The word *cause* is not in the vocabulary of standard probability theory. It is an embarrassing yet inescapable fact that probability theory, the official mathematical language of many empirical sciences, does not permit us to express sentences such as "Mud does not cause rain"; all we can say are that the two events are mutually correlated, or dependent – meaning that if we find one, we can expect to encounter the other. Scientists seeking causal explanations for complex phenomenon or rationales for policy decisions must therefore supplement the language of probability with a vocabulary for causality, one in which the symbolic representation for "Mud does not cause rain" is distinct from the symbolic representation for "Mud is independent of rain". Oddly, such distinctions have yet to be incorporated into standard scientific analysis.

– Judea Pearl, *Causality*, p 134.

Overview

At the end of the class you should be able to:

- explain the predictions of a causal model
- model a user's preferences and utility when there is uncertainty

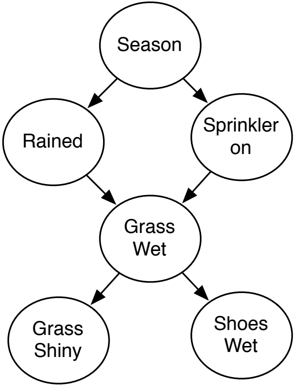
Causality

- An intervention on a variable changes its value by some mechanism outside of the model.
- A causal model is a directed model which predicts the effects of interventions.
- The parents of a node are its direct causes.
- We would expect that a causal model to obey the independence assumption of a belief network.
  - All causal networks are belief networks.
  - Not all belief networks are causal networks.

Sprinkler Example

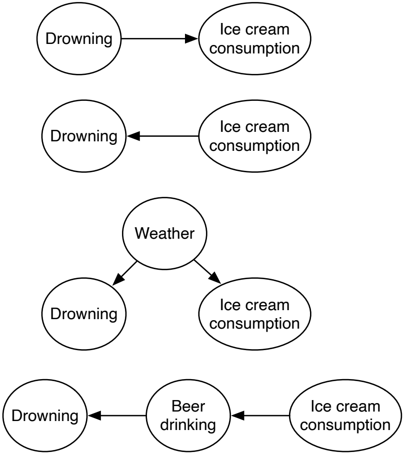
Variables:

- Season dry or wet
- Rained last night
- Sprinkler was on last night
- Grass wet
- Grass shiny and appears to be wet
- Shoes wet after walking on grass



- Which probabilities change if we observe sprinkler on?
- Which probabilities change if we turn the sprinkler on?

Example: drowning and icecream.



- Ice cream consumption and drowning are correlated.
- The top two can be made to fit the data
- Which is a better causal model?
- What experiments could be used to test the models?

Causality

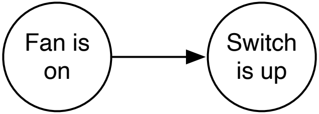
In a causal model:

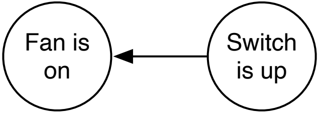
- Intervening on a variable only affects its descendants.
- $P(y \mid do(x))$ is probability of y given intervene on x.
  $P(ShoesWet \mid do(Sprinkler\_on{=}true))$
  $\neq P(ShoesWet \mid Sprinkler\_on{=}true)$
- $do(X)$ be modelled by having an extra parent, "$ForceX$", with $domain(ForceX) = domain(X) \cup \{\perp\}$ and
  if $ForceX = \perp$, the variable X gets it value from other parents,
  otherwise X takes the value from $ForceX$.
- An intervention has a different effect than an observation:
  An observation only affects descendents.

In AISpace try
`https://artint.info/tutorials/sprinklerseason.xml`

Which is a reasonable causal model?

A switch is connected to a fan so they are strongly correlated...

A) 

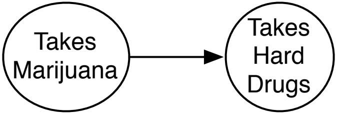
B) 

C) both

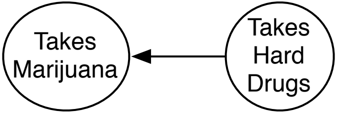
D) neither

Which is a reasonable causal model?

A)



B)



C) both

D) neither

These models cannot be distinguished by observations — but can be distinguished by interventions in controlled studies.
Try:
`https://artint.info/tutorials/causality/marijuana.xml`
in aispace.org Bayes net applet

Clicker Question

Which of the following is not true:

A All belief networks are causal networks

B All causal networks are belief networks

C A causal network predicts the effect of an intervention

D An intervention changes the value of a variable by some mechanism external to the model

E Intervening on a variable only affects the descendents of the variable

Goals and Preferences

Alice . . . went on "Would you please tell me, please, which way I ought to go from here?"
"That depends a good deal on where you want to get to," said the Cat.
"I don't much care where —" said Alice.
"Then it doesn't matter which way you go," said the Cat.

Lewis Carroll, 1832–1898
Alice's Adventures in Wonderland, 1865
Chapter 6

STRIPS Planning (Review)

- deterministic or stochastic dynamics
- fully observable or partially observable
- explicit states or features or individuals and relations
- static or finite stage or indefinite stage or infinite stage
- goals or complex preferences
- perfect rationality or bounded rationality
- flat or modular or hierarchical
- single agent or multiple agents
- knowledge is given or knowledge is learned
- reason offline or reason while interacting with environment

Decision Networks

- deterministic or stochastic dynamics
- fully observable or partially observable
- explicit states or features or individuals and relations
- static or finite stage or indefinite stage or infinite stage
- goals or complex preferences
- perfect rationality or bounded rationality
- flat or modular or hierarchical
- single agent or multiple agents
- knowledge is given or knowledge is learned
- reason offline or reason while interacting with environment

Making Decisions Under Uncertainty

What an agent should do depends on:

- The agent's ability — what options are available to it.
- The agent's beliefs — the ways the world could be, given the agent's knowledge.
  Sensing updates the agent's beliefs.
- The agent's preferences — what the agent wants and tradeoffs when there are risks.

Decision theory specifies how to trade off the desirability and probabilities of the possible outcomes for competing actions.

Preferences

- Actions result in outcomes
- Agents have preferences over outcomes
- A rational agent will do the action that has the best outcome for them
- Sometimes agents don't know the outcomes of the actions, but they still need to compare actions
- Agents have to act.
  (Doing nothing is (often) an action).

Lotteries

- An agent may not know the outcomes of its actions, but only have a probability distribution of the outcomes.
- A lottery is a probability distribution over outcomes. It is written

$$[p_1 : o_1, p_2 : o_2, \ldots, p_k : o_k]$$

where the o_i are outcomes and $p_i \geq 0$ such that

$$\sum_i p_i = 1$$

The lottery specifies that outcome o_i occurs with probability p_i.

Utility for a person

- Utility specifies someone's preferences over possible worlds
- Each possible world has a real-valued utility that measures the desirability of the world
- Suppose the best world "*best*" has a utility of 1, the worst world "*worst*" has a utility of 0
  A world w has utility $u(w)$ means the decision maker:
  - prefers w to the lottery $[p : best, 1 - p : worst]$ for $p < u(w)$
  - prefers the lottery $[p : best, 1 - p : worst]$ to w for $p > u(w)$
  - is indifferent between w and the lottery $[p : best, 1 - p : worst]$ for $p = u(w)$
- Utilities can be any real number: linearly scale to [0,1] and use above definition. Often use $[0, 100]$.
- Different people can have different utilities

Umbrella Utility

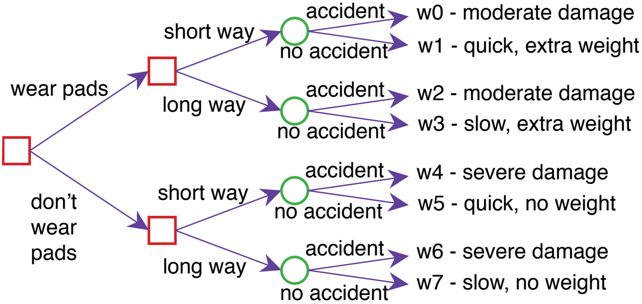
We can take an umbrella or leave it at home.
It could rain or be dry.

	Weather	Umbrella	Utility
A	No Rain	Take it	20
B	No Rain	Leave at home	100
C	Rain	Take it	70
D	Rain	Leave at home	0

Which is the worst outcome?
Which is the best outcome?
How good/bad is it if you take it and it doesn't rain?
How good/bad is it if you take it and it does rain?

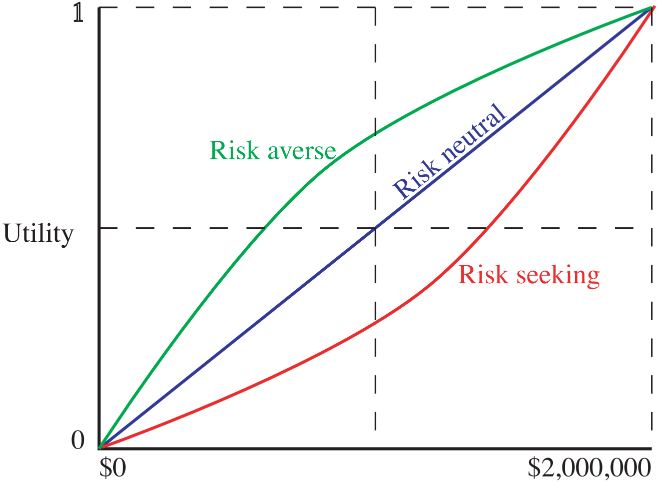
Utility for Delivery Robot

The robot can choose to wear pads to protect itself or not.
The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
There might be an accident.



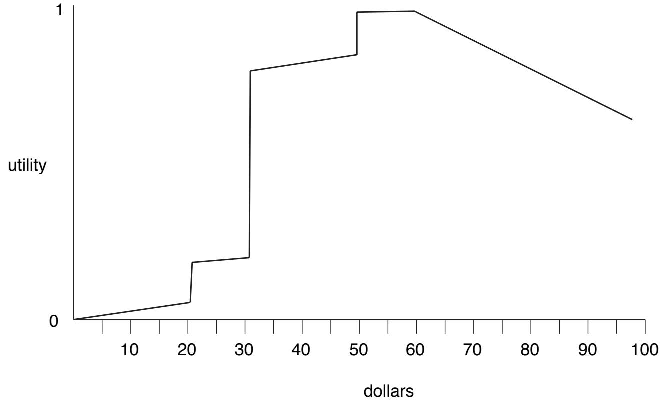
Utility as a function of money

For what p would you be indifferent between \$1m and lottery $[p : \$2m, (1-p) : \$0]$?



Possible utility as a function of money

Someone who really wants a toy worth \$30, but who would also like one worth \$20:



Clicker Question

In a decision for earthquake proofing of schools where

- the worst outcome is 500 die, with a utility of 0
- the best outcome is no money spent and no injuries, with a utility of 100

$1,000,000 spent (on earthquake proofing) has a utility of 80 means:

A If you spent $1,000,000 only 20% of students would die

B You would prefer to spend $1,000,000 and get the advantages of earthquake proofing if and only if there is a greater than 20% probability of the worst outcome.

C You would be prepared to spend $5,000,000 to guarantee to save all of the children

D You are prepared to spend $1,000,000 to save 80% of students

\$1m vs $[p : best, 1 - p : worst]$ Preference flips at $p = 0.8$.

Expected Values

- The expected value of a function of possible worlds is its average value, weighting possible worlds by their probability.
- Suppose $f(\omega)$ is the value of function f on world ω.
  Ω is the set of all worlds.
  - The expected value of f is
    $$\mathcal{E}(f) = \sum_{\omega \in \Omega} P(\omega) \times f(\omega).$$
  - The conditional expected value of f given e is
    $$\mathcal{E}(f|e) = \sum_{\omega \models e} P(\omega|e) \times f(\omega).$$

Decision Variables

- Decision variables are like random variables that an agent gets to choose a value for.
- A possible world specifies a value for each decision variable and each random variable.
- For each assignment of values to all decision variables, the measure of the set of worlds satisfying that assignment sum to 1.
- The probability of a proposition is undefined unless the agent conditions on the values of all decision variables.

Decision Tree for Delivery Robot

The robot can choose to wear pads to protect itself or not.
The robot can choose to go the short way past the stairs or a long way that reduces the chance of an accident.
There is one random variable of whether there is an accident.