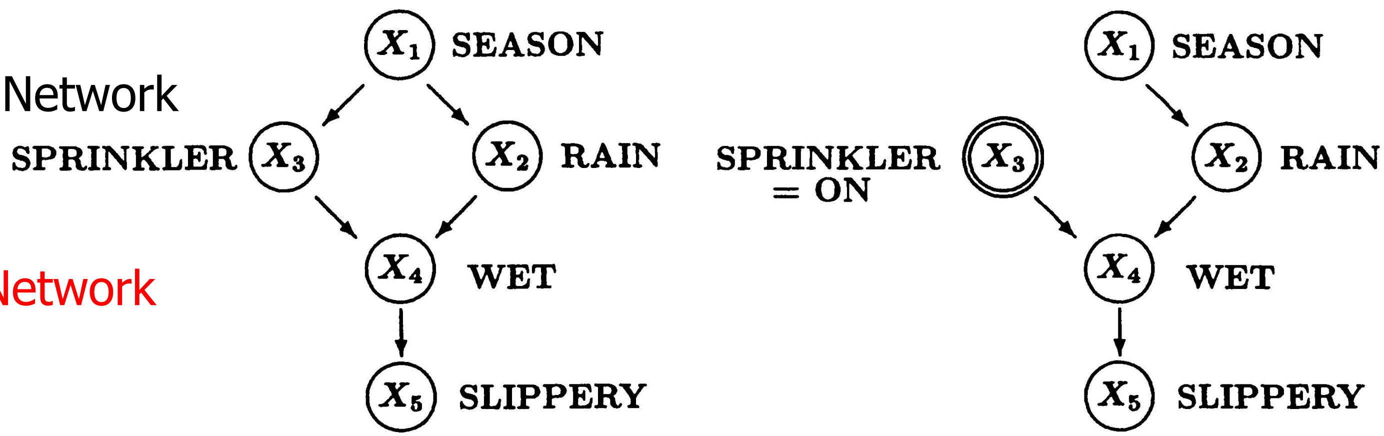
Needed: algebra of doing
e.g., What is the chance it rained if we make the grass wet?

$P\,(rain \mid do(wet)) = ?$ $\{= P\,(rain)\}$

Seeing vs. Doing

Bayesian Network

Causal Network



Effect of turning the sprinkler ON: P(season | do(sprinkler=on))

Pearl *do-calculus* leads to a complete mathematical framework for formulating causal models and for analyzing data to determine causal relationships.

Bayesian network

Bayesian networks are *Directed Acyclic Graphs* (DAGs) whose nodes represent random variables:

$$p(x_1, \ldots, x_d) = \prod_{j=1}^{d} p(x_j | x_{pa(j)})$$

Assumes that a variable is independent of previous non-parents given the parents, that is, $p(x_j | x_1, \ldots, x_d) = p(x_j | x_{pa(j)})$

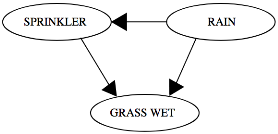
Captures how variables are conditionally dependent: If there are no any arrows between the nodes then they are independent:

$$p(A, B) = p(A)p(B)$$

The joint probability distribution:

$$p(G, S, R) = p(G|S, R)p(S|R)p(R)$$

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

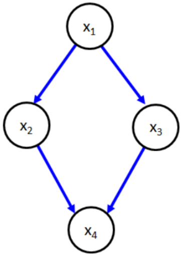
Sampling from a DAG

Ancestral sampling:

- Sample x_1 from $p(x_1)$
- If x_1 is a parent of x_2, sample x_2 from $p(x_2|x_1)$ Otherwise, sample x_2 from $p(x_2)$
- Go through the subsequent j in order sampling x_j from $p(x_j|x_{pa(j)})$

Conditional Sampling:

- easy if condition on the first variables: fix these and run ancestral sampling
- Hard if condition on the last variables: Conditioning on descendent makes ancestors dependent



It's hard to separate out causality from correlation

DAGs can be viewed as a causal process: the parents "cause" the children to take different values

The below equations are equivalent and the graphs have same conditional independences, but the causalities are not the same. Graphs tells us something useful that equations miss.

$$Y = X + 1$$
$$Z = 2Z$$

$$X = Y - 1$$
$$Y = Z/2$$

There is *observational data* ("seeing") and *interventional data* ("doing")

Usually the DAG is designed for observational data, but that ignores the possibility of unobserved variables, also without interventional data you can't distinguish the direction of causality.

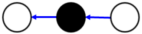
Simplest external intervention: a single variable is forced to take some fixed value (in a graph remove arrows entering that variable)

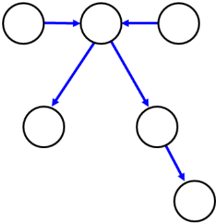
D-separation

d-separation is a criterion for deciding, from a causal graph, whether a set A of variables is independent of another set B (given a third set C)

$$A \perp\!\!\!\perp B | C$$

A and B are d-separated if for all paths P from A to B, at least one of the following holds:

- P includes a "chain" with an observed middle node
  
- P includes a "fork" with an observed parent node
  
- P includes a "v-structure" or "collider"



A and B are *d-separated*, give C, iff corresponding random variables are conditionally independent:

$$p(A, B|C) = p(A|C)p(B|C)$$

If A and B are not d-separated they are *d-connected*

The Causal Calculus (do-calculus, Pearl's Causal Calculus, Calculus of Actions)

Shortly: Calculus to discuss causality in a formal language by Judea Pearl

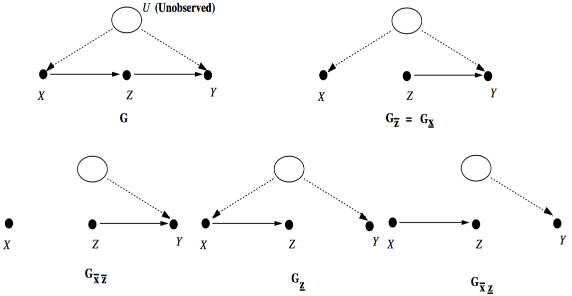
A new operator, *do()*, marks an action or an intervention in the model. In an algebraic model we replace certain functions with a constant $X = x$, and in a graph we remove edges going into the target of intervention, but preserve edges going out of the target.

The causal calculus uses Bayesian conditioning, $p(y|x)$, where x is observed variable, and causal conditioning, $p(y|do(x))$, where an action is taken to force a specific value x.

Goal is to generate probabilistic formulas for the effect of interventions in terms of the observed probabilities.

Judea Pearl, *Causality: Models, reasoning, and inference.* Cambridge University Press, 2000.

Notations



Notation: a graph G, W, X, Y, Z are disjoint subsets of the variables. $G_{\overline{X}}$ denotes the perturbed graph in which all edges pointing to X have been deleted, and $G_{\underline{X}}$ denotes the perturbed graph in which all edges pointing from X have been deleted. $Z(W)$ denote the set of nodes in Z which are not ancestors of W

Image: Judea Pearl

Pearl's 3 rules

Pearl's 3 rules

- Ignoring observations

$$p(y|do(x), z, w) = p(y|do(x), w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}$$

- Action/Observation exchange (the back-door criterion)

$$p(y|do(x), do(z), w) = p(y|do(x), z, w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \underline{Z}}}$$

- Ignoring actions/interventions

$$p(y|do(x), do(z), w) = p(y|do(x), w) \text{ if } (Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}, \overline{Z(W)}}}$$

Notation: a graph G, W, X, Y, Z are disjoint subsets of the variables. $G_{\overline{X}}$ denotes the perturbed graph in which all edges pointing to X have been deleted, and $G_{\underline{X}}$ denotes the perturbed graph in which all edges pointing from X have been deleted. $Z(W)$ denote the set of nodes in Z which are not ancestors of W

Intuition behind the Pearl's first rule

With condition $(Y \perp\!\!\!\perp Z|X, W)_{G_{\overline{X}}}$ we have

$$p(y|do(x), z, w) = p(y|do(x), w)$$

- Let's start with a simple case where we assume that there are no W or X. We get a condition $(Y \perp\!\!\!\perp Z)_G$, so Y is independent of Z, that is, $p(y|z) = p(y)$
- In the second case assume we have passively observed W, but no variable X: $(Y \perp\!\!\!\perp Z|W)_G$. Earlier we mentioned connection of d-separation and conditionally independent, that is, $p(y|z, w) = p(y|w)$
- The third case assume we don't know W, but we have X that's value is set by intervention: $(Y \perp\!\!\!\perp Z|X)_{G_{\overline{X}}}$. By the same theorem, that is, $p(y|z, do(x)) = p(y|do(x))$

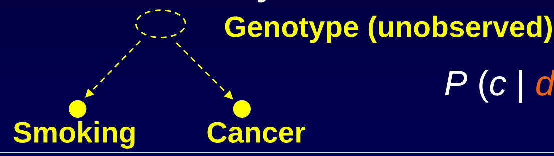
Combining these gives the full rule.

PREDICTING THE EFFECTS OF POLICIES

1. **Surgeon General (1964):**

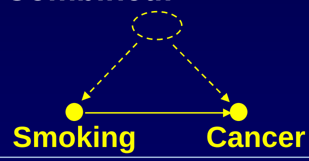
$$P\,(c \mid do(s)) \approx P\,(c \mid s)$$

2. **Tobacco Industry:**



$$P\,(c \mid do(s)) = P\,(c)$$

3. **Combined:**



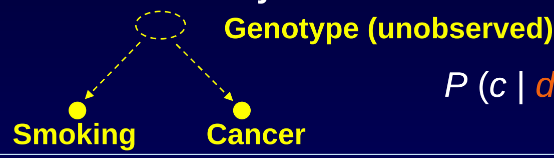
$$P\,(c \mid do(s)) = \text{noncomputable}$$

PREDICTING THE EFFECTS OF POLICIES

1. **Surgeon General (1964):**

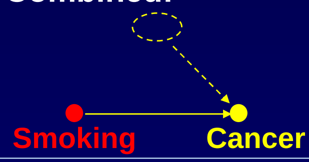
$$P(c \mid do(s)) \approx P(c \mid s)$$

2. **Tobacco Industry:**



$$P(c \mid do(s)) = P(c)$$

3. **Combined:**



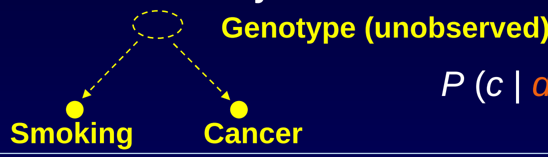
$$P(c \mid do(s)) = \text{noncomputable}$$

PREDICTING THE EFFECTS OF POLICIES

1. Surgeon General (1964):

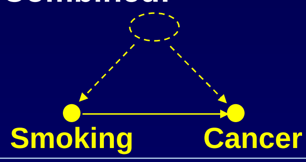
$P\,(c \mid do(s)) \approx P\,(c \mid s)$

2. Tobacco Industry:



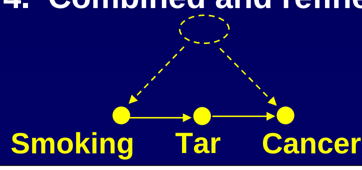
$P\,(c \mid do(s)) = P\,(c)$

3. Combined:



$P\,(c \mid do(s))$ = noncomputable

4. Combined and refined:



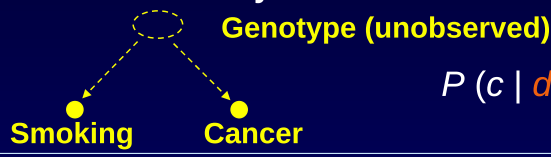
$P\,(c \mid do(s))$ = computable

PREDICTING THE EFFECTS OF POLICIES

1. Surgeon General (1964):

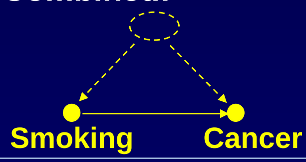
$P(c \mid do(s)) \approx P(c \mid s)$

2. Tobacco Industry:



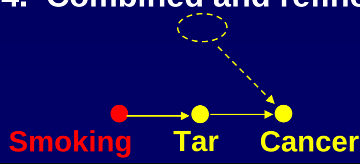
$P(c \mid do(s)) = P(c)$

3. Combined:

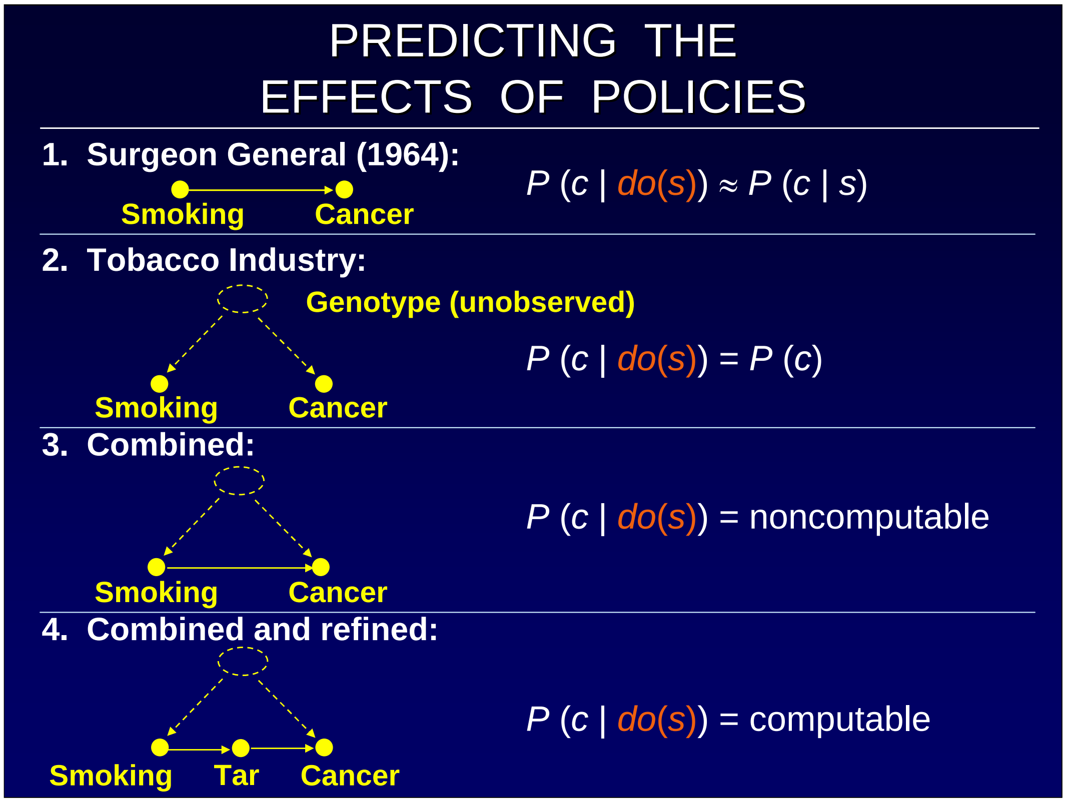


$P(c \mid do(s))$ = noncomputable

4. Combined and refined:



$P(c \mid do(s))$ = computable


PREDICTING THE
EFFECTS OF POLICIES
1. Surgeon General (1964):
Smoking
Cancer
$P(c \mid do(s)) \approx P(c \mid s)$
2. Tobacco Industry:
Genotype (unobserved)
Smoking
Cancer
$P(c \mid do(s)) = P(c)$
3. Combined:
Smoking
Cancer
$P(c \mid do(s))$ = noncomputable
4. Combined and refined:
Smoking
Tar
Cancer
$P(c \mid do(s))$ = computable

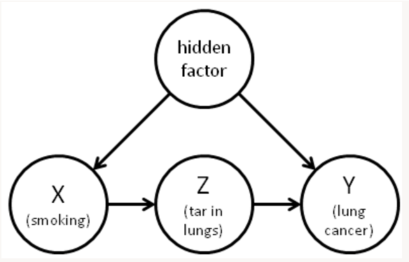
Example: Smoking and lung cancer

Randomized Controlled Trials (RCT)

- AKA Randomized Control Trial, Randomized clinical trial
- The participants in the trial are randomly allocated to either the group receiving the treatment under investigation or to the control group
- The control group removes the confounding factor of the placebo effect
- Double-blind studies remove further confounding factors
- Sometimes impractical or impossible

We can try to use causal calculus to analyze the probability that someone would get cancer given that they are smoking, without doing an actual RCT:

$$p(y|do(x))$$

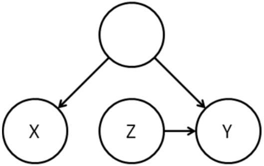


Note: We have no information about the hidden variable that could cause both smoking and cancer

Example

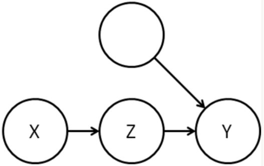
We can't try to apply rule 1 because there is no observations to ignore, we would just have $p(y|do(x)) = p(y|do(x))$.

Try apply rule 2: We would have $p(y|do(x)) = p(y|x)$, that is, the intervention doesn't matter. It's condition is $(Y \perp\!\!\!\perp X)_{G_{\underline{X}}}$:



Y and X are not d-separated, because they have a common ancestor.
$\Longrightarrow$ Rule 2 can't be applied

Try apply rule 3: We would have $p(y|do(x)) = p(y)$, that is, an intervention to force someone to smoke has no impact on whether they get cancer. It's condition is $(Y \perp\!\!\!\perp X)_{G_{\overline{X}}}$:

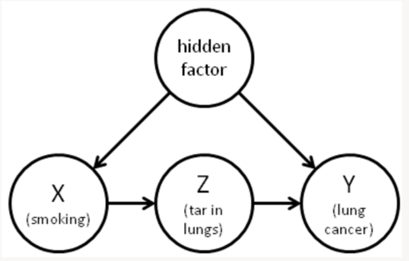


Y and X are not d-separated, because we have unblocked path between them.
$\Longrightarrow$ Rule 3 can't be applied

Example

New attempt:

$$
\begin{aligned}
p(y|do(x)) &= \sum_z p(y|z, do(x))p(z|do(x)) \\
&= \sum_z p(y|z, do(x))p(z|x) && \text{(rule 2: } (Z \perp\!\!\!\perp X)_{G_{\underline{X}}}) \\
&= \sum_z p(y|do(z), do(x))p(z|x) && \text{(rule 2: } (Y \perp\!\!\!\perp Z|X)_{G_{\overline{X},\underline{Z}}}) \\
&= \sum_z p(y|do(z))p(z|x) && \text{(rule 3: } (Y \perp\!\!\!\perp X|Z)_{G_{\overline{Z},\overline{X}}})
\end{aligned}
$$



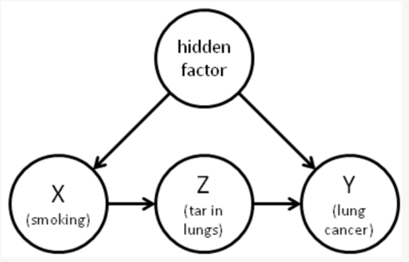
Example

We can use the same approach to the first term on the right hand side:

$$
\begin{aligned}
p(y|do(z)) &= \sum_x p(y|x, do(z))p(x|do(z)) \\
&= \sum_x p(y|x, z)p(x) \qquad\qquad \text{(rule 2 + rule 3)}
\end{aligned}
$$

Finally we can combine these results:

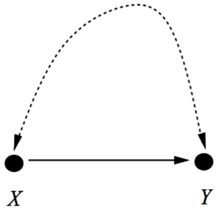
$$p(y|do(x)) = \sum_{z,x'} p(y|x', z)p(z|x)p(x')$$



We can now compare $p(y)$ and $p(y|x)$. The needed probabilities can be observed directly from experimental data: What part of smokers have lung cancer, how many of them have tar in their lungs etc.

Example: end

- The analysis would have not worked if the graph had missed the tar variable, Z, because there is no general way to compute $p(y|do(x))$ from any observed distributions whenever the causal model includes subgraph shown the figure below
- Causal Calculus can be used to analyze causality in more complicated (and more unethical) situations than RCT
- Causal Calculus can also be used to test whether unobserved variables are missed by removing all do terms from the relation
- Not all models are acyclic. See for example Modeling Discrete Interventional Data Using Directed Cyclic Graphical Models (UAI 2009) by Mark Schmidt and Kevin Murphy



RULES OF CAUSAL CALCULUS

Rule 1: Ignoring observations

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

if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}}}$

Rule 2: Action/observation exchange

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

if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}\underline{Z}}}$

Rule 3: Ignoring actions

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

if $(Y \perp\!\!\!\perp Z \mid X, W)_{G_{\overline{X}\,\overline{Z(W)}}}$

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

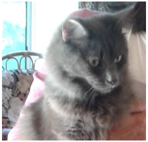
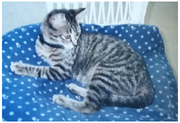
ACTIVITY: Seeing, Observing

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?

Example from Pearl's book

Suppose that we have:

- a cat, Oscar 
- another cat: Bastet 
- a bird feeder outside  that we assume is always populated with birds unless at least one cat is outside
- and a window. 

If the door is open the cats will always go outside. The door is normally closed but if the temperature outside goes above 20C, someone will open it. There are always birds at the feeder unless there is at least one cat outside in which case they will all leave the feeder. We have the following propositions:

- T = The temperature is above 20C.
- D = Someone opens the door.
- O = Oscar goes outside.
- B = Bastet goes outside.
- L = All the birds leave the feeding station.

Some sentences:

1. *Prediction* If Bastet did not go outside then there are birds at the feeding station.

   $\neg B \Rightarrow \neg L.$

2. *Abduction* If there are birds at the feeder then no one opened the door.

   $\neg L \Rightarrow \neg D.$ (Given $D \Rightarrow O \wedge B$, and $B \vee O \Rightarrow L$, then $D \Rightarrow L$, so its contrapositive is true.)

3. *Transduction* If Oscar went outside then so did Bastet.

   $O \Rightarrow B$ (Given $D \Leftrightarrow B$ and $D \Leftrightarrow O$, then $O \Rightarrow D$. So $O \Rightarrow B$.)

4. *Action* If no one opened the door and Bastet snuck outside through a window then all the birds will leave the feeder and Oscar will remain inside.

   $\neg D \Rightarrow L_B \ \& \ \neg O_B$

5. *Counterfactual* If the birds have left the feeder then they still would have left the feeder even if Bastet had not gone outside.

   $L \Rightarrow L_{\neg B}$

Sentences 1 - 3 can be handled by standard logical deduction.

Pearl p.209.:

> *"The feature that renders S1 - S3 manageable in standard logic is that they all deal with epistemic inference – that is, inference from beliefs to beliefs about a static world."*
>
> ⋮
>
> *"From our discussion of actions . . . , any such action must violate some premises, or mechanisms, in the initial theory of the story. To formally identify what remains invariant under the action, we must incorporate causal relationships into the theory; logical relationships alone are not sufficient."*

Equality to show two-way inference

Pearl uses equality rather than implication in order to permit two-way inference. The independent variable is given in brackets in the second column below, to demonstrate the causal asymmetry.

Here is the causal model so far:

Model M

	(T)	
$D = T$	(D)	(Door opens iff temp > 20C.)
$O = D$	(O)	(Oscar goes out iff door opens.)
$B = D$	(B)	(Bastet goes out iff door opens.)
$L = O \vee B$	(L)	(Birds leave iff Oscar or Bastet goes out)

A submodel

To evaluate S4, $(\neg D \Rightarrow L_B \ \& \ \neg O_B)$ we form submodel M_B in which the equation $B = D$ is replaced by B.

If no one opened the door and Bastet snuck outside through a window then all the birds will leave the feeder and Oscar will remain inside.

Model M_B

		(T)
	$D = T$	(D)
	$O = D$	(O)
	B	(B)
	$L = O \vee B$	(L)
Facts: $\neg D$		
Conclusions: $B, L, \neg O, \neg T, \neg D$		

$\neg D \Rightarrow \neg O$ by contrapositive but L is still true since $B \Rightarrow L$.

Pearl's view

Pearl, pp. 209-210:

> *"It is important to note that problematic sentences like S4, whose antecedent violates one of the basic premises in the story,* [*in this case, that Bastet got outside without the door being opened*] *are handled naturally in the same deterministic setting in which the story is told. Traditional logicians and probabilists tend to reject sentences like S4 as contradictory and insist on reformulating the problem probabilistically so as to tolerate exceptions to* [*a*] *law. ...Such reformulations are unnecessary; the structural approach permits us to process commonplace causal statements in their natural deterministic habitat without first immersing them in nondeterministic decor. In this framework, all laws are understood to represent defeasible default expressions subject to breakdown by deliberate intervention."*

Evaluating counterfactuals: step 1

If the birds have left the feeder then they still would have left the feeder even if Bastet had not gone outside.

The counterfactual $L_{\neg B}$ stands for the value of L in submodel $M_{\neg B}$ below. The value L depends on the value of T, which is not specified in $M_{\neg B}$ The observation L removes the ambiguity: if we see the birds have left the feeder, we can infer that the temperature rose above 20C and thus the door was opened. If Bastet had not gone outside then Oscar would have, scaring the birds away from the feeder. We can derive $L_{\neg B}$ as follows.

We add the fact L to the original model and evaluate T. Then we form submodel $M_{\neg B}$ and reevaluate L in $M_{\neg B}$ using the value of T found in the first step.

Model M (Step 1)

	(T)	
$D = T$	(D)	(Door opens iff temp > 20C.)
$O = D$	(B)	(Oscar goes out iff door opens.)
$B = D$	(O)	(Bastet goes out iff door opens.)
$L = O \vee B$	(L)	(Birds leave iff Oscar or Bastet go out)
Facts: L		(Birds leave the feeder.)
Conclusions: T, B, O, D, L		

Evaluating counterfactuals: step 2

Step 2
Model $M_{\neg B}$

	(T)	
$D = T$	(D)	(Door opens iff temp > 20C.)
$O = D$	(O)	(Oscar goes out iff door opens.)
$\neg B$	(B)	(Bastet does not go out.)
$L = O \vee B$	(L)	(Birds leave iff Oscar or Bastet go out)

Facts: T

Conclusions: $T, \neg B, O, D, L$

Pearl remarks that it is only the value of T which he refers to as a 'background variable' that is carried over to Step 2. Everything else must be re-evaluated.

Pearl's next step is to combine steps 1 and 2 into one by using an asterisk to denote variables whose truth value pertains to the hypothetical world created by the modification – in this case $\neg B$. So we rewrite S5 as follows:

Combined theory

$$\begin{array}{lll}
 & & (T) \\
D^* = T & D = T & (D) \\
\neg B^* & B = D & (B) \\
O^* = D^* & O = D & (O) \\
L^* = O^* \vee B^* & L = O \vee B & (L) \\
\hline
\text{Facts: } L & & \\
\hline\hline
\text{Conclusions: } T, B, O, D, L, \neg B^*, O^*, D^*, L^* & &
\end{array}$$

Given L, we have $O \vee B$. Since at least one of O or B is true, we must have D and therefore T, which exists in both worlds. Therefore D^*. Therefore L^*, therefore L^* in spite of $\neg B^*$

Why is S4 ‘action’ and S5 ‘counterfactual’?

- In S4, the fact given (no one opened the door) is not affected by the antecedent (Bastet snuck outside through a window.)
- In S5 we were asking if changing B to $\neg B$ would affect the outcome L vs. $\neg L$. To determine this we had to calculate the potential impact of $\neg B$ on L and route the impact of $\neg B$ through T.

Probabilistic evaluation of counterfactuals

Suppose that ...

1. There is a probability $P(T) = p$ that the temperature goes above 20C.
2. Bastet has a probability q of sneaking out through a window.
3. Bastet's inclination to sneak out a window is independent of T.

We want to compute the probability $P(\neg L_{\neg B}|L)$, the probability that the birds would not have left the feeder if Bastet had not gone outside, given that the birds have in fact left the feeder.

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.
- The only way that the birds would have all left if the temperature had not gone above 20C is that Bastet snuck out a window. (We are assuming that Oscar is not capable of doing so.)

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.
- The only way that the birds would have all left if the temperature had not gone above 20C is that Bastet snuck out a window. (We are assuming that Oscar is not capable of doing so.)
- So the numerator is the probability that Bastet snuck out a window times the probability that the temperature did not rise above 20C (the two events independent) which is $q(1-p)$.

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.
- The only way that the birds would have all left if the temperature had not gone above 20C is that Bastet snuck out a window. (We are assuming that Oscar is not capable of doing so.)
- So the numerator is the probability that Bastet snuck out a window times the probability that the temperature did not rise above 20C (the two events independent) which is $q(1-p)$.
- The denominator is 1 minus the probability that the birds are still there, the latter only being possible if the temperature did not rise above 20C and Bastet did not sneak out a window – i.e. $1-(1-p)(1-q)$

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.
- The only way that the birds would have all left if the temperature had not gone above 20C is that Bastet snuck out a window. (We are assuming that Oscar is not capable of doing so.)
- So the numerator is the probability that Bastet snuck out a window times the probability that the temperature did not rise above 20C (the two events independent) which is $q(1-p)$.
- The denominator is 1 minus the probability that the birds are still there, the latter only being possible if the temperature did not rise above 20C and Bastet did not sneak out a window – i.e. $1-(1-p)(1-q)$
- So $P(\neg L_{\neg B}|L) = \frac{q(1-p)}{1-(1-p)(1-q)}$.

Intuitive calculation

- Intuitively, $\neg L_{\neg B}$ is true, given $\neg B$ iff the temperature did not go above 20C. So we want to compute $P(\neg T|L) = \frac{P(\neg T \wedge L)}{P(L)}$
- This comes to the probability that the birds all left under the circumstances that the temperature did not rise above 20C divided by the probability that the birds all left.
- The only way that the birds would have all left if the temperature had not gone above 20C is that Bastet snuck out a window. (We are assuming that Oscar is not capable of doing so.)
- So the numerator is the probability that Bastet snuck out a window times the probability that the temperature did not rise above 20C (the two events independent) which is $q(1-p)$.
- The denominator is 1 minus the probability that the birds are still there, the latter only being possible if the temperature did not rise above 20C and Bastet did not sneak out a window – i.e. $1-(1-p)(1-q)$
- So $P(\neg L_{\neg B}|L) = \frac{q(1-p)}{1-(1-p)(1-q)}$.

A probabilistic causal model

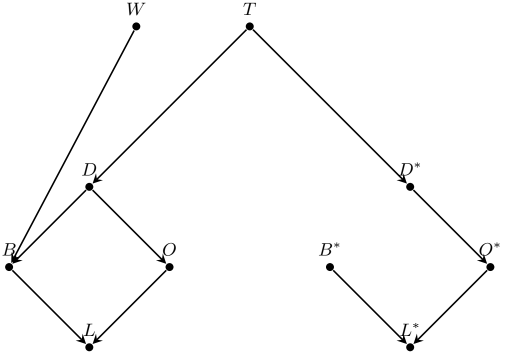
Pearl comments that we can calculate this using a probabilistic causal model using two background variables T (temperature rises above 20C) and W (Bastet decides to go out through a window.)

$$P(t,w) = \begin{cases} pq & \iff t=1, w=1, \\ p(1-q) & \iff t=1, w=0, \\ (1-p)q & \iff t=0, w=1, \\ (1-p)(1-q) & \iff t=0, w=0 \end{cases}$$

We need to first compute the posterior probability $P(t, w|L)$. This can become a problem computationally to compute and store if there are a lot of background variables. And conditioning on some variable e normally destroys the mutual independence of the background variables so that we have to maintain the joint distribution of all the background variables.

Solution: Balke and Pearl 1994: Twin network graphical model

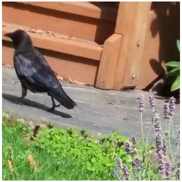
Two networks: one to represent the actual world and one to represent the hypothetical world.



Since we are conditioning on Bastet not going outside, there is no path from D^* to B^*.

A different example

We now look at a different example from Balke and Pearl that illustrates the calculations in more detail.

- There is a crow that sometimes comes to the yard to look for worms but only if it is raining. 
- Bastet goes outside if the crow is out there but otherwise almost never goes outside if it is raining.
- Oscar likes going outside as much as possible even if it is raining but, strangely, is afraid of the crow, so avoids going outside if the crow is there.
- If Bastet and Oscar are both outside, one will likely chase the other away. There is also a slight chance that if both are inside, one will chase the other.

Variables for this example

We have the following variables:

- C The crow is outside or not outside.
- B Bastet is outside or not outside.
- O Oscar is outside or not outside.
- W One of the cats chases the other away.

$$c \in \left\{ \begin{array}{lcl} c_0 & \equiv & \text{The crow is not outside.} \\ c_1 & \equiv & \text{The crow is outside.} \end{array} \right\}$$

$$b \in \left\{ \begin{array}{lcl} b_0 & \equiv & \text{Bastet is not outside.} \\ b_1 & \equiv & \text{Bastet is outside.} \end{array} \right\}$$

$$o \in \left\{ \begin{array}{lcl} o_0 & \equiv & \text{Oscar is not outside.} \\ o_1 & \equiv & \text{Oscar is outside.} \end{array} \right\}$$

$$w \in \left\{ \begin{array}{lcl} w_0 & \equiv & \text{There is no cat chase.} \\ w_1 & \equiv & \text{There is a cat chase.} \end{array} \right\}$$

A conversation by observers

Imagine the following conversation by observers who notice that Bastet is inside even though it is raining.

A: The crow must not be outside, or Bastet would be there instead of inside.

B: That must mean that Oscar is outside!

A: If Bastet were outside, then Bastet and Oscar would surely chase each other.

B: No. If Bastet was there, then Oscar would not be there, because the crow would have been outside.

A: True. But if Bastet were outside even though the crow was not, then Bastet and Oscar would be chasing each other.

B: I agree.

Balke & Pearl p. 232

‘In the fourth sentence, B tries to explain away A’s conclusion by claiming that Bastet’s presence would be evidence that the crow was outside which would imply that Oscar was not outside. B, though, analyzes A’s counterfactual statement as an indicative sentence by imagining that she had observed Bastet’s presence outside; this allows A to use the observation for abductive reasoning. But A’s subjunctive (counterfactual) statement should be interpreted as leaving everything in the past as it was [e.g. that Bastet is inside] (including conclusions obtained from abductive reasoning from real observations [e.g. that the crow must be outside and therefore Oscar must be inside]) while forcing variables to their counterfactual values. This is the gist of A’s last statement.

Unknown factors

Suppose that we have the following probabilities:

$p(b_1|c_1) = 0.9$

$p(b_0|c_0) = 0.9$

We observe that neither Bastet nor the crow is outside and ask whether Bastet would be there if the crow were there: $p(b_1^*|\hat{c}_1^*, c_0, b_0)$. The answer depends on what causes Bastet not to go outside even when the crow is there.

We model the influence of A on B by a function: $b = F_b(a, \epsilon_b)$ where ϵ represents all the unknown factors that could influence B as quantified by the prior distribution $P(\epsilon_b)$. For example, possible components of ϵ_b could be Bastet being sick or Bastet being sulky about never being able to catch the crow.

Response function variables

Each value in ϵ_b's domain specifies a response function that maps each value of A to some value in B's domain.

r_b : domain(ϵ_b) $\to$ N:

$$r_b(\epsilon_b) = \begin{cases} 0 \text{ if } F_b(a_0, \epsilon_b) = 0 \ \& \ F_b(a_1, \epsilon_b) = 0 \ (b = b_0 \text{ regardless of a}) \\ 1 \text{ if } F_b(a_0, \epsilon_b) = 0 \ \& \ F_b(a_1, \epsilon_b) = 1 \ (b = b_1 \Longleftrightarrow a = a_1) \\ 2 \text{ if } F_b(a_0, \epsilon_b) = 1 \ \& \ F_b(a_1, \epsilon_b) = 0 \ (b \text{ has opposite value of } a) \\ 3 \text{ if } F_b(a_0, \epsilon_b) = 1 \ \& \ F_b(a_1, \epsilon_b) = 1 \ (b = b_1 \text{ regardless of a}) \end{cases}$$

r_b is a random variable that can take on as many values as there are functions between a and b. Balke and Pearl call this a *response function variable.*

Response functions for this example

Specifically for this example:

$b = f_b(c, r_b) = h_{b,r_b}(c)$

Whether Bastet goes outside or not is a function of whether the crow is there and of the response function that accounts for other factors that can influence Bastet's behaviour. We can also think of a function h of c that returns a value of b given the value of c and the value of the response variable:

$h_{b,0}(c)$	$=$	b_0	Bastet doesn't go outside regardless of whether the crow is there. e.g. Bastet is ill.
$h_{b,1}(c)$	$=$	$\begin{cases} b_0 \text{ if } c = c_0 \\ b_1 \text{ if } c = c_1 \end{cases}$	Bastet goes outside only if the crow is there.
$h_{b,2}(c)$	$=$	$\begin{cases} b_1 \text{ if } c = c_0 \\ b_0 \text{ if } c = c_1 \end{cases}$	Bastet goes outside only if the crow isn't there.
$h_{b,3}(c)$	$=$	b_1	Bastet goes outside regardless.

An example counterfactual

If we have the prior probability $P(r_b)$ we can calculate $P(b_1^*|\hat{c}_1^*, c_0, b_0)$: i.e. 'Given that the crow is not outside and Bastet is not outside, if the crow were outside, what is the probability that Bastet would be outside?'

We crucially assume that:

> *'...the disturbance ϵ_b, and hence the response-function r_b, is unaffected by the actions that force the counterfactual values; therefore, what we learn about the response-function from the observed evidence is applicable to the evaluation of belief in the counterfactual consequent.'*

If we observe (c_0, b_0) (neither Bastet nor the crow is outside), then it must be that $r_b \in \{0, 1\}$, an event with prior probability $P(r_b = 0) + P(r_b = 1)$. This updates the posterior probability of r_b as follows, letting $\vec{P}(r_b) = \langle P(r_b = 0), P(r_b = 1), P(r_b = 2), P(r_b = 3)\rangle$:

$$\begin{aligned} \vec{P}(r_b) &= \vec{P}(r_b|c_0, b_0) \\ &= \left\langle \frac{P(r_b = 0)}{P(r_b = 0) + P(r_b) = 1}, \frac{P(r_b = 1)}{P(r_b = 0) + P(r_b) = 1}, 0, 0 \right\rangle \end{aligned}$$

Calculating this counterfactual

From the definition of $r_b(\epsilon_b)$ above, if C were forced to c_1 (the crow is outside), then B would have been b_1 (Bastet would have also been outside iff $r_b \in \{1, 3\}$, whose probability is $P'(r_b = 1) + P'(r_b = 3) = P'(r_b = 1)$. ($P'(r_b = 3)$ must be zero since we have determined that $r_b \in \{0, 1\}$.) This gives the solution to the counterfactual query:

$$P(b_1^* | \hat{c}_1^*, c_0, b_0) = P'(r_b = 1) = \frac{P(r_b=1)}{P(r_b=0)+P(r_b=1)}$$

The probability of external influence 1 that causes Bastet to go outside if the crow is there divided by the probability of external influence 1 plus exernal influence 0, the latter causing Bastet to stay inside regardless.

Representation with a DAG

We can represent the causal influences over a set of variables in this example through a DAG. If the set of variables is $\{X_1, X_2, \ldots X_n\}$, for each x_i there is a functional mapping $x_i = f_i(\text{pa}(x_i), \epsilon_i)$, where $\text{pa}(x_i)$ is the value of X_i's parents in the graph and there is a prior probability distribution $P(\epsilon_i)$ for each 'disturbance' ϵ_i.

A counterfactual query will be of the form: 'What is $P(c_* | \hat{a}^*, obs)$, where c^* is a set of counterfactual values for $C \subset X$, $\hat{a}^*$ is a set of forced values in the counterfactual antecedent and *obs* represents observed evidence.

For our example, we assume that Bastet is not outside ($b = b_0$) and want to ask 'what is $P(c_1^* | \hat{b}_1^*, b_0)$?' Or further, what is the probability that the cats will chase each other under those conditions?

A possible causal theory with response variables

Suppose that we have the following of what Balke and Pearl call a 'causal theory':

$$
\begin{array}{llcll}
c = & f_c(r_c) & = & h_{c,r_c}() & \text{(crow's presence depends only on } r_c) \\
b = & f_b(c, r_b) & = & h_{c,r_b}(c) & \text{(Bastet's presence depends on } r_b, \text{ crow)} \\
o = & f_o(c, r_o) & = & h_{c,r_c}(c) & \text{(Oscar's presence depends on } r_o, \text{ crow)} \\
w = & f_w(b, o, r_w) & = & h_{w,r_w}(b, o) & \text{(chase depends on } r_w, \text{ Bastet and Oscar)}
\end{array}
$$

$$
P(r_c) = \begin{cases} 0.40 & \text{if} \quad r_c = 0 \\ 0.60 & \text{if} \quad r_c = 1 \end{cases} \qquad \text{(60\% chance crow is there)}
$$

$$
P(r_b) = \begin{cases} 0.07 & \text{if} \quad r_b = 0 \\ 0.90 & \text{if} \quad r_b = 1 \\ 0.03 & \text{if} \quad r_b = 2 \\ 0 & \text{if} \quad r_b = 3 \end{cases} \qquad \text{(90\% chance Bastet there if crow is)}
$$

$$
P(r_o) = \begin{cases} 0.05 & \text{if} \quad r_o = 0 \\ 0 & \text{if} \quad r_o = 1 \\ 0.85 & \text{if} \quad r_o = 2 \\ 0.10 & \text{if} \quad r_o = 3 \end{cases} \qquad \text{(85\% chance Oscar there if crow isn't)}
$$

$$
P(r_w) = \begin{cases} 0.05 & \text{if} \quad r_w = 0 \\ 0.90 & \text{if} \quad r_w = 8 \\ 0.05 & \text{if} \quad r_w = 9 \\ 0 & \text{otherwise} \end{cases} \qquad \text{(90\% chance chase if B \& O there)}
$$

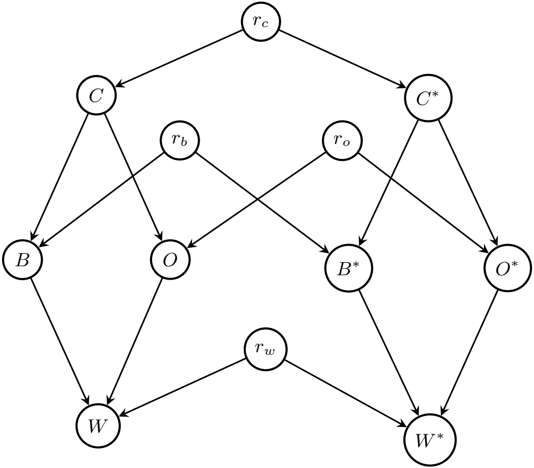
$h_{c,0}() = c_0$ (if $r_c = 0$ the crow is not there)
$h_{c,1}() = c_1$ (if $r_c = 1$ the crow is there)

$h_{w,0}(b,o) = s_0$ (if $r_s = 0$, there is no chase regardless)

$$h_{w,8}(b,o) = \begin{cases} s_0 & \text{if} \quad (b,o) \neq (b_1,o_1) & \text{no chase unless both B,O outside} \\ s_1 & \text{if} \quad (b,o) = (b_1,o_1) & \text{chase if both B,O outside} \end{cases}$$

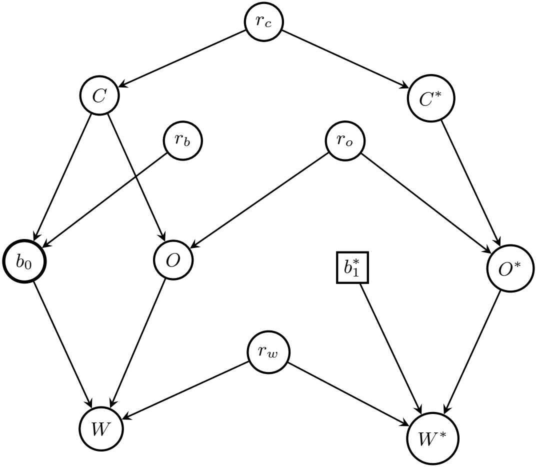
$$h_{w,9}(b,o) = \begin{cases} s_0 \text{ if } (b,o) \in \{(b_1,o_0),(b_0,o_1)\} & \text{no chase if 1 cat present} \\ s_1 \text{ if } (b,o) \in \{(b_0,o_0),(b_1,o_1)\} & \text{chase if B,O meet in or out} \end{cases}$$

DAG



Variables marked with $*$ indicate the counterfactual world and those without the factual world. The r variables are response functions.

DAG for counterfactual evaluation



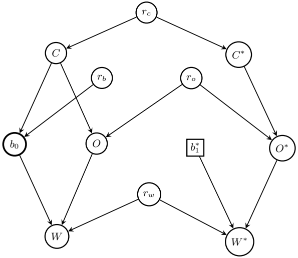
To evaluate $P(w_1^*|\hat{b}_1^*, b_0)$, instantiate B as b_0 and B^* as b_1^*. Sever links pointing to b_1^*

Evaluating W^*

Balke and Pearl comment:

> *If a variable X_j^* in the counterfactual world is not a causal descendant of any of the variables mentioned in the counterfactual antecedent $\hat{a}^*$, then X_j and X_j^* will always have identical distributions, because the causal influences that functionally determine X_j and X_j^* are identical.*

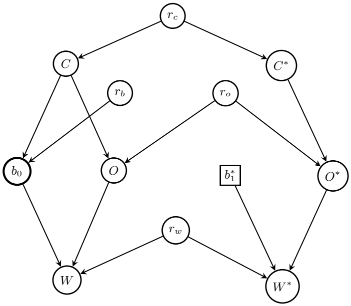
To evaluate W^*, we can start by looking at the graph in the factual world to see what values of parents of b_0 could lead to that value. We consider all the possible combinations of values of parents of b_0. The probability of each combination is the product of their probabilitites and the total prior probability of b_0 is the sum of probabilities of combinations that result in $B = b_0$.



Evaluating W^*

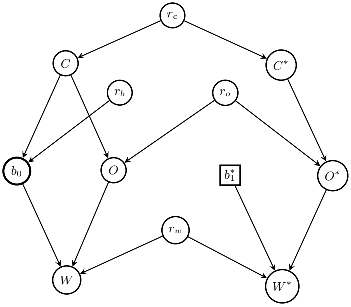
- $r_c = 0$ (0.4) and $r_b = 0$ (0.07) $\rightarrow C = c_0$ and $B = b_0$ (0.028)
- $r_c = 0$ (0.4) and $r_b = 1$ (0.90) $\rightarrow C = c_0$ and $B = b_0$ (0.36)
- $r_c = 0$ (0.4) and $r_b = 2$ (0.03) $\rightarrow C = c_0$ and $B = b_1$ (0.012)
- $r_c = 1$ (0.6) and $r_b = 0$ (0.07) $\rightarrow C = c_1$ and $B = b_0$ (0.042)
- $r_c = 1$ (0.6) and $r_b = 1$ (0.90) $\rightarrow C = c_1$ and $B = b_1$ (0.54)
- $r_c = 1$ (0.6) and $r_b = 2$ (0.03) $\rightarrow C = c_1$ and $B = b_0$ (0.018)

The prior probability $P(B = b_0) = 0.028 + 0.36 + 0.042 + 0.018 = 0.448$. So $p(C = c_0|B = b_0) = \frac{0.028+0.36}{0.448} = 0.86607$



Evaluating W^*

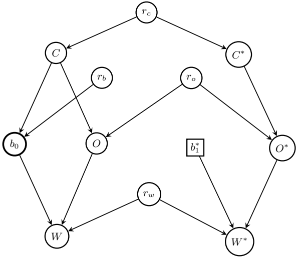
Given that C^* is not a causal descendant of any B variables, we can give counterfactual C^* the same probability as C. We can now work down on the counterfactual side of the DAG and calculate O^*. We calculate the probability of each possible combination of values of r_o and C^* and determine for each the value of O^* that results. We add to get the total probability of o_1^*.



Evaluating W^*

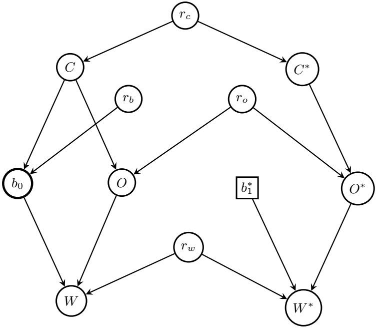
- $r_o = 0$ (0.05) and $C^* = c_0^*$ (0.86607) $\rightarrow O^* = o_0^*$ (0.043304)
- $r_o = 0$ (0.05) and $C^* = c_1^*$ (0.13393) $\rightarrow O^* = o_0^*$ (0.0066965)
- $r_o = 1$ (0) and $C^* = c_0^*$ (0.86607) $\rightarrow O^* = o_0^*$ (0)
- $r_o = 1$ (0) and $C^* = c_1^*$ (0.13393) $\rightarrow O^* = o_1^*$ (0)
- $r_o = 2$ (0.85) and $C^* = c_0^*$ (0.86607) $\rightarrow O^* = o_1^*$ (0.73616)
- $r_o = 2$ (0.85) and $C^* = c_1^*$ (0.13393) $\rightarrow O^* = o_0^*$ (0.1138405)
- $r_o = 3$ (0.10) and $C^* = c_0^*$ (0.86607) $\rightarrow O^* = o_1^*$ (0.086607)
- $r_o = 3$ (0.10) and $C^* = c_1^*$ (0.13393) $\rightarrow O^* = o_1^*$ (0.013393)

$P(O^* = o_1^*) = 0 + 0.73616 + 0.086607 + 0.013393 = 0.83616$



Evaluating W^*

Given that $P(b_0^*) = 1$, we can now calculate $P(W^* = 1 | b_0^*, O^*)$, moving further down the graph. We look at all the possible combinations of possible values of parents of W^*. Since B^* is set at b_1^*, we need not include it in the set of combinations.



Evaluating W^*

- $r_w = 0$ (0.05) and $O^* = o_0^*$ (0.16184) $\rightarrow W^* = w_0^*$ (0.008092)
- $r_w = 0$ (0.05) and $O^* = o_1^*$ (0.83616) $\rightarrow W^* = w_0^*$ (0.041808)
- $r_w = 8$ (0.90) and $O^* = o_0^*$ (0.16184) $\rightarrow W^* = w_0^*$ (0.145656)
- $r_w = 8$ (0.90) and $O^* = o_1^*$ (0.83616) $\rightarrow W^* = w_1^*$ (0.752544)
- $r_w = 9$ (0.05) and $O^* = o_0^*$ (0.16184) $\rightarrow W^* = w_0^*$ (0.008092)
- $r_w = 9$ (0.05) and $O^* = o_1^*$ (0.83616) $\rightarrow W^* = w_1^*$ (0.041808)

So $P(W^* = 1 | b_0^*, O^*) = 0.75254 + 0.041808 = 0.79435$, which to two decimal places is the value given by Balke and Pearl.