CS 486/686 Independence and Bayesian Networks II

Yuntian Deng

Lecture 8

RN 13.2 · PM 8.3

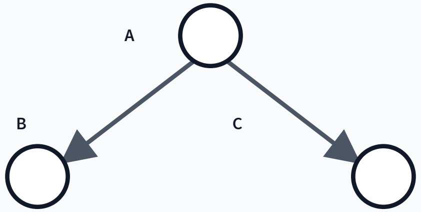
Learning goals

- Apply **d-separation** to check (conditional) independence in any Bayes net
- Given a BN and a variable ordering, **construct a correct Bayes net**
- Distinguish **correlation** from **causation**

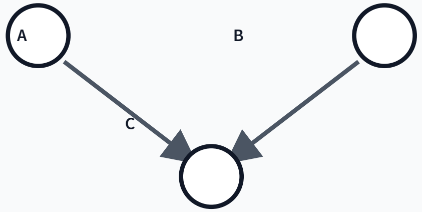
Recap: the three structures from L7

Chain

Common cause



V-structure



Today: generalize the 3 structures to **arbitrary paths** using **d-separation**.

D-separation: the rule

D-separation. A set of observed variables E *d-separates* X and Y iff E **blocks every undirected path** between X and Y.

If E d-separates X and Y, then X and Y are **conditionally independent given** E.

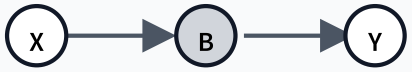
A single open path is enough to break independence. We have to check every path and ask: is it blocked?

The three blocking rules

A path through middle node B is **blocked** iff:

Chain

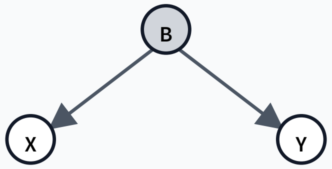
$$X \to B \to Y$$



Blocked when $B \in E$ (middle is observed)

Common cause

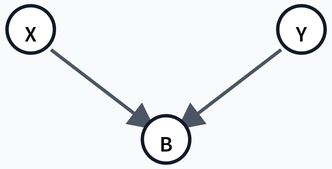
$$X \leftarrow B \to Y$$



Blocked when $B \in E$ (middle is observed)

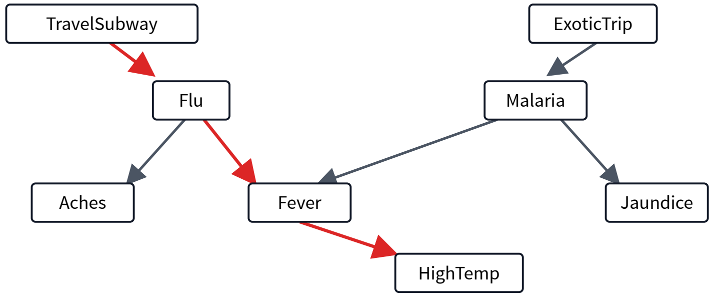
V-structure

$$X \to B \leftarrow Y$$



Reversed: blocked when neither B nor any of its descendants are in E

Quiz pack 1: chain blocking



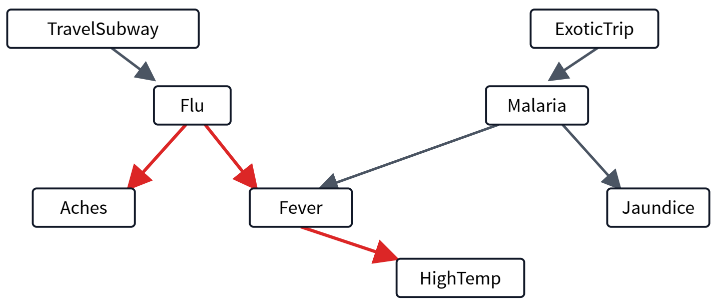
Q1. Are **TravelSubway** and **HighTemp** independent?

> **No.** The path Subway $\to$ Flu $\to$ Fever $\to$ HighTemp has two chain middles (Flu, Fever); both unobserved $\Rightarrow$ path open.

Q2. Are **TravelSubway** and **HighTemp** independent *given Flu*?

> **Yes.** Observing Flu blocks the chain at the first middle $\Rightarrow$ path closed.

Quiz pack 2: another chain test



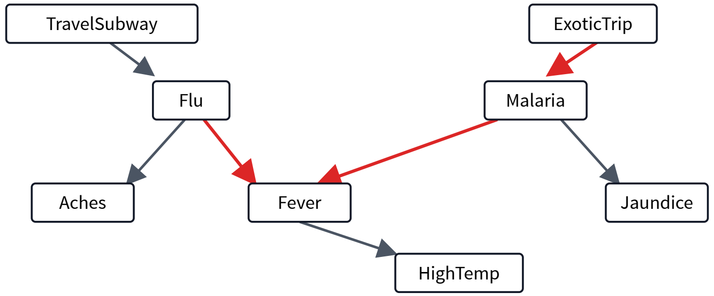
Q3. Are **Aches** and **HighTemp** independent?

> **No.** Path Aches $\leftarrow$ Flu $\rightarrow$ Fever $\rightarrow$ HighTemp: common cause (Flu, unobserved) and chain (Fever, unobserved). Both open $\Rightarrow$ path open.

Q4. Are **Aches** and **HighTemp** independent *given Flu*?

> **Yes.** Observing Flu (the common cause) blocks the path at the first middle.

Quiz pack 3: v-structure with a descendant



Q5. Are **Flu** and **ExoticTrip** independent?

> **Yes.** Path Flu $\longrightarrow$ Fever $\longleftarrow$ Malaria $\longleftarrow$ ExoticTrip. Fever is a v-structure middle; neither Fever nor its descendant HighTemp is observed $\Rightarrow$ v-structure blocks the path.

Q6. Independent *given HighTemp*?

> **No.** HighTemp is a descendant of the v-structure middle Fever, so the v-structure now *opens* the path (explaining-away).

Many correct Bayes nets exist

A BN is **correct** if every independence it claims also holds in the true distribution.

Missing an independence: OK

extra edges — still encodes the joint

Missing a dependence: NOT OK

too sparse — forces a wrong independence

So prefer the BN with the **fewest edges** (fewest probabilities to store).

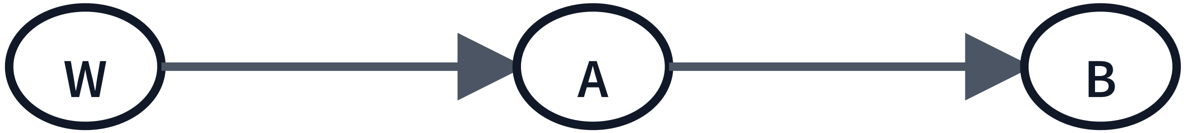
Construction algorithm

1. Order the variables $X_1, \ldots, X_n$.
2. For each X_i, choose the **smallest** subset of $\{X_1, \ldots, X_{i-1}\}$ such that, given those parents, X_i is independent of the rest.
3. Add edges from each parent to X_i; write down $P(X_i \mid \mathrm{Parents}(X_i))$.

Ordering matters: a bad order can force every later variable to depend on many earlier ones.

Example 1: chain BN, order W, A, B

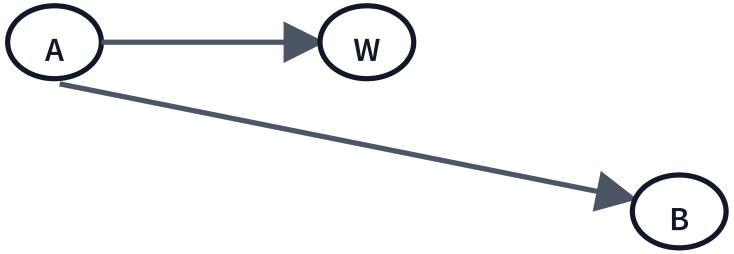
Original BN: 



Final: $W \to A \to B$. Same shape as the original (2 edges).

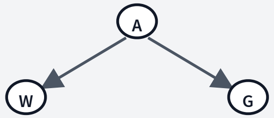
Example 1 alt: same BN, order A, W, B

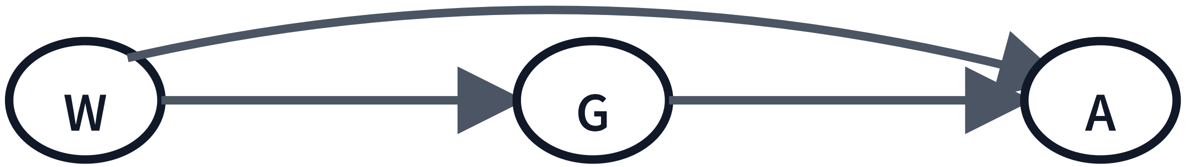
Original BN: 



Final: $A \to W$ and $A \to B$. Different shape, still 2 edges — *different but equally compact.*

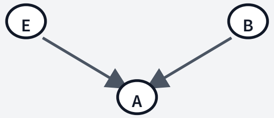
Example 2: common cause, order W, G, A

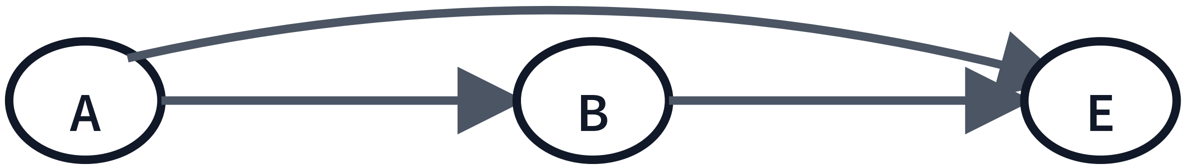
Original BN: 



Final: 3 edges ($W \to G, W \to A, G \to A$) — *more than the original's 2 edges.* Suboptimal!

Example 3: v-structure, order A, B, E

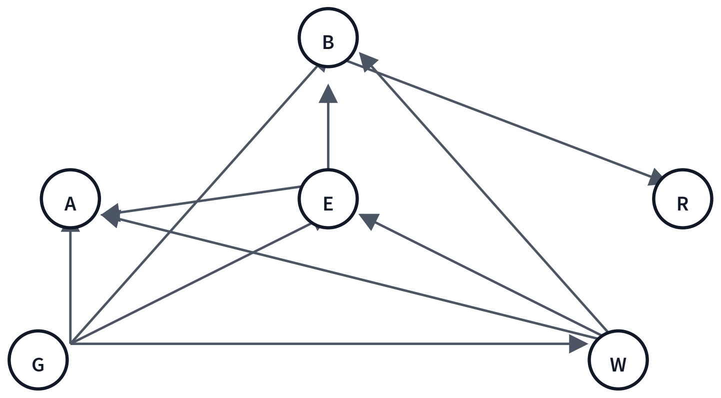
Original BN: 



Final: 3 edges ($A \to B, A \to E, B \to E$) — *more than the original's 2*. Reversed v-structures are expensive.

Holmes with a bad order: G, W, E, B, A, R

Adding effects before causes forces every later node to depend on every earlier one.



$1 + 2 + 4 + 8 + 16 + 2 = \mathbf{33}$ probabilities — vs **12** with the causal order.

Pick a causal order

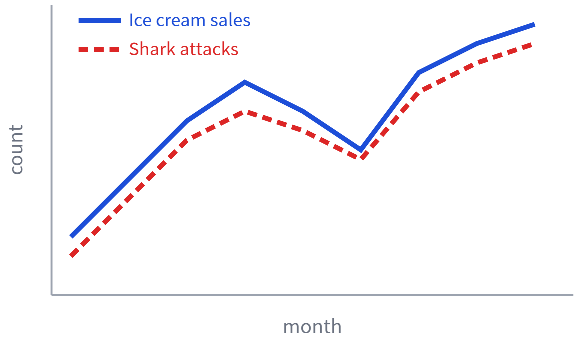
Causes precede effects. Add root causes first; effects last.

Example	Original	Order	Reconstructed	Result
Ex 1	$B \to A \to W$	W, A, B	$W \to A \to B$	2 edges, same shape
Ex 1 alt	$B \to A \to W$	A, W, B	$A \to W, A \to B$	2 edges, different shape
Ex 2	$A \to W, A \to G$	W, G, A	$W \to G, W \to A, G \to A$	**3 edges — worse**
Ex 3	$E \to A, B \to A$	A, B, E	$A \to B, A \to E, B \to E$	**3 edges — worse**

Finding the most compact BN is NP-hard in general — but a causal ordering is a good heuristic.

Correlation ≠ causation

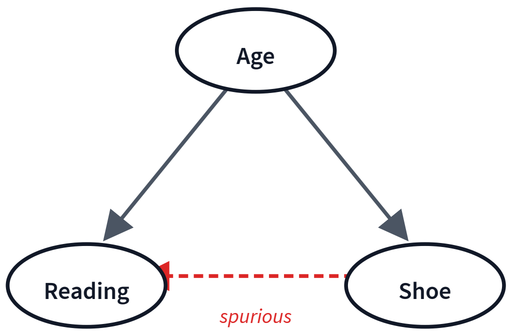
Two variables can be highly correlated without one causing the other.



both peak in summer (hidden cause)

- **Ice cream sales** correlate with **shark attacks**. Does ice cream attract sharks?
- No — both rise in **summer**. Temperature is a hidden common cause.
- An edge in a Bayes net is associational, *not necessarily causal*.

Confounding variables

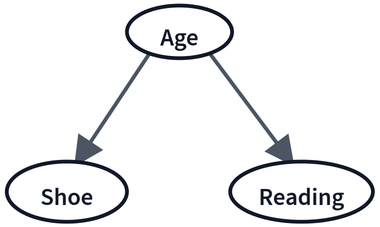


- Bigger shoes correlate with better reading scores.
- **Hidden cause:** Age. Older children have bigger feet *and* read better.
- The Shoe $\longrightarrow$ Reading "effect" is **spurious**; controlling for Age would erase it.

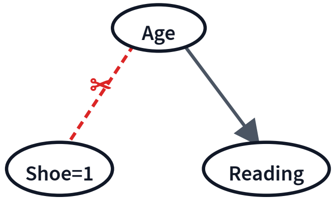
Causal intervention: the do operator

- **Condition** $P(R \mid S{=}1)$: what we *see* — large feet correlate with reading.
- **Intervene** $P(R \mid \mathrm{do}(S{=}1))$: what happens if we *force* big shoes on a random child.

condition on S



$\mathrm{do}(S = 1)$



Intervention *severs* the incoming edges of S — the confounder can no longer reach S.

Adjusting for confounders

The **average treatment effect** averages the effect over the confounder A:

$$\mathrm{ATE} = \sum_A P(R \mid S{=}1, A)\, P(A) - \sum_A P(R \mid S{=}0, A)\, P(A)$$

$$\mathrm{ATE} \approx 0$$

Shoe size doesn't *cause* reading skill — randomised experiments confirm it.

Learning goals (recap)

- ✓ Apply d-separation to check (conditional) independence
- ✓ Construct a correct Bayes net from a variable ordering
- ✓ Distinguish correlation from causation

Next: decision theory

We can now reason under uncertainty. Next, we combine these probabilities with **utilities** to *act* — choose the action that maximizes expected reward.