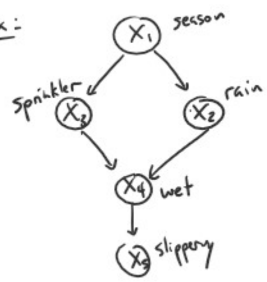
Causality

Monday, November 28, 2022 1:25 PM

# Causal Bayesian Networks

## Bayes Net

Ex:



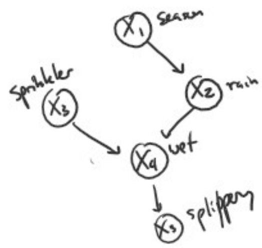

$$p(x_4 \mid x_3 = ON) \propto \sum_{x_1} \sum_{x_2} \sum_{x_5} p(x_1, \dots, x_5)$$

$$= \sum_{x_1} \sum_{x_2} p(x_4 \mid x_3 = ON, x_2)\, p(x_3 = ON \mid x_1)\, p(x_2 \mid x_1)\, p(x_1)$$

depend on the season

## Causal Bayes Net

Ex:
- Suppose we intervene by turning on the sprinkler denoted by, $do(X_3 = ON)$
- Remove dependence of $X_3$ on $X_1$




$$p(x_4 \mid do(X_3 = ON)) \propto \sum_{x_1} \sum_{x_2} \sum_{x_5} p(x_1)\, p(x_2 \mid x_1)\, p(x_4 \mid x_2, do(x_3 = ON))\, p(x_5 \mid x_4)$$

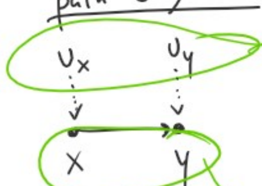
notably absent $p(x_3)$ or $p(x_3 \mid x_1)$

$$= \sum_{x_1} \sum_{x_2} p(x_3 \mid x_2, do(x_3 = ON))\, p(x_2 \mid x_1)\, p(x_1)$$

# Structural Causal Model

## Ex#1: Linear SCM (structural equations model)

path diagram




equations

$$X = U_x$$

$$Y = \beta x + U_y$$

unobserved vars.

observed vars

Legend:
- $X$ = severity of disease (smartphone use)
- $Y$ = severity of symptom (depression)
- $U_x, U_y$ = all other factors (noise/error)
- $\beta \in \mathbb{R}$ = quantifies causal effect of $X$ on $Y$

Model:
- Nature examines values of $X$ and $U_y$ and assigns var $Y$ the value $y = \beta x + U_y$
- Process is <u>not</u> invertible (not true that $x = (y - U_y)/\beta$
- Causal assumptions:
  - ↳ not encoded in links in path diagram

- Causal assumptions:
  - ↳ not encoded in links in path diagram
  - ↳ encoded in the absence of a link

    no direct arrow $\Rightarrow$ claim zero influence

    direct arrow $\Rightarrow$ <u>could</u> be a causal effect

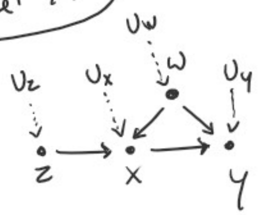
<u>Inference</u>: Given (a) certain conditions (e.g. $cov(U_X, U_Y) = 0$)

(b) trust in causal assumptions of model

one can infer magnitude $\beta$ of causal effects from observed non-experimental data

<u>Ex#2</u>: Nonlinear SCM

Model $M$




$w = f_W(U_W)$

$z = f_Z(U_Z)$

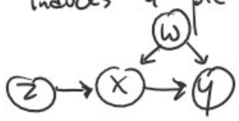
$x = f_X(z, w, U_X)$

$y = f_Y(x, w, U_Y)$

claim of model: changes in $Z$ will leave $Y$ unchanged given $X, W, U_Y$

PGM

- Model $M$ induces a pre-intervention distribution $p_M(w, x, y, z) = p(z)p(w)p(x|z,w)p(y|x,w)$




- We learn $p(v|\text{parents}(v))$ $\forall v \in \{w, x, y, z\}$ from observational data alone

Model $M_X$

<u>Intervention</u>: