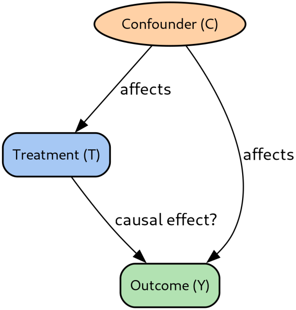
UNIVERSITY OF MARYLAND 18 56 SCIENCE ACADEMY

Estimating Causal Effects

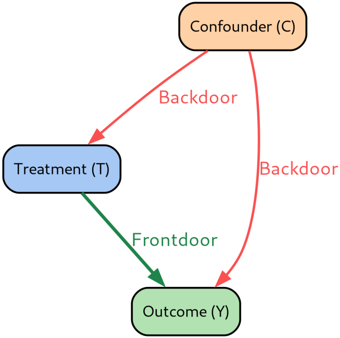
- **Goal**: Determine the causal effect of a treatment (aka intervention) variable T on an outcome Y
- **Example:**
  - $T =$ "takes drug"
  - $Y =$ "recovers"
  - $C =$ "overall health"
  - Healthier people may both take medicine and recover faster $\implies$ correlation without causation
- In **observational data**
  - Confounding variable C affects both treatment T and outcome Y
  - C creates *spurious correlation* between T and Y



- **Problem**
  - There is a "backdoor path" from *Treatment* $\rightarrow$ *Confounder* $\rightarrow$ *Outcome*

Frontdoor and Backdoor Paths: Intuition

- **Backdoor path**
  - A backdoor path is any path from T to Y starting with an arrow into T
    - E.g., backdoor path: $T \leftarrow C \rightarrow Y$
  - Interpretation:
    - C is a common cause of both T and Y and it confounds the relationship between T and Y
    - Controlling for (condition) C blocks the backdoor path, allowing identification of the causal effect of T on Y



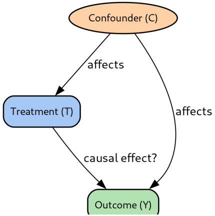
- **Frontdoor path**
  - A frontdoor path goes directly or indirectly from T to Y through mediators, following the direction of causal flow
    - E.g., front-door path: $T \rightarrow Y$
  - Interpretation:
    - This is the direct causal path of interest
    - No mediators here, so the front-door path is the direct causal effect of T on Y

Randomized Controlled Trials (RCTs)

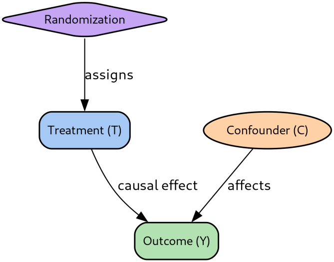
- **Randomized Controlled Trial (RCT)** is an experimental study to assess the causal effect of an intervention or treatment
  - Determine whether an intervention causes an effect, not just associated with it
  - Eliminate selection bias and confounding variables through randomization
- **Key Components**
  - *Randomization:* ensures groups are statistically equivalent at baseline
  - *Control Group:* receives a placebo or standard treatment
  - *Blinding:* participants and/or researchers do not know the assignment to avoid bias
  - *Outcome Measurement:* pre-defined metrics assess the intervention's effect
- **Example**: testing a new drug
  - Treatment group receives the new drug
  - Control group receives a placebo
  - Compare recovery rates after a fixed period
- **Pros**
  - Provides clear causal inference due to randomization
- **Cons**
  - Expensive and time-consuming
  - Ethical or practical constraints may prevent randomization

RCTs Solve the Problem of Confounders

- In **observational data**
  - Confounding variable C affects both treatment T and outcome Y
  - C creates *spurious correlation* between T and Y



- In **experimental settings**
  - Randomization (R) breaks the link between C and T
  - Everyone is assigned randomly, so nothing can influence both treatment and outcome
  - Now T is independent of C: $T \perp C$
  - The only open path between T and Y is the causal path $T \to Y$



Causal Graphs and Interventions

- Simply **observing correlations** between variables *does not reveal causality*
  - $\Pr(Y|T)$ confounds direct and indirect influences
- **Randomized Controlled Trials (RCTs)** provide the *gold standard* for causal inference
  - Randomization breaks all back-door (confounding) paths
  - But RCTs are expensive, slow, or ethically impossible
- **Alternative solution**
  - Can we estimate the *causal effect* from *observational data alone*?
  - Under *what conditions* and using *which variables*?
- **Idea**: If we identify and condition on the right *confounders*, you can
  - Block spurious associations between T and Y
  - Recover the true causal effect $P(Y|do(T))$

Intervention in Structural Equations

- **Purpose of Structural Equations**
  - Capture causal mechanisms among variables
  - Go beyond correlations: predict the impact of external interventions
- **Effect of Intervention** $do(X_j = x_j)$
  - Original equation: $X_j = f_j(Parents(X_j), U_j)$
  - Modified by intervention: $X_j = x_j$ (fixed value)
  - "Mutilate" the causal network by *removing incoming edges* to X_j
  - Use the modified structure to recompute the joint distribution of all variables
- **Intuition**
  - The *do*-operator enforces a variable's value externally, breaking causal dependencies
  - Enables reasoning about "what would happen if. . . ?" scenarios

SCIENCE ACADEMY

Adjustment Formula in Causal Networks

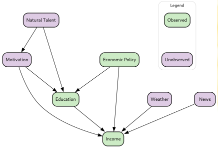
- **Goal**
  - Estimate causal effect of intervention $do(X_j = x_{jk})$ on another variable X_i
- **The Adjustment Formula**
  - Derived from the post-intervention joint distribution:

$$\Pr(X_i = x_i | do(X_j = x_{jk})) = \sum_{Parents(X_j)} \Pr(x_i | x_{jk}, Parents(X_j)) \Pr(Parents(X_j))$$

  - The mechanism for X_j is *removed*: it is treated as a fixed cause, not a random variable
- **Interpretation**
  - Computes a **weighted average** of effects of X_j and its parents on X_i
  - Weights come from prior probabilities of the parents' values
- **Back-Door Criterion**
  - A set Z is a valid adjustment set if it blocks *all back-door paths* from X_j to X_i
  - Ensures $X_i \perp\!\!\!\perp \text{Parents}(X_j) \mid X_j, Z$
- **Why It Matters**
  - Enables causal inference from observational data
  - Forms the basis for estimating treatment and policy effects *without randomized trials*

Intervention: Sprinkler Example

- "Intervene" by turning the sprinkler on, i.e., in do-calculus $do(Sprinkler = T)$
  - Now the sprinkler variable s is not dependent on whether it's a cloudy day c
- The structural equations after the intervention become:



$$
\begin{aligned}
C &= f_C(U_C) \\
R &= f_R(C, U_R) \\
S &= True \\
W &= f_W(R, S, U_W) \\
G &= f_G(W, U_G)
\end{aligned}
$$

- Then the $\Pr(s|c) = 1$ and $\Pr(w|r, s) = \Pr(w|r, s = T)$ and the joint probability becomes:

$$\Pr(c, r, w, g | do(S = True)) = \Pr(c) \Pr(r|c) \Pr(w|r, s = True) \Pr(g|w)$$

- The variables that are affected are only the descendants of the manipulated variable *Sprinkler*

Intervention vs Observation

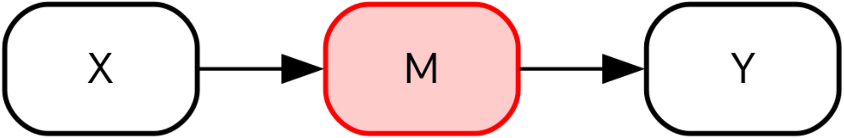
- **Intervention** "breaks" normal causal link between *Weather* and *Sprinkler*
  - Causal graph shows no influence of *Sprinkler* on *Weather*
- Difference between $do(Sprinkler = T)$ (intervention) and $Sprinkler = T$ (observation)
  - Observing sprinkler on:
    - Less likely weather is cloudy
  - Turning on sprinkler:
    - Weather unaffected, probability of cloudy unchanged

Back Door

- A causal network can predict the effect of an intervention using the adjustment formula
- The problem is that it requires accurate knowledge of the conditional distributions of the model
- E.g., in the Sprinkler example, why does someone turn on the sprinkler? Maybe they check the weather, but how do they make their decision?
- Besides the direct route, we needs to take in account the "back door" route

TBD

- We want to estimate the effect of *Sprinkler* on *GreenerGrass*
  - We intervene and set $Sprinkler = True$
- Besides the direct route, we needs to take in account the “back door” route $Cloudy \rightarrow Rain$



- If we knew the value of *Rain*, this back door path would be blocked and we could condition on *Rain* instead of *Cloudy*
- In formal way we need to find a set of Z variables such that X_i is conditionally independent of $Parents(X_j)$ given X_j and Z
  - TODO: ?

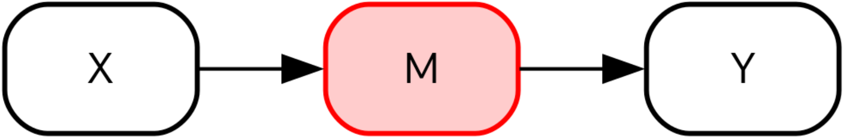
SCIENCE ACADEMY

Backdoor Path

- Z blocks all backdoor paths from X to Y
- A backdoor path is any path from X to Y that starts with an arrow pointing into X. These paths create confounding relationships that can bias the estimate of X's effect on Y

Backdoor Criteria: Condition

- Variables Z satisfy the backdoor criterion for X and Y in a causal network if:
  1. No element of Z is a descendant of X
     - Z do not capture the effect of X on Y through any causal pathway
  2. Z "blocks" every path from X to Y
- The idea is to estimate the causal effect of X on Y without confounding relationships, by controlling variables Z satisfying the backdoor criterion
- Then we can use non-experimental (i.e., observational data), assigning X randomly



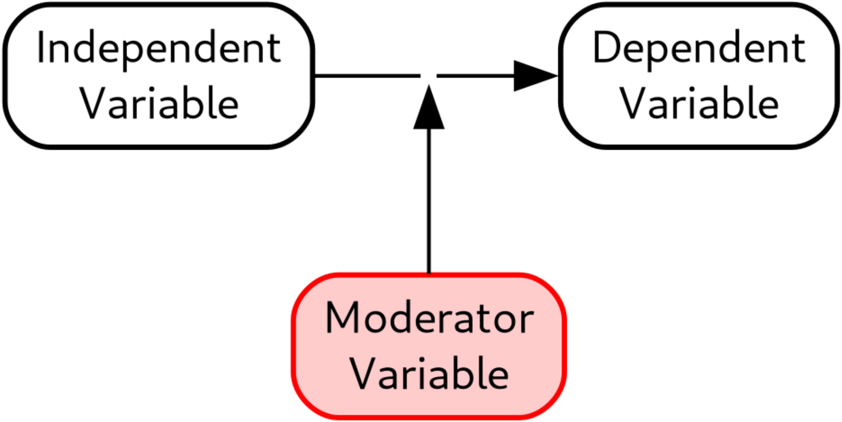
- ***Type of Variables in Causal AI***

Mediator Variable

- A **mediator variable** M lies on the causal path between a treatment X and an outcome Y

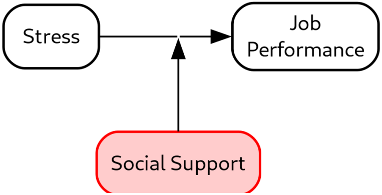
Mediator Variable: Example

- Does a Training Program increase Employee Productivity?
- Training Program may not directly increase productivity
  - It might increase job satisfaction (mediator), leading to greater productivity



Moderator variable

- **Moderator variable** affects the strength or direction between two other variables



- **E.g.,**
  - Study the relationship between stress X and job performance Y
  - The level of social support an individual receives M could be a moderator
    - If social support is high, the negative effect of stress on job performance might be weaker
    - If social support is low, the negative effect might be stronger