# Independence and Bayesian Networks (Part 2)

Yuntian Deng

Lecture 8

Readings: RN 13.2. PM 8.3.

# Outline

# Learning Goals

- Determine whether an independence relationship holds by applying d-separation.
- Given a Bayesian network and an order of the variables, construct a Bayesian network that correctly represents the independence relationships among the variables.
- Understand the difference between correlation and causality.

# D-Separation

Are two variables $X$ and $Y$ independent
given the set of observed variables $E$?

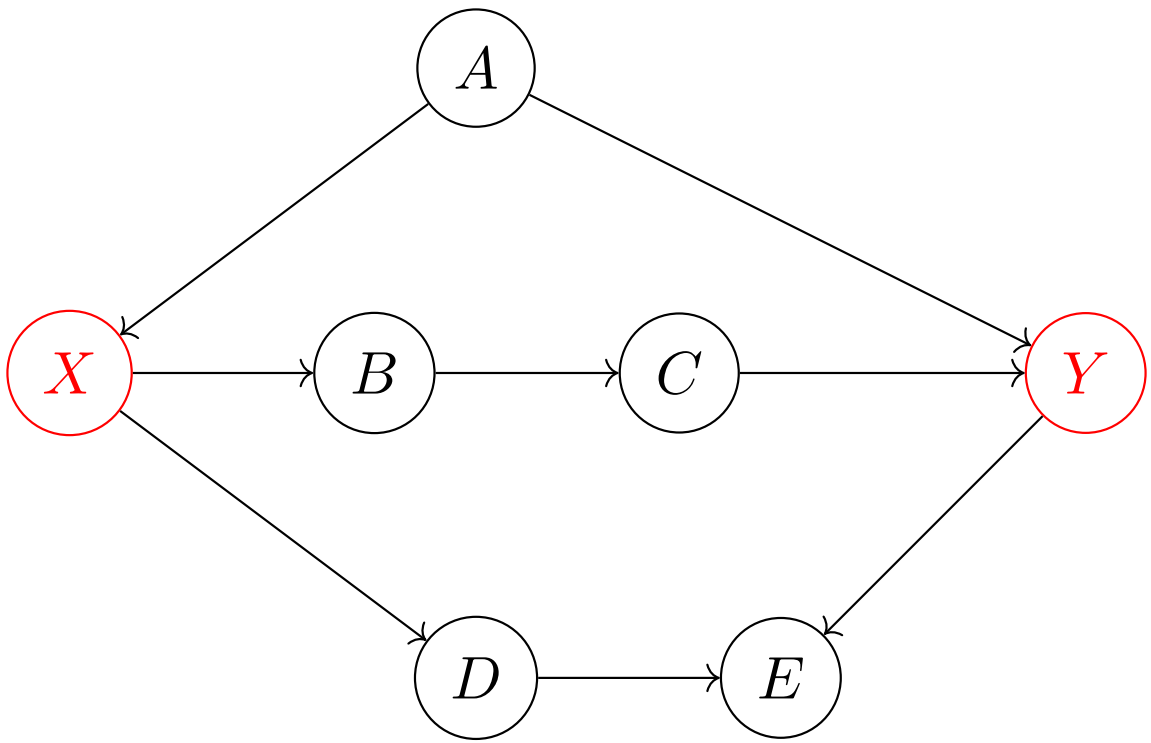
## Definition (D-Separation)

$E$ d-separates $X$ and $Y$
iff $E$ blocks every un-directed path between $X$ and $Y$.

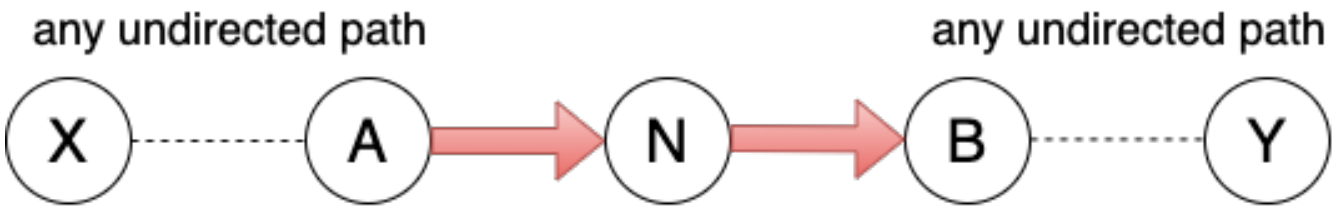
If $E$ d-separates $X$ and $Y$,
then $X$ and $Y$ are conditionally independent given $E$.

# D-Separation

- Un-directed paths between $X$ and $Y$.
- Multiple paths need to be considered if they exist.
- One of the nodes on all the paths blocking the connection.



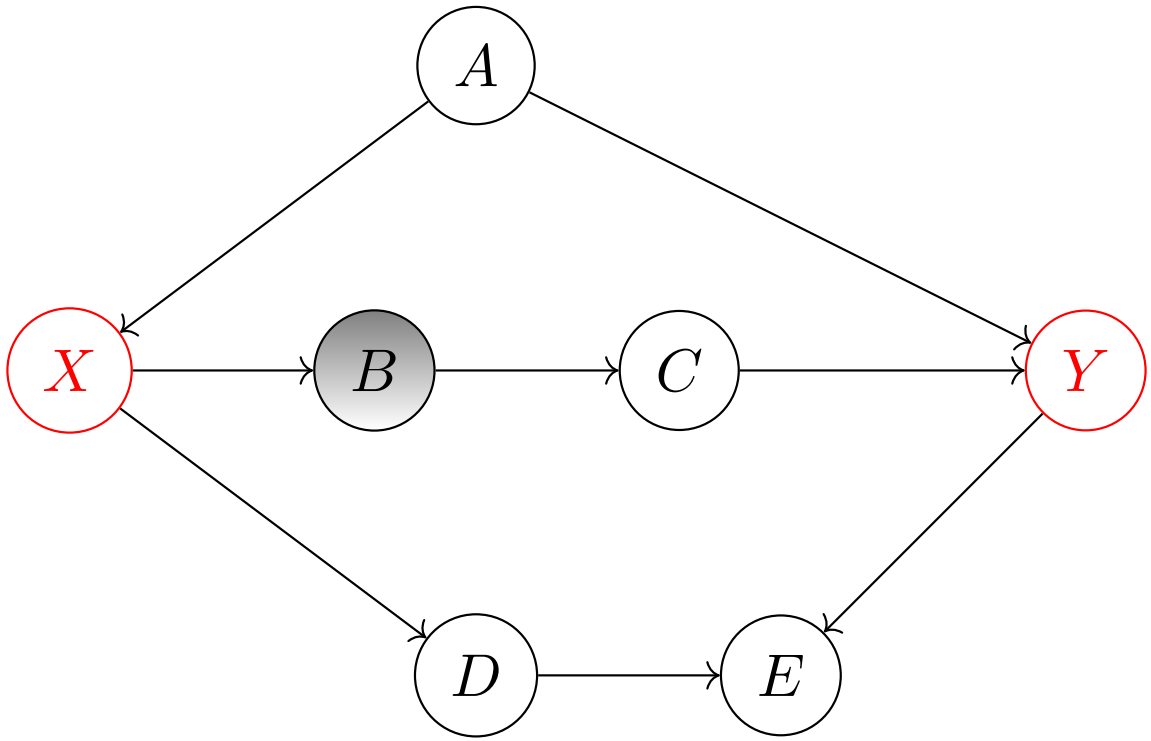
# Blocked Path - Scenario 1/3



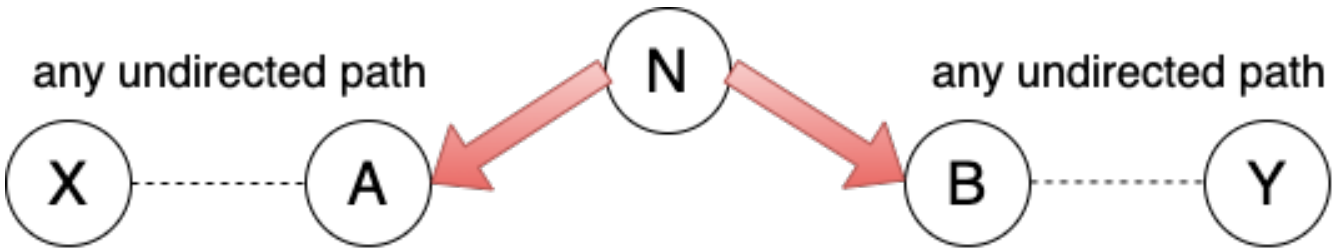

If N is observed,
then it blocks the path between X and Y.

# D-Separation (Scenario 1/3)

- $B$ blocks the path between $X$ and $Y$, which follows scenario 1.



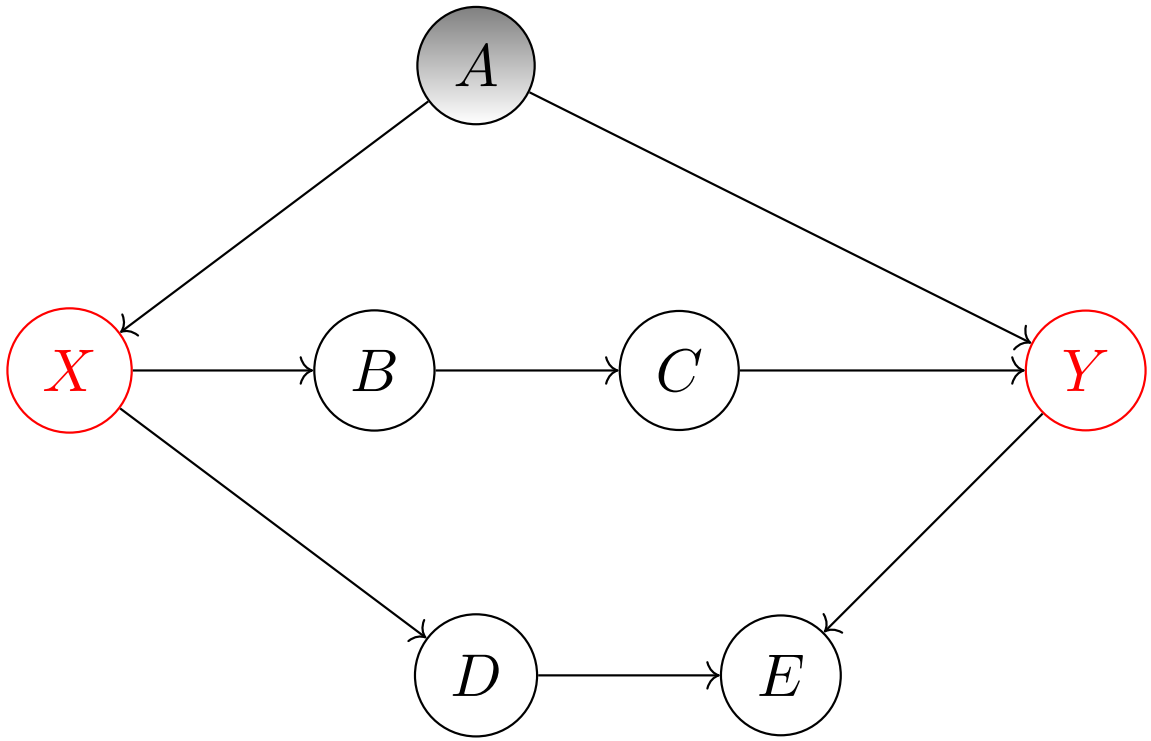
# Blocked Path - Scenario 2/3



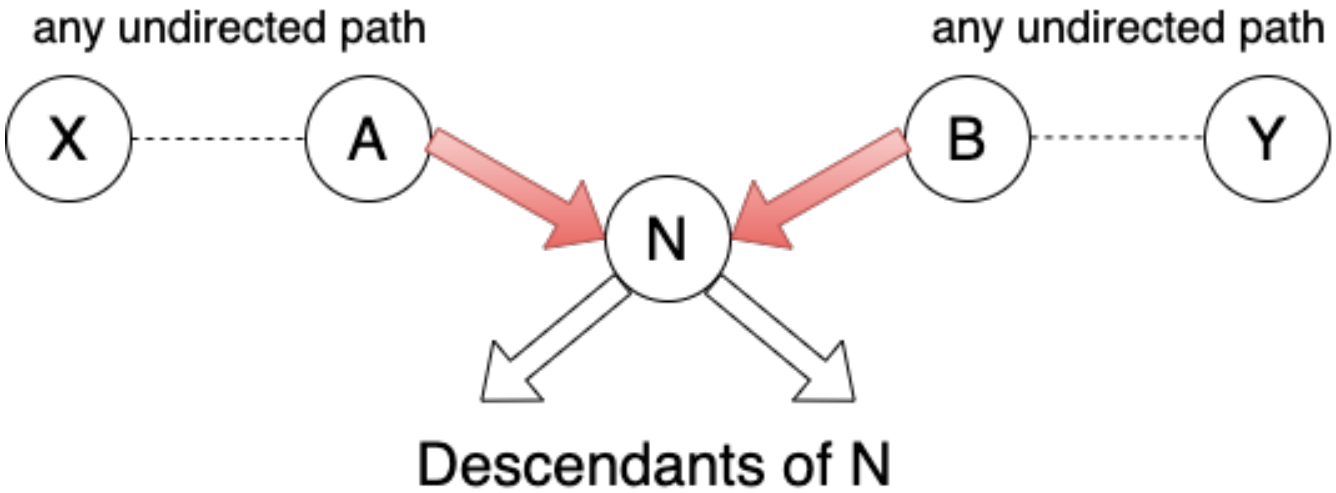

If N is observed,
then it blocks the path between X and Y.

# D-Separation (Rule 2)

- $A$ blocks the path between $X$ and $Y$, which follows scenario 2.



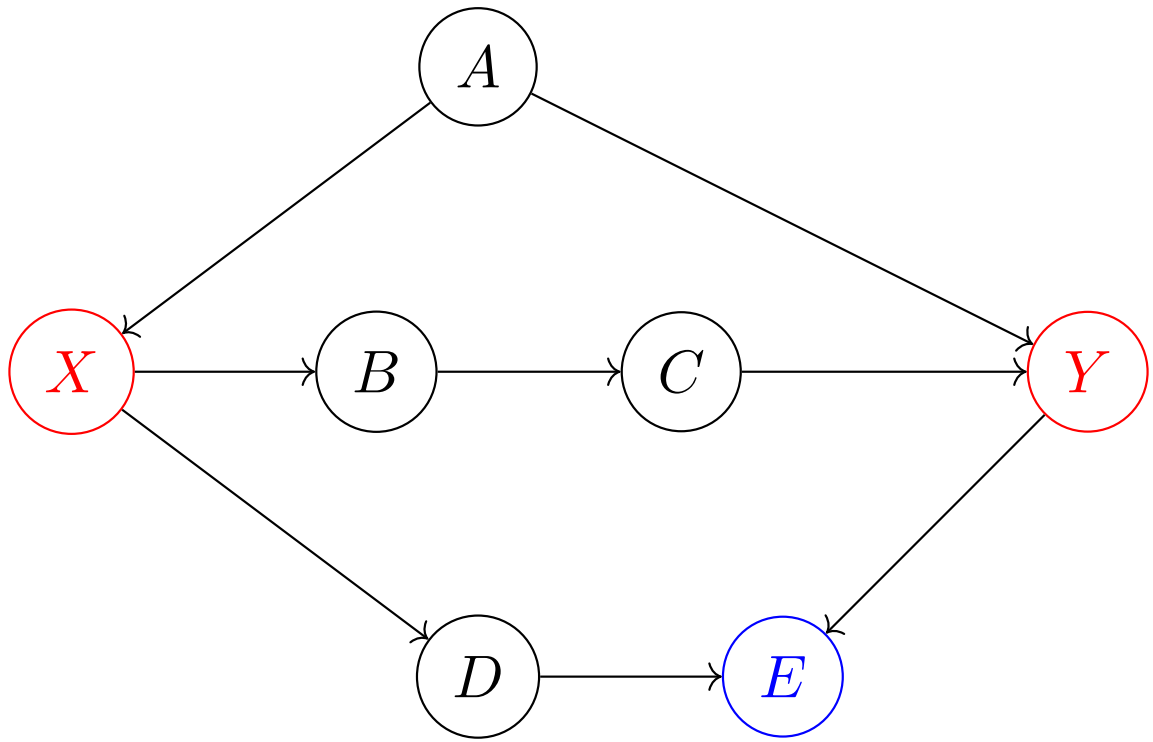
# Blocked Path - Scenario 3/3




If N and N's descendants are NOT observed,
then they block the path between X and Y.

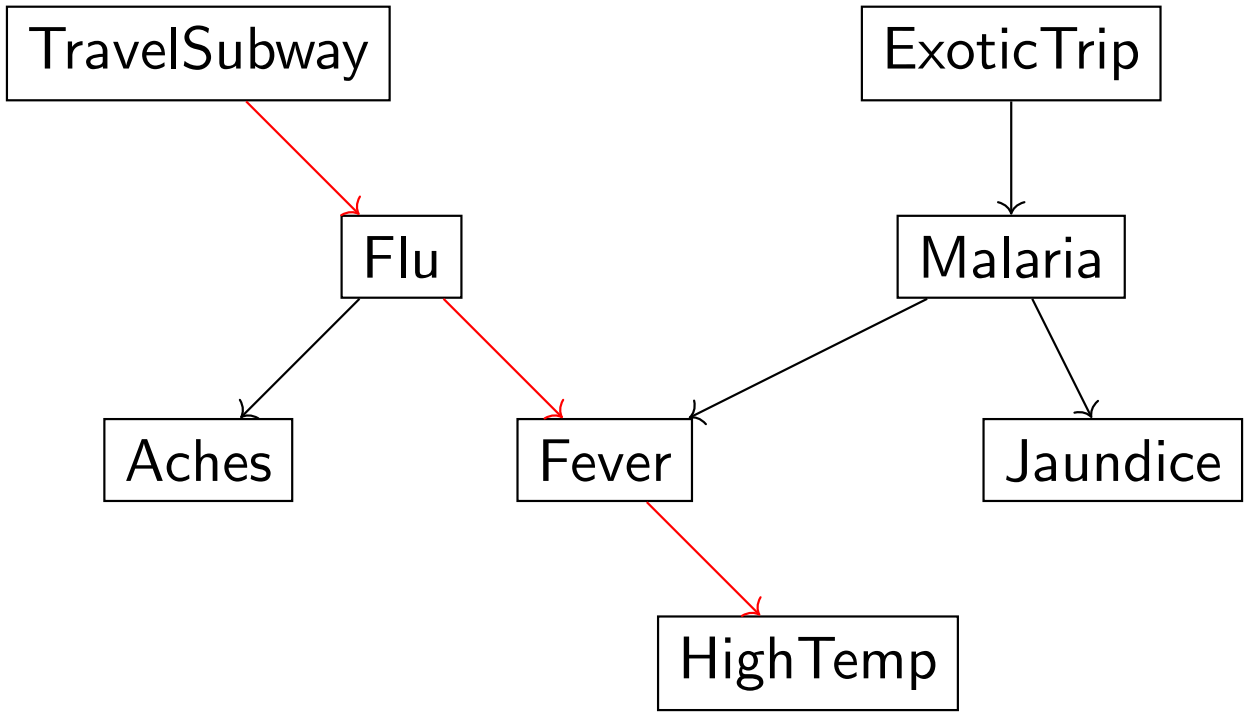
# D-Separation (Rule 3)

- $E$ blocks the path between $X$ and $Y$, which follows scenario 3.
- If $E$ is not observed, then $X$ and $Y$ are d-separated.



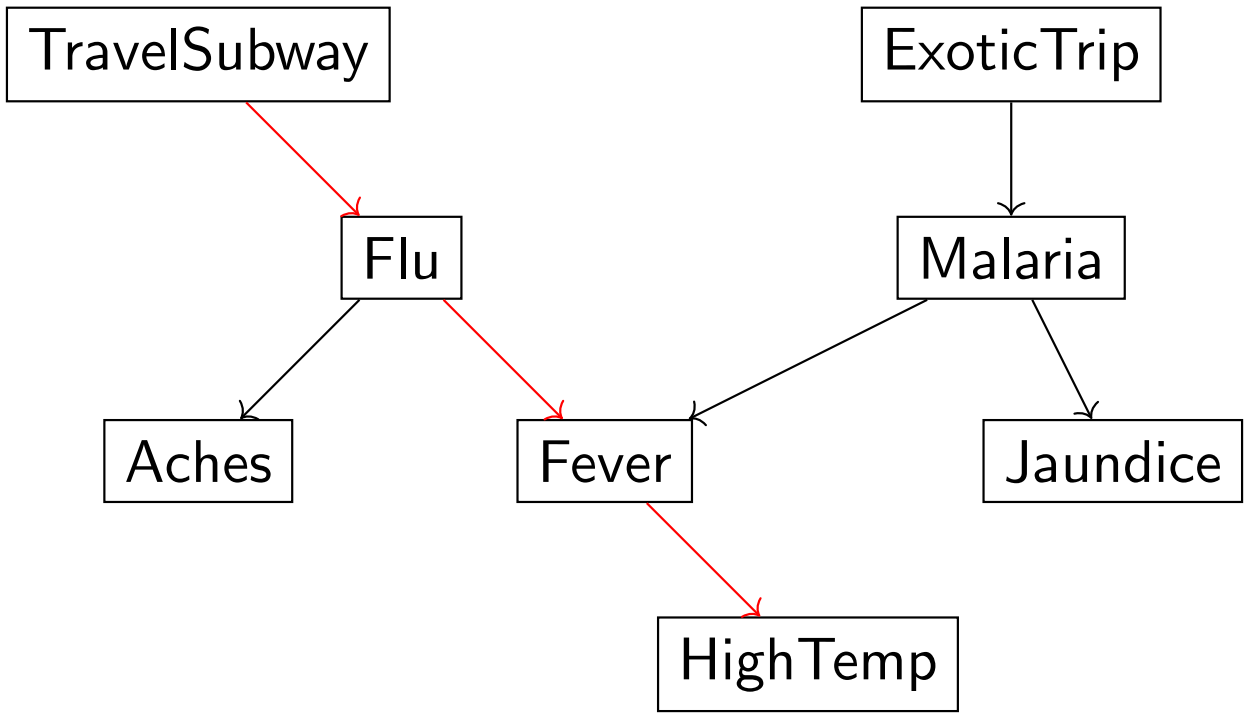
# Q: Applying D-separation

**Q #1:** Are **TravelSubway** and **HighTemp** independent?



# Q: Applying D-separation

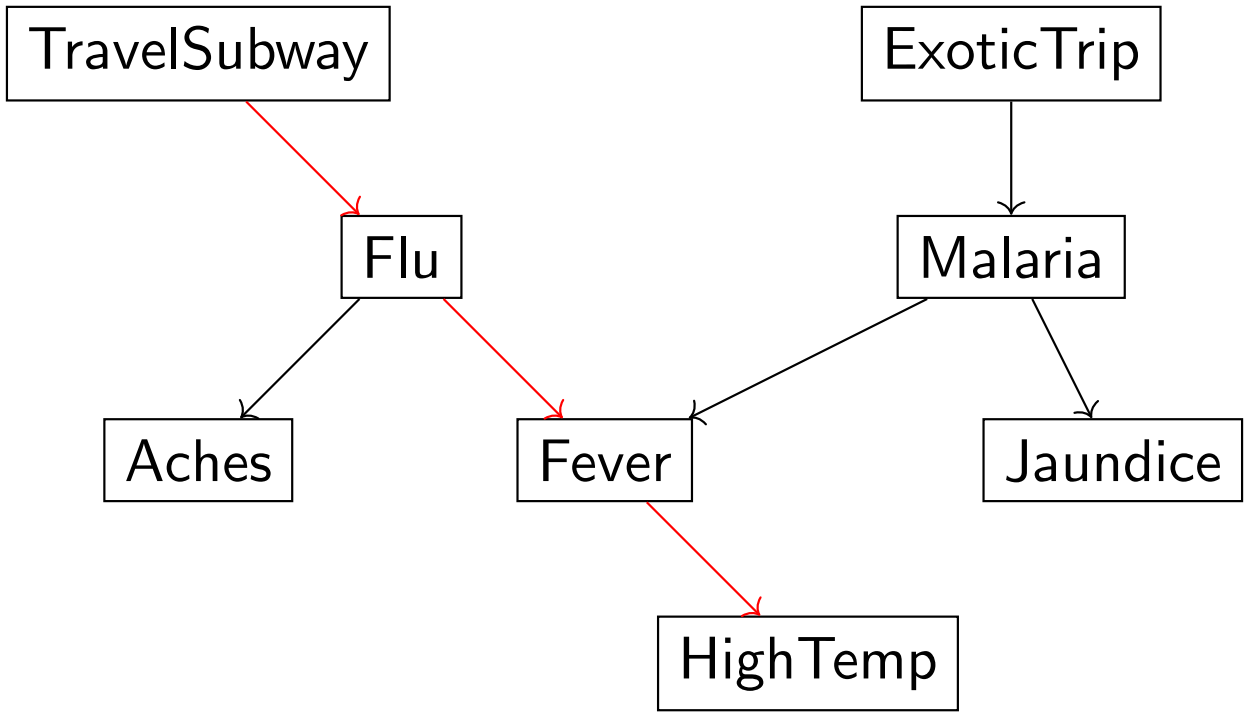
**Q #1:** Are **TravelSubway** and **HighTemp** independent?




- A path from TravelSubway to Flu to Fever to HighTemp.
- Two nodes on the path: Flu and Fever.
- Apply rule 1 to Flu, Flu is not observed, no blocking
- Apply rule 1 to Fever, Fever is not observed, no blocking

# Q: Applying D-separation

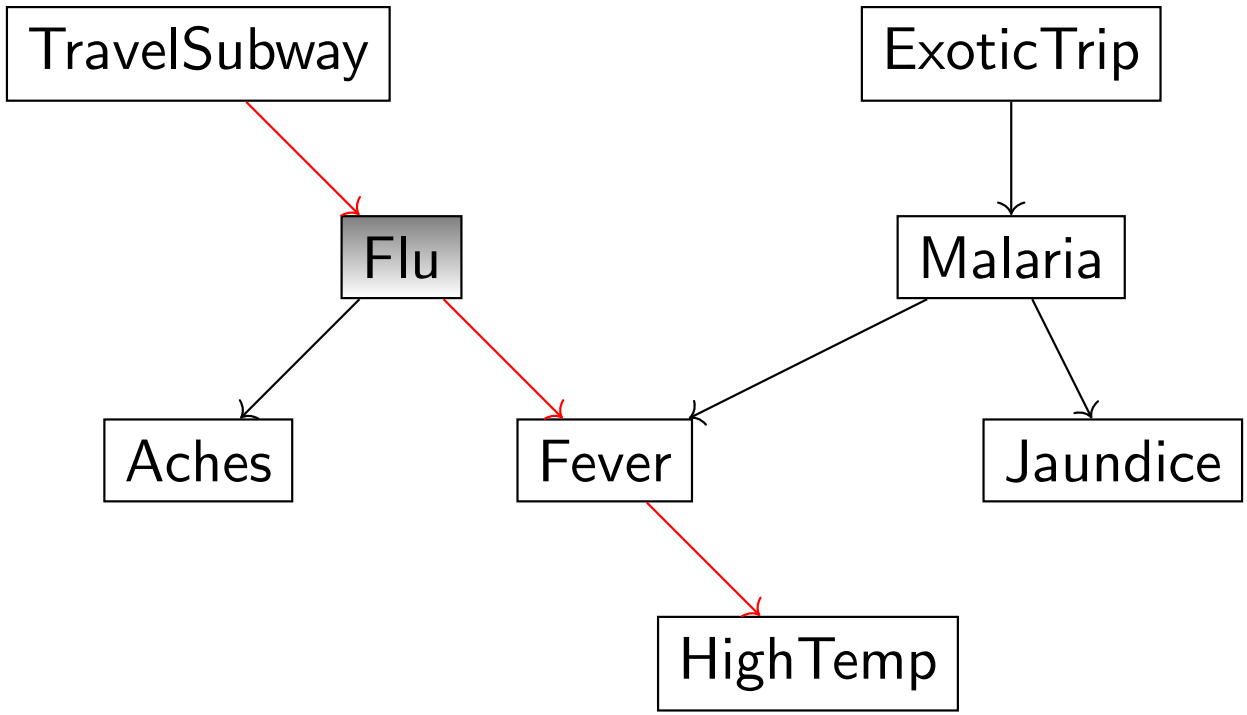
**Q #1:** Are **TravelSubway** and **HighTemp** independent?




→ No, they not independent.

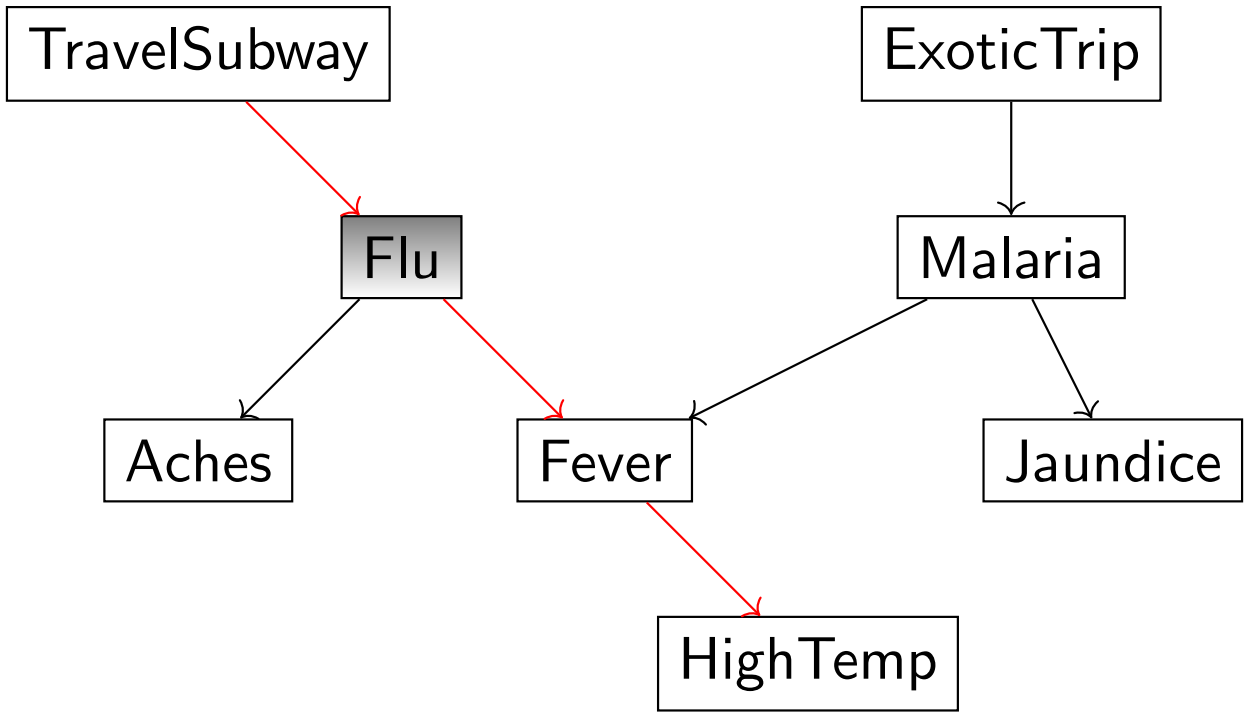
# Q: Applying D-separation

**Q #2:** Are **TravelSubway** and **HighTemp** independent given **Flu**?



# Q: Applying D-separation

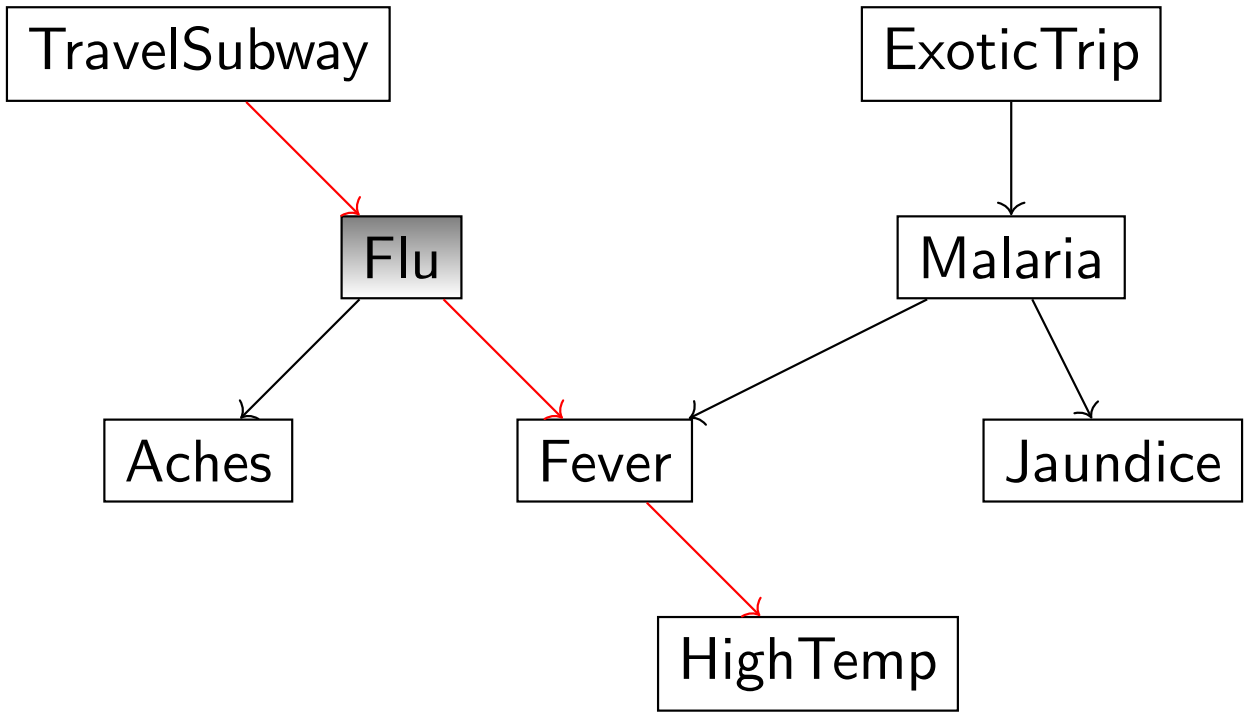
**Q #2:** Are **TravelSubway** and **HighTemp** independent given **Flu**?




- A path from TravelSubway to Flu to Fever to HighTemp.
- Flu and Fever are the nodes on the path.
- Apply rule 1 to Fever, Fever is not observed, no blocking.
- Apply rule 1 to Flu, Flu is observed, the path is blocked.

# Q: Applying D-separation

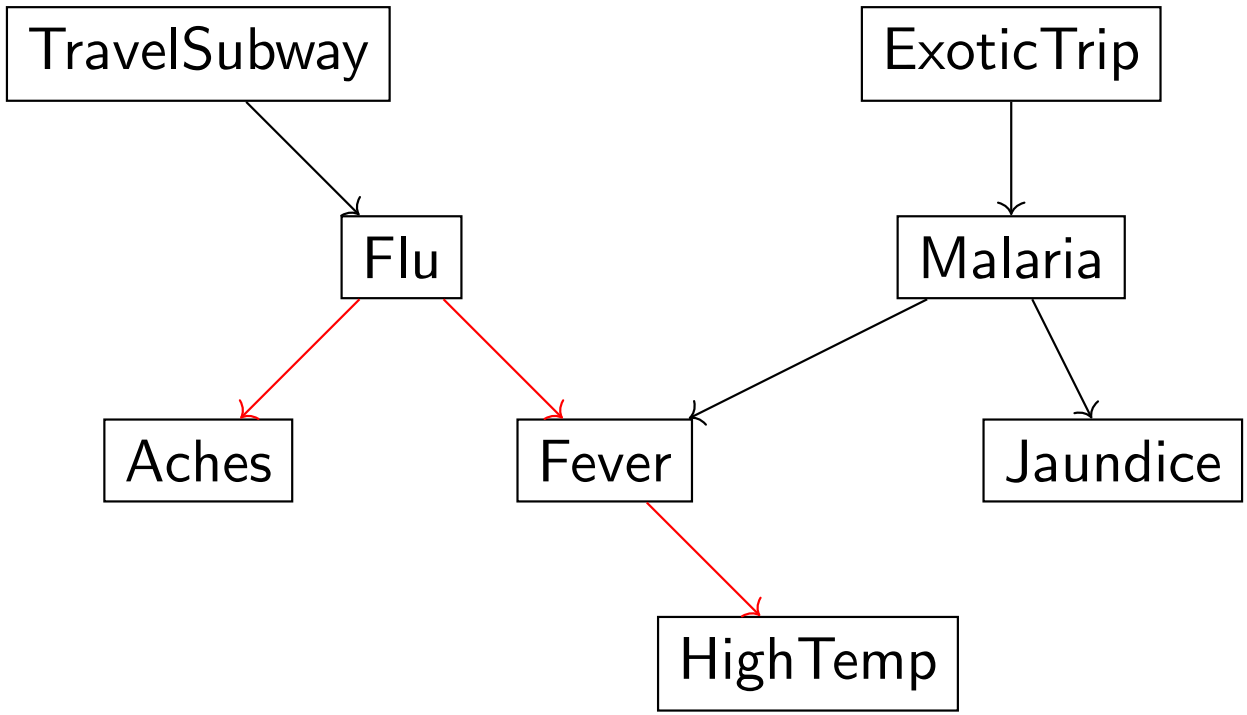
**Q #2:** Are **TravelSubway** and **HighTemp** independent given **Flu**?




→ Yes, they are independent given Fever

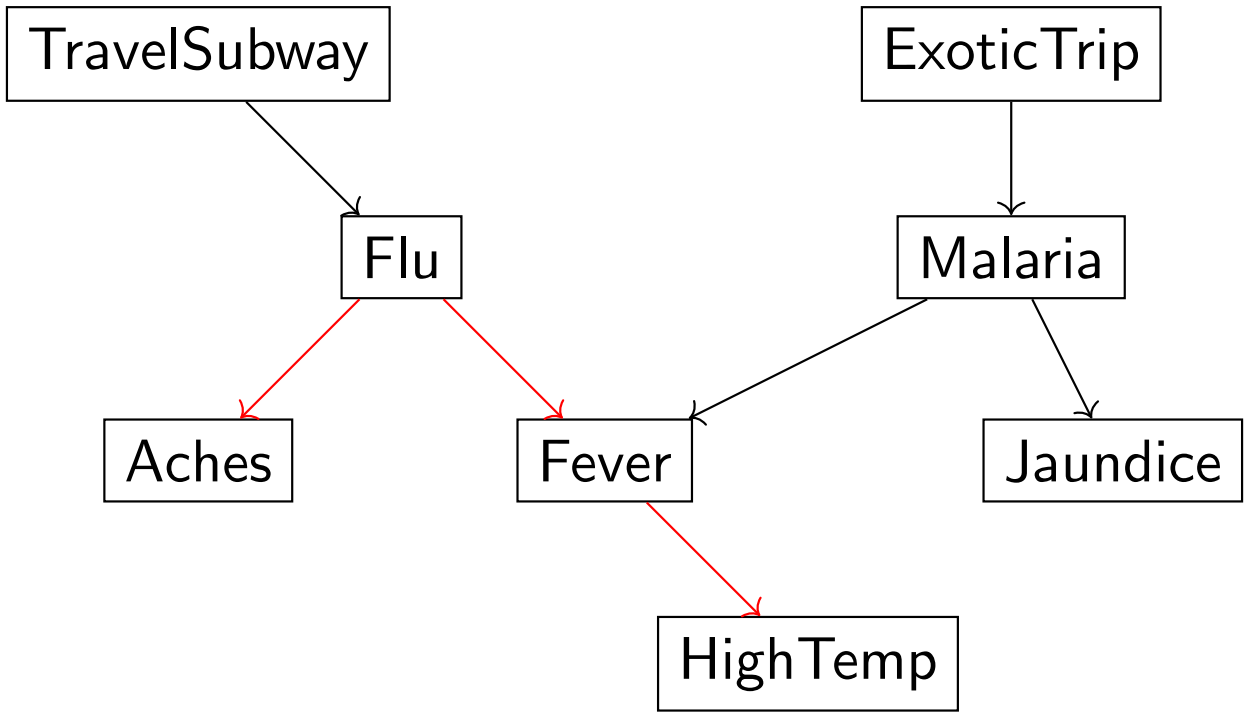
# Q: Applying D-separation

**Q #3:** Are **Aches** and **HighTemp** independent?



# Q: Applying D-separation

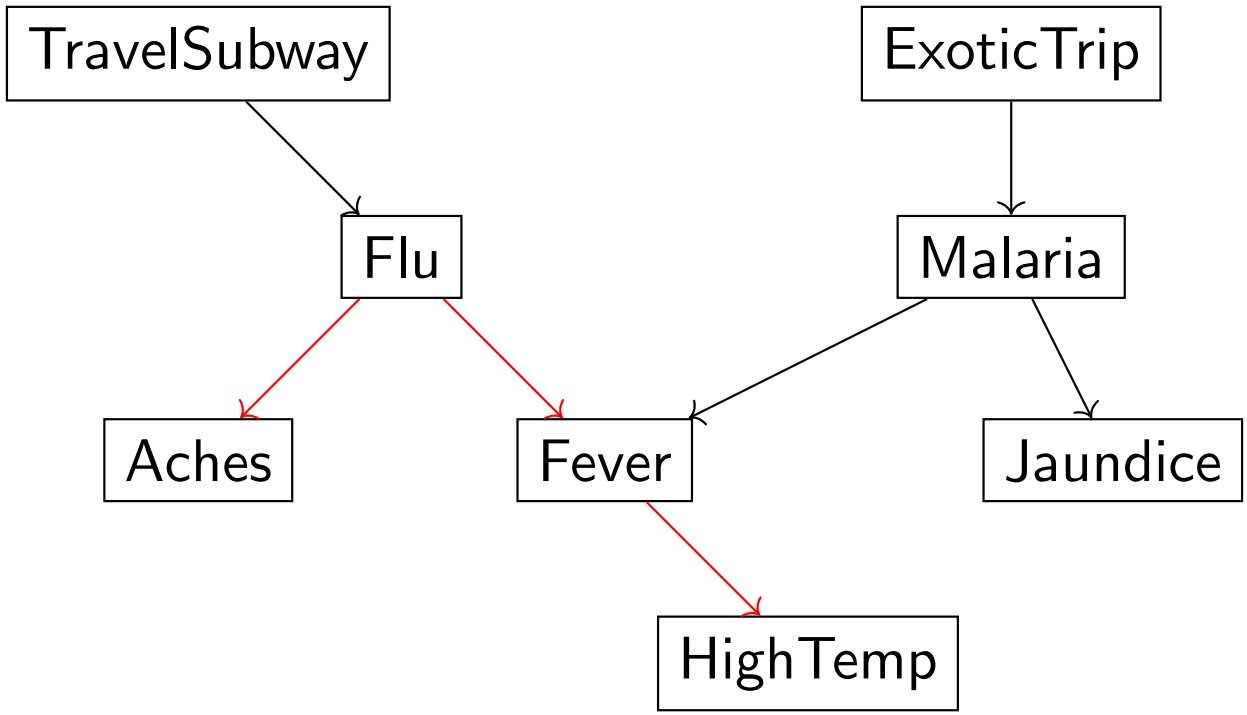
**Q #3:** Are **Aches** and **HighTemp** independent?




- One path from Aches to Flu to Fever to High Temp
- Apply Rule 1 to Fever, Fever is not observed, no blocking
- Apply Rule 2 to Flu, Flu is not observed, no blocking

# Q: Applying D-separation

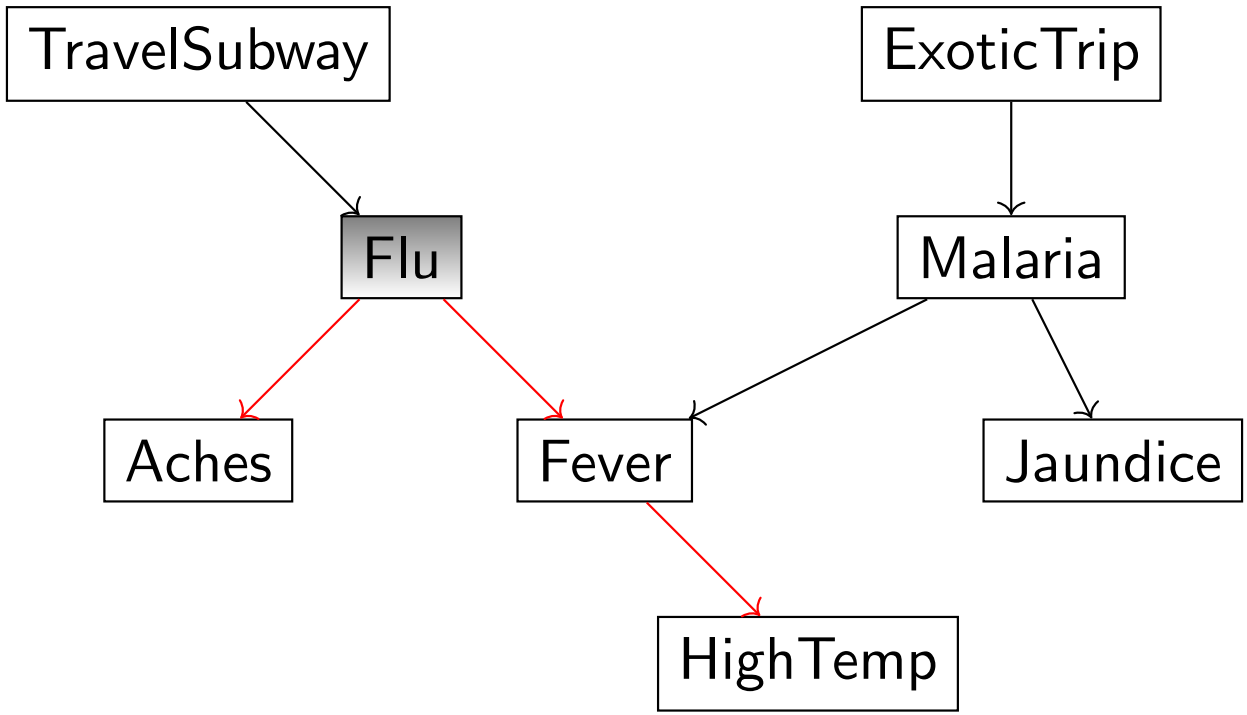
**Q #3:** Are **Aches** and **HighTemp** independent?




→ No, they are not independent.

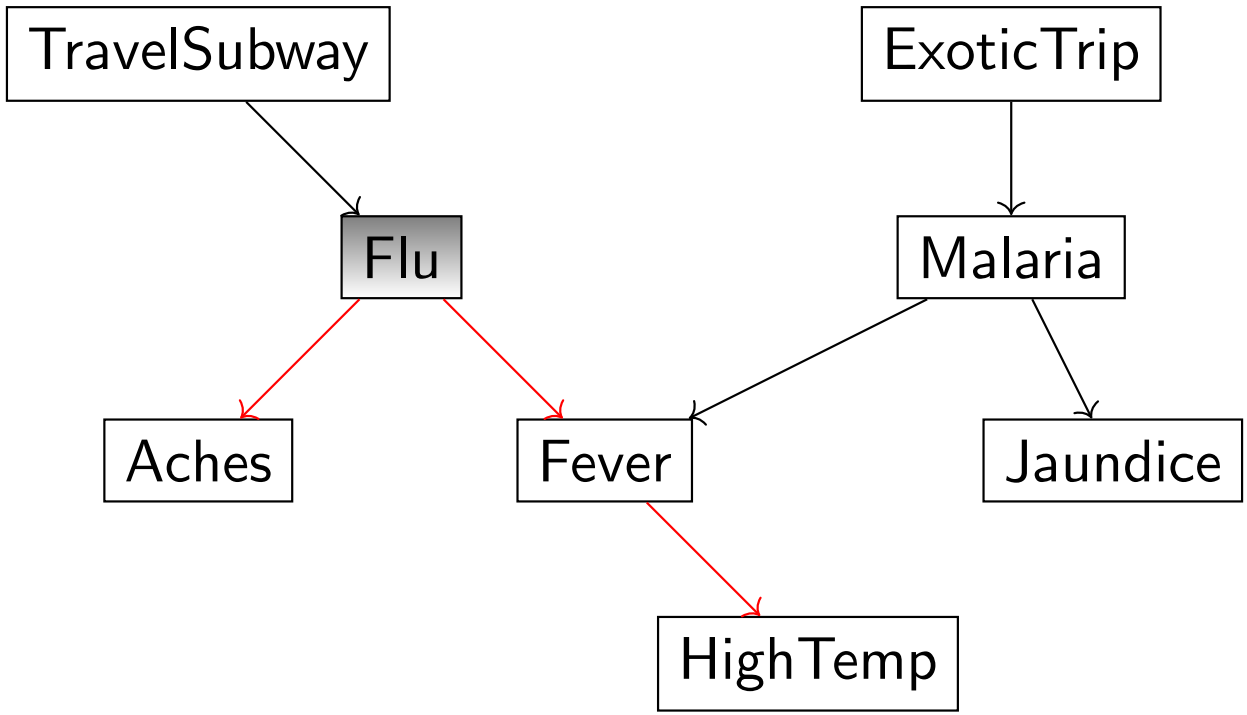
# Q: Applying D-separation

**Q #4:** Are **Aches** and **HighTemp** independent given **Flu**?



# Q: Applying D-separation

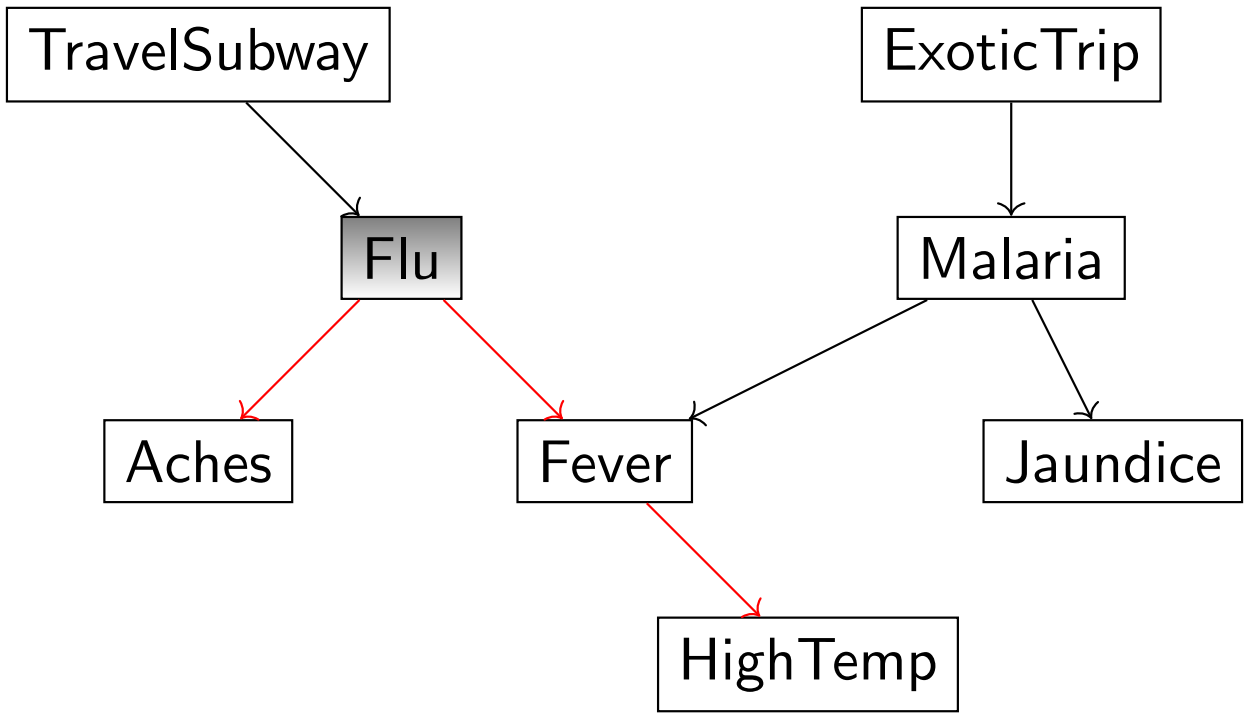
**Q #4:** Are **Aches** and **HighTemp** independent given **Flu**?




- One path from Aches to Flu to Fever to High Temp.
- Apply Rule 1 to Fever, Fever is not observed, no blocking.
- Apply Rule 2 to Flu, Flu is observed, and the path is blocked.

# Q: Applying D-separation

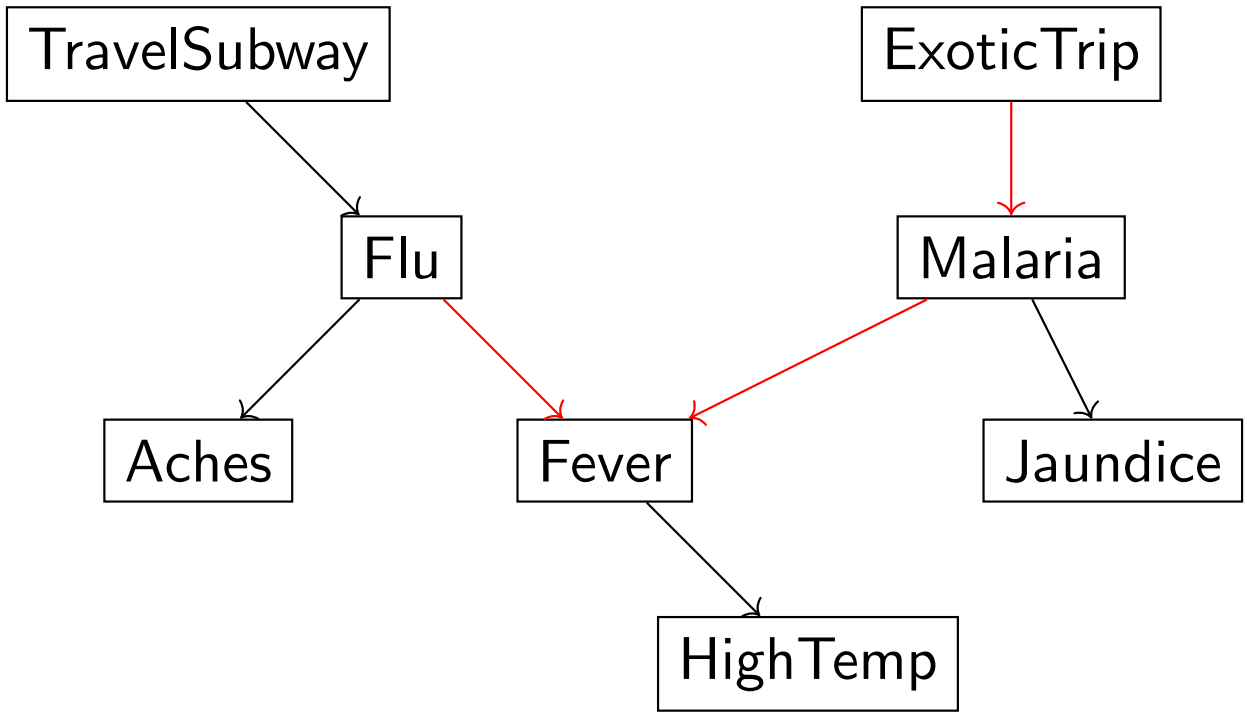
**Q #4:** Are **Aches** and **HighTemp** independent given **Flu**?




→ Yes, they are independent given Flu.

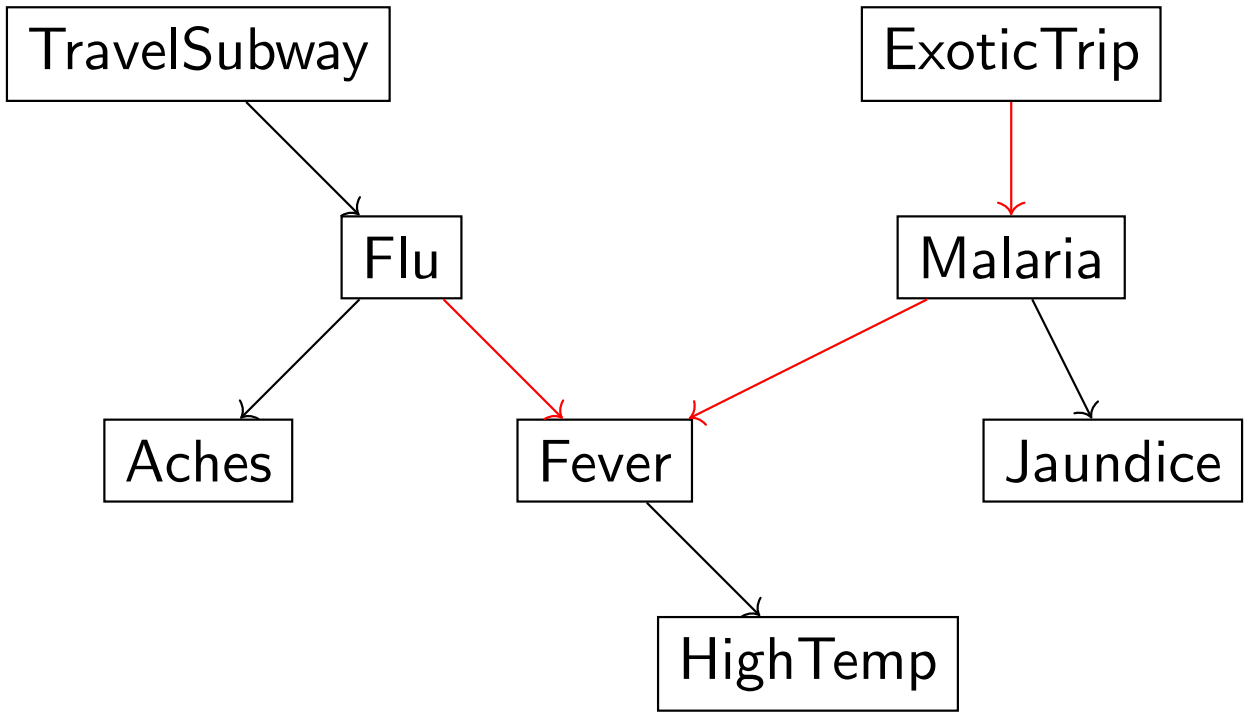
# Q: Applying D-separation

**Q #5:** Are **Flu** and **ExoticTrip** independent?



# Q: Applying D-separation

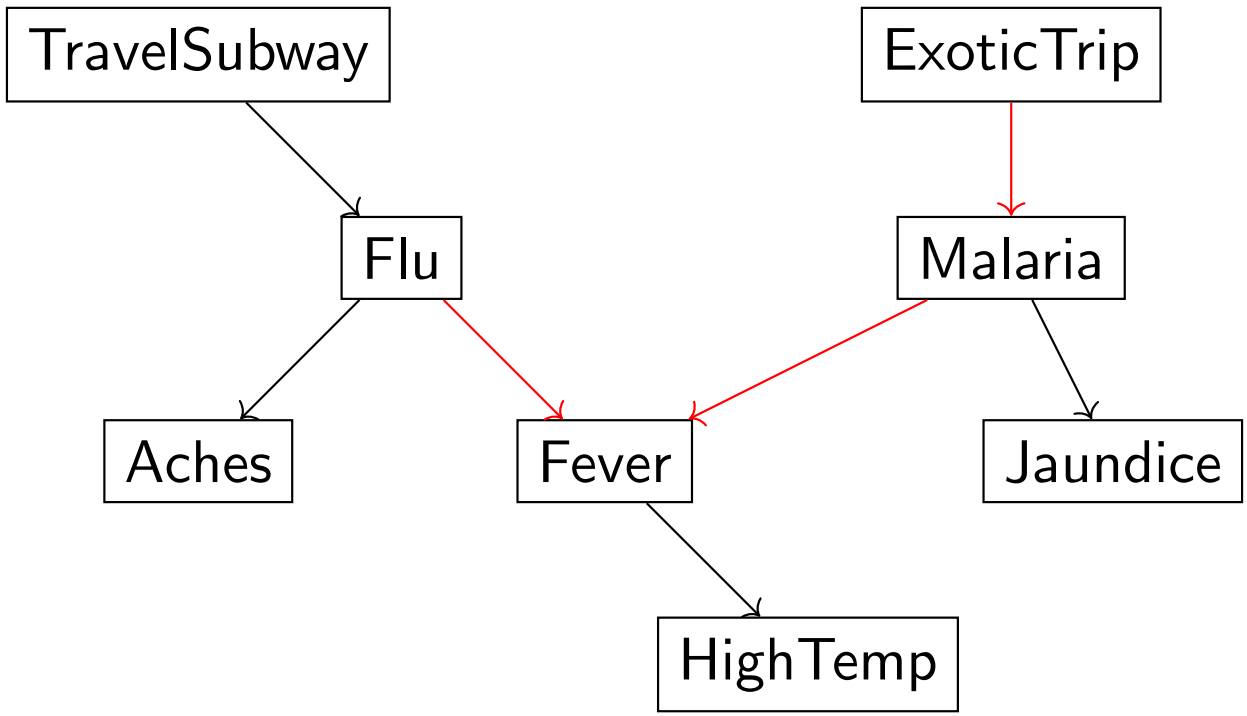
**Q #5:** Are **Flu** and **ExoticTrip** independent?




- One path from Flu to Fever to Malaria to ExoticTrip.
- Apply rule 1 to Malaria, which is not observed, not blocking.
- Apply rule 3 to Fever, itself and descendent are not observed, the path is being blocked.

# Q: Applying D-separation

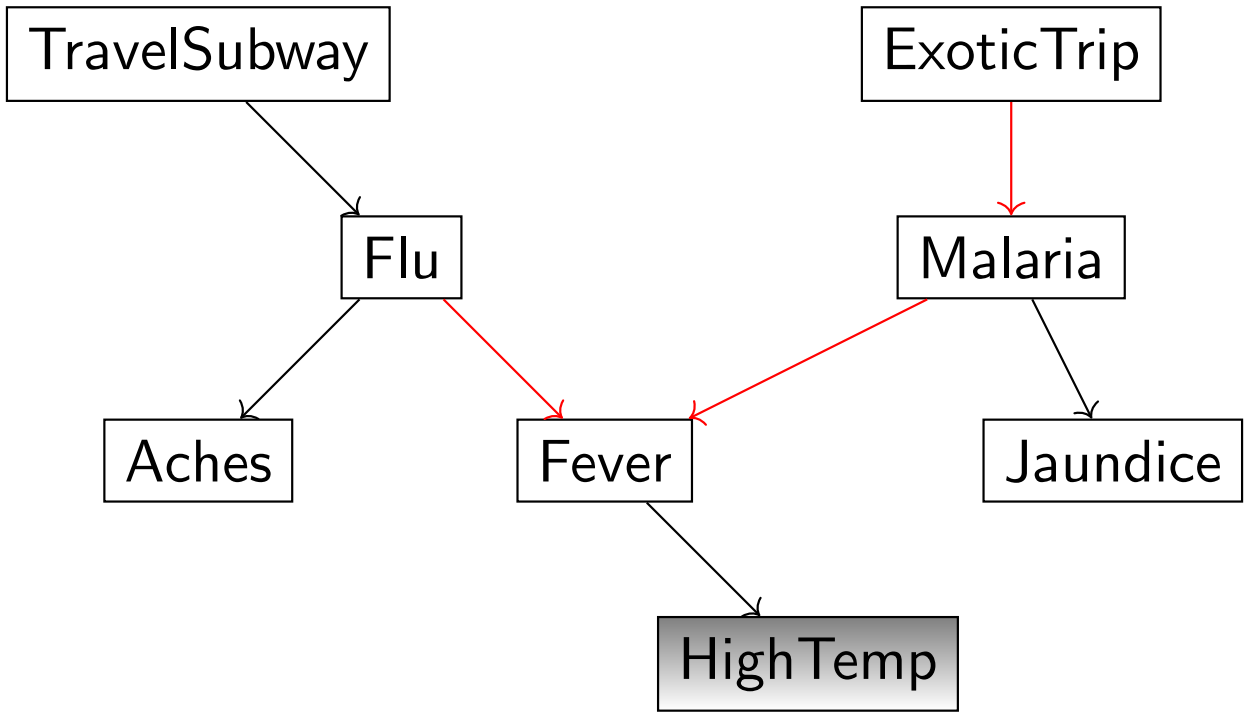
**Q #5:** Are **Flu** and **ExoticTrip** independent?




→ Yes, they are indepent.

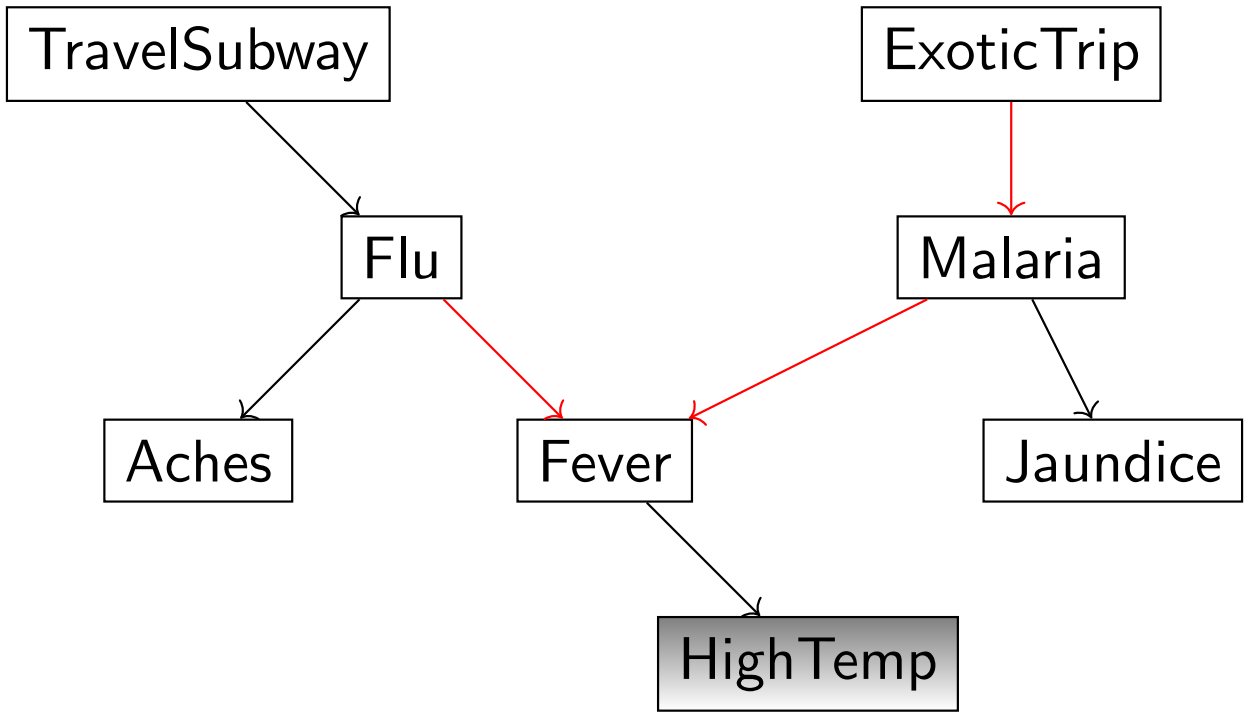
# Q: Applying D-separation

**Q #6:** Are **Flu** and **ExoticTrip** independent given **HighTemp**?



# Q: Applying D-separation

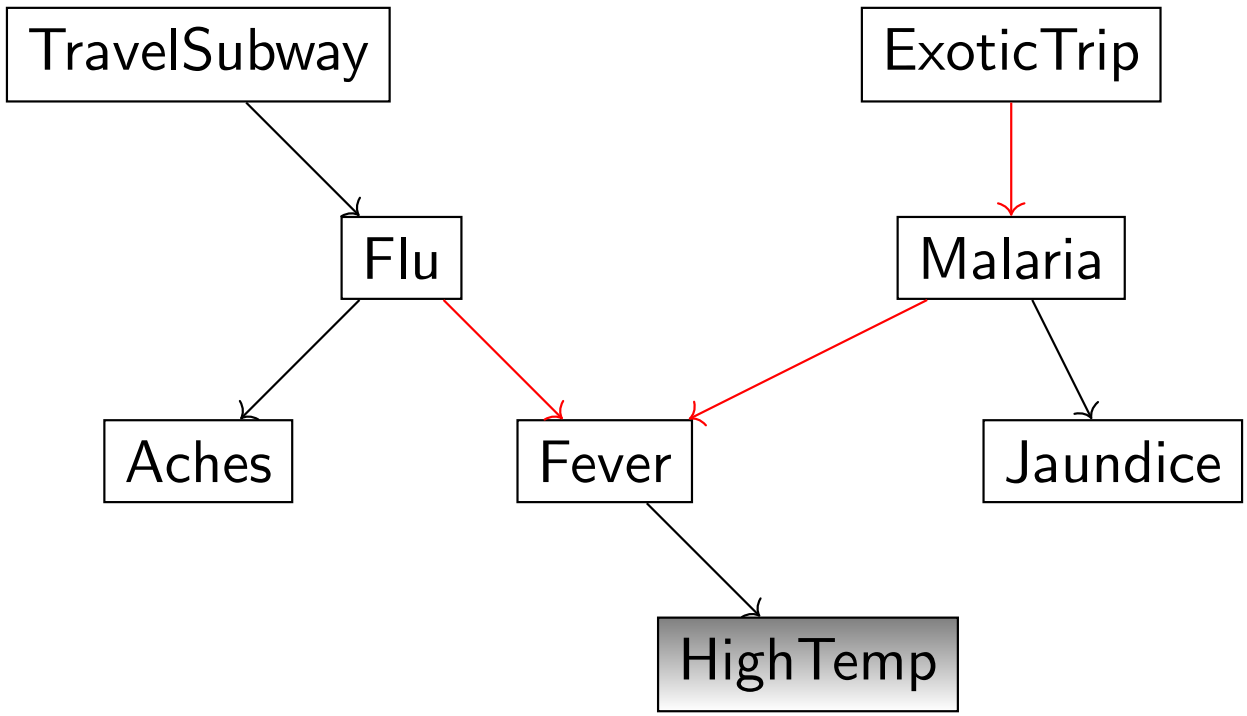
**Q #6:** Are **Flu** and **ExoticTrip** independent given **HighTemp**?




- One path from Aches to Flu to Fever to High Temp.
- Apply rule 1 to Fever, Fever is not observed, no blocking.
- Apply rule 3 to Fever, its descendent is observed, the path is not being blocked.

# Q: Applying D-separation

**Q #6:** Are **Flu** and **ExoticTrip** independent given **HighTemp**?




→ They are not independent given HighTemp.

# Constructing Bayesian Networks

- For a joint probability distribution, there are many correct Bayesian networks.
- Given a Bayesian network A, a Bayesian network B is correct if and only if the following is true:

  *If Bayesian network B requires two variables to satisfy an independence relationship, Bayesian network A must also require the two variables to satisfy the same independence relationship.*
- Bayesian network B could miss independence from Network A, but it cannot miss dependence.
- We prefer a Bayesian network that requires fewer probabilities.

# Constructing a Correct Bayesian Network

1. Order the variables $\{X_1, \ldots, X_n\}$.
2. For each variable $X_i$ in the ordering,
   - 2.1 **Choose the node's parents:**

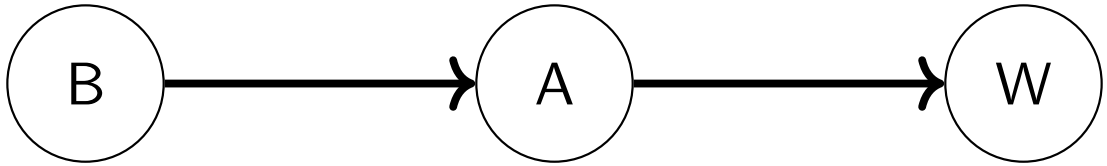
     Choose the smallest set of parents from $\{X_1, \ldots, X_{i-1}\}$ such that given $Parents(X_i)$, $X_i$ is independent of all the nodes in $\{X_1, \ldots, X_{i-1}\} - Parents(X_i)$. Formally,

     $$P(X_i|Parents(X_i)) = P(X_i|X_{i-1} \wedge \cdots \wedge X_1).$$

   - 2.2 Create a link from each parent of $X_i$ to the node $X_i$.
   - 2.3 Write down the conditional probability table $P(X_i|Parents(X_i))$.

# Example 1: Construct a Bayes Net

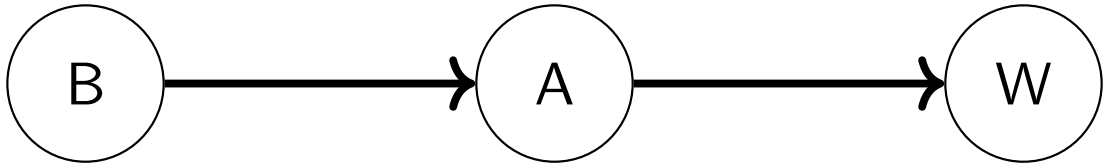
Consider the Bayesian network.




Construct a correct Bayesian network by adding the variables in the order: W, A, and B.

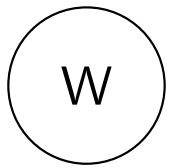
# Example 1: Construct a Bayes Net

Consider the Bayesian network.



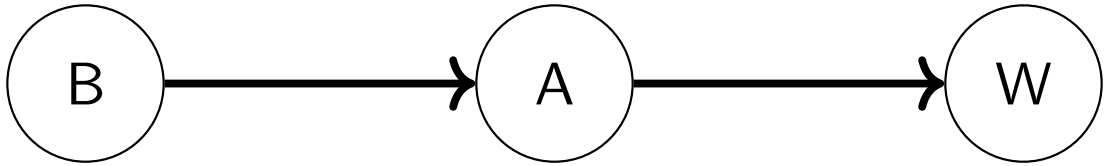

Construct a correct Bayesian network by adding the variables in the order: W, A, and B.

Set: {}



# Example 1: Construct a Bayes Net

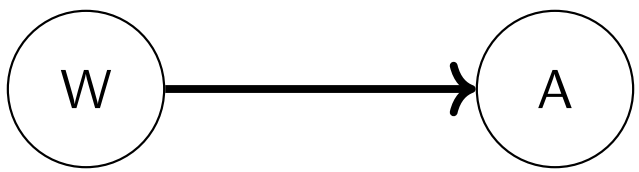
Consider the Bayesian network.




Construct a correct Bayesian network by adding the variables in the order: W, A, and B.

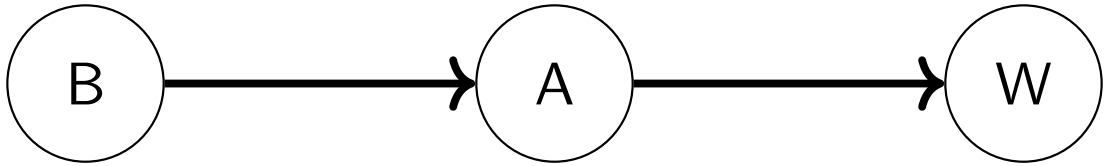
Set: $\{W\}$

- Is A dependent on W?



# Example 1: Construct a Bayes Net

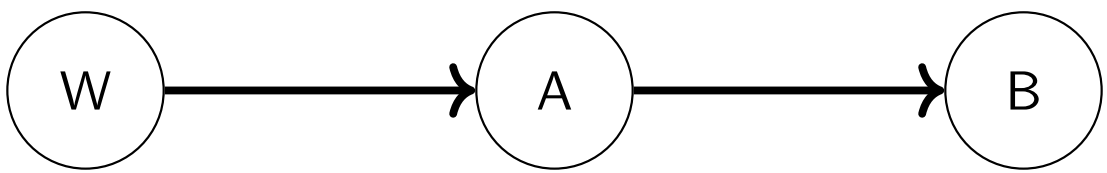
Consider the Bayesian network.




Construct a correct Bayesian network by adding the variables in the order: W, A, and B.

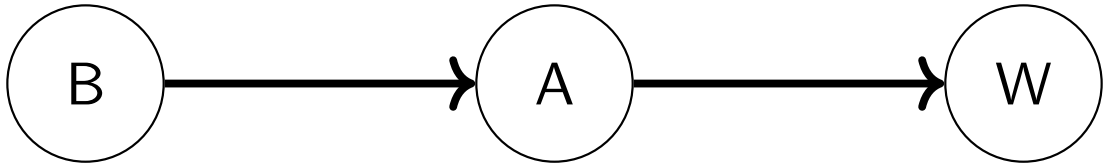
Set: {W, A}

- Is B independent from A given W?
- Is B independent from W given A?



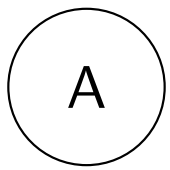
# Example 1: Construct a Bayes Net

Consider the Bayesian network.



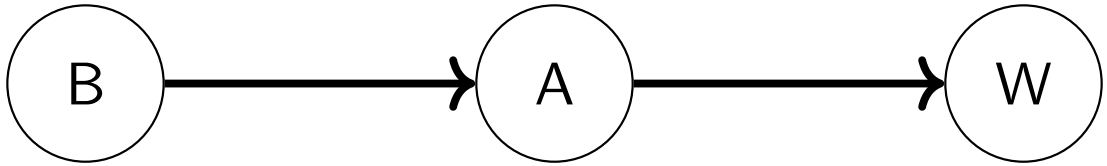

Construct a correct Bayesian network by adding the variables in the order: A, W, and B.

Set: $\{\}$



# Example 1: Construct a Bayes Net

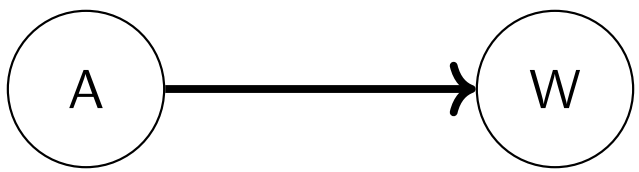
Consider the Bayesian network.




Construct a correct Bayesian network by adding the variables in the order: A, W, and B.

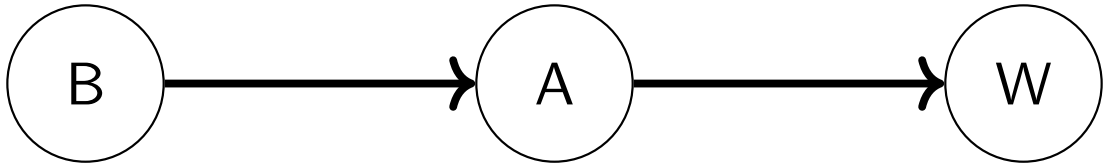
Set: $\{A\}$

- is W dependent on A?



# Example 1: Construct a Bayes Net

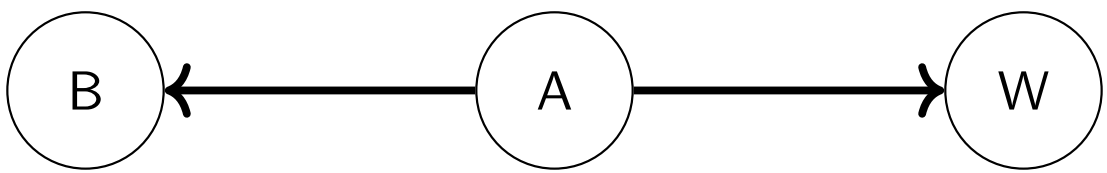
Consider the Bayesian network.




Construct a correct Bayesian network by adding the variables in the order: A, W, and B.

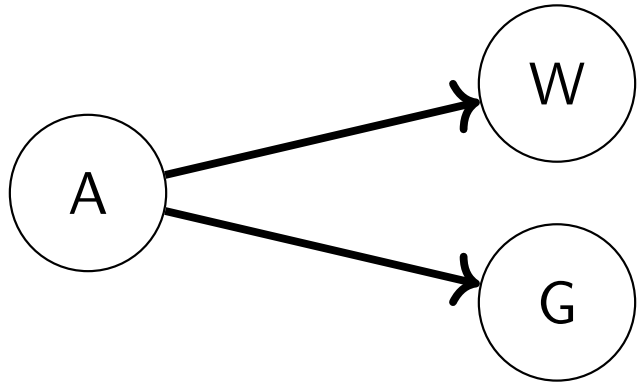
Set: {A, W}

- B is independent from A given W?
- B is independent from W given A?



# Example 2: Construct a Bayes Net

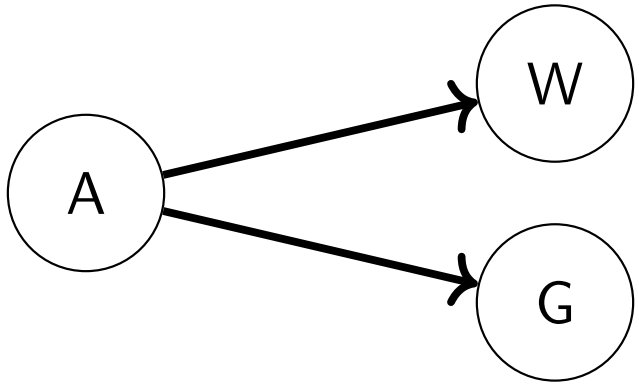
Consider the Bayesian network:



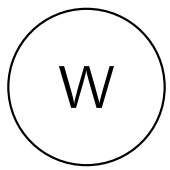

Construct a correct Bayesian network by adding the variables in the order: W, G, and A.

# Example 2: Construct a Bayes Net

Consider the Bayesian network:

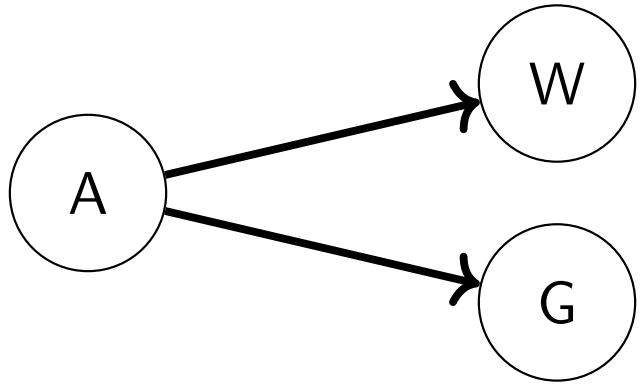



Set: {}



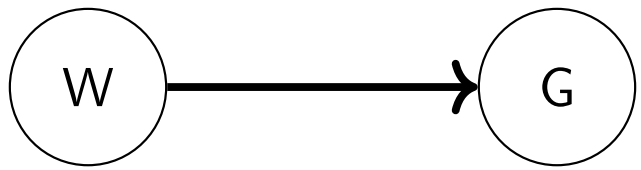
# Example 2: Construct a Bayes Net

Consider the Bayesian network:



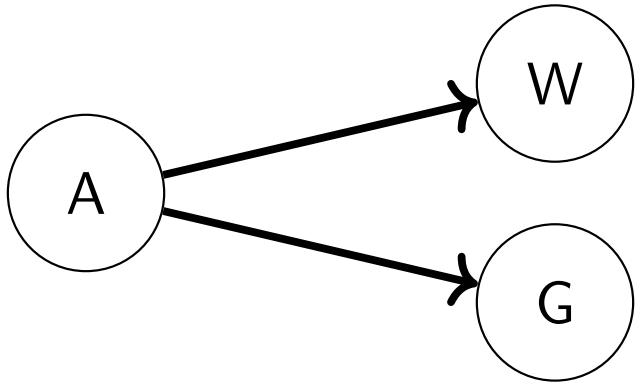

Set: $\{W\}$

- Is G dependent on W?



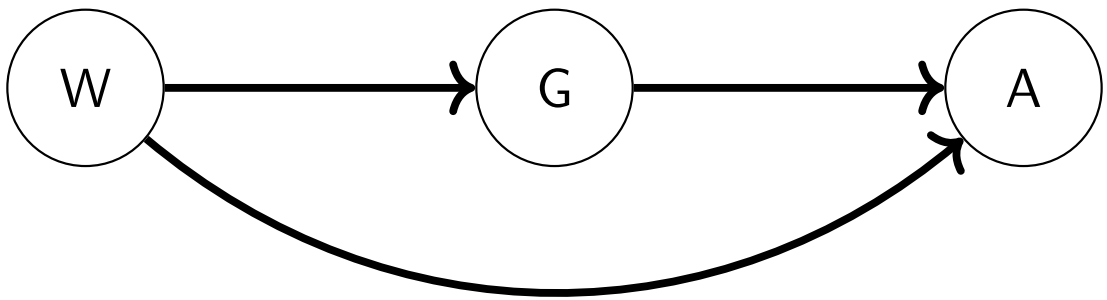
# Example 2: Construct a Bayes Net

Consider the Bayesian network:



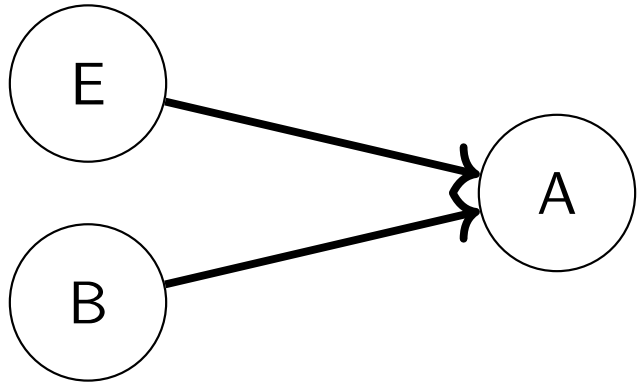

- Is A independent on G given W?
- Is A independent on W given G?

Set: {W, G}



# Example 3: Construct a Bayes Net

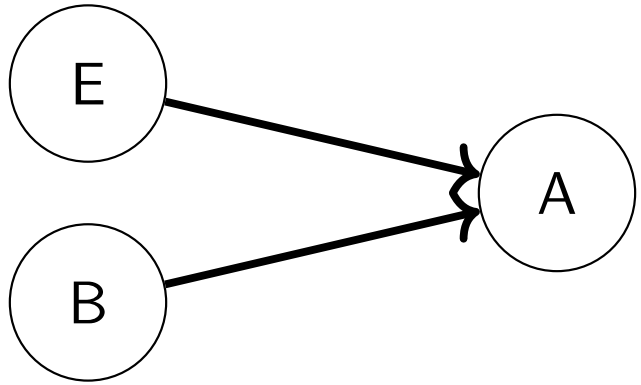
Consider the Bayesian network.



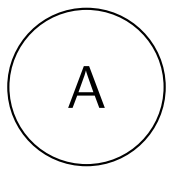

Construct a correct Bayesian network by adding the variables in the order: A, B, and E.

# Example 3: Construct a Bayes Net

Consider the Bayesian network.

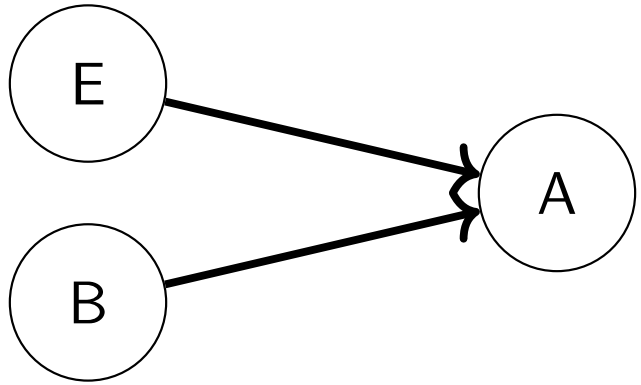



set: $\{\}$



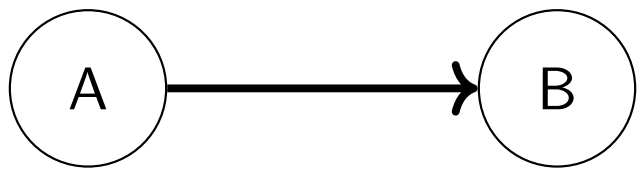
# Example 3: Construct a Bayes Net

Consider the Bayesian network.



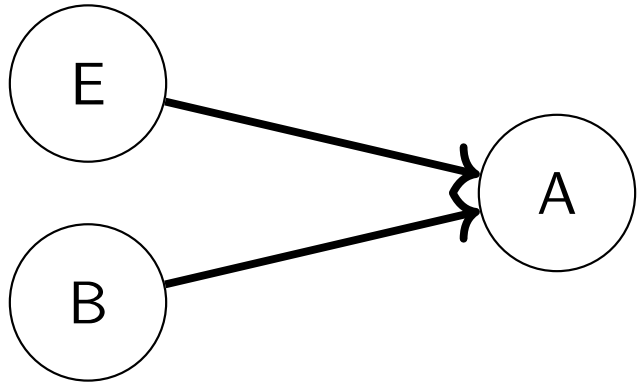

set: {A}

- Is B dependent on A?



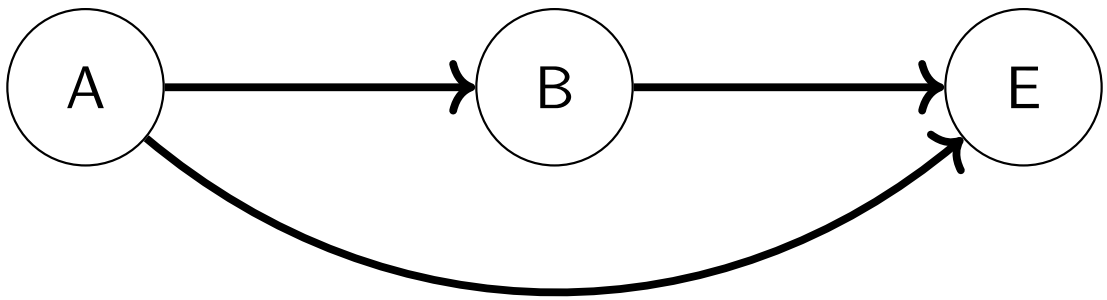
# Example 3: Construct a Bayes Net

Consider the Bayesian network.



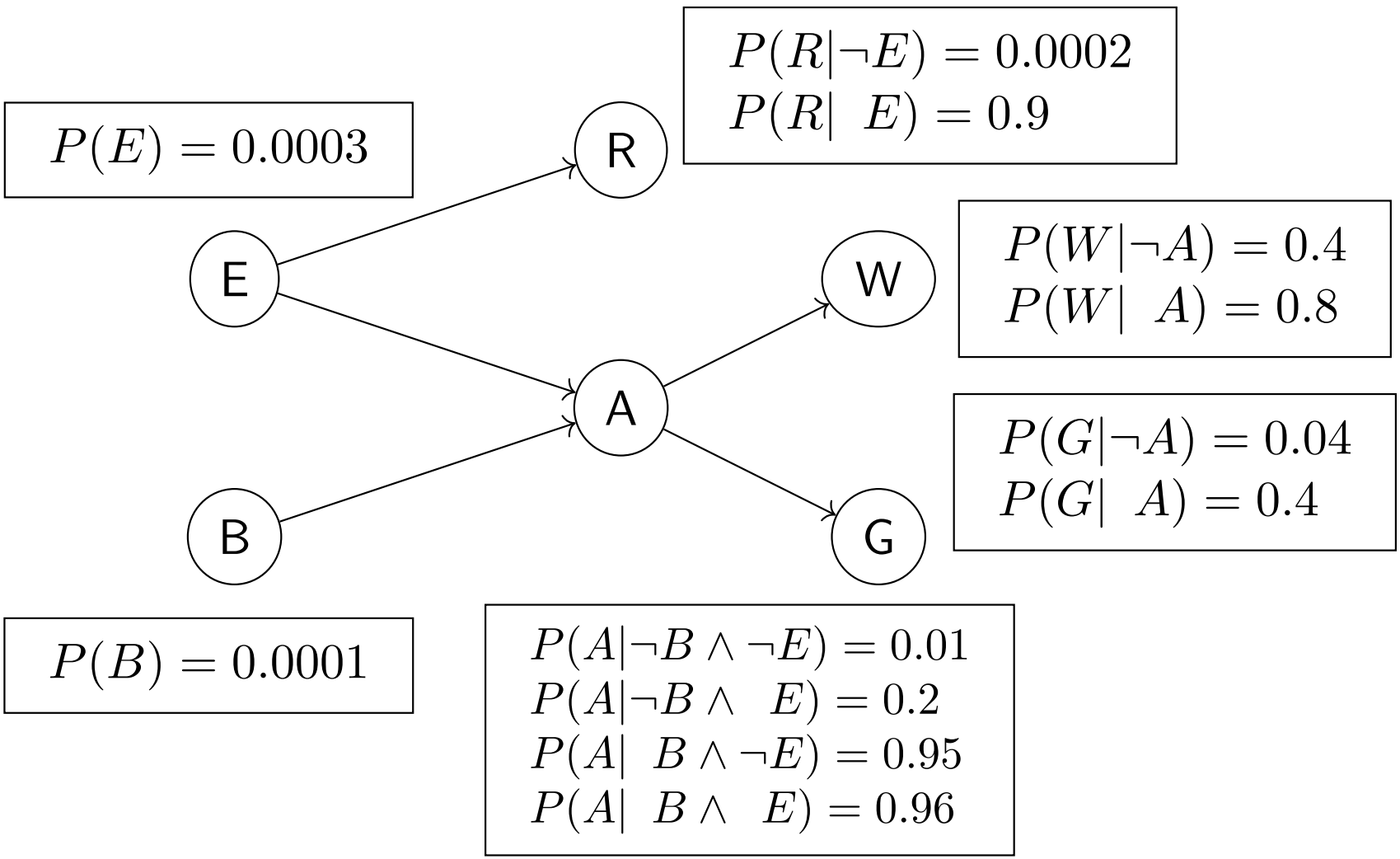

set: {A, B}

- Is E independent on A given B?
- Is E independent on B given A?



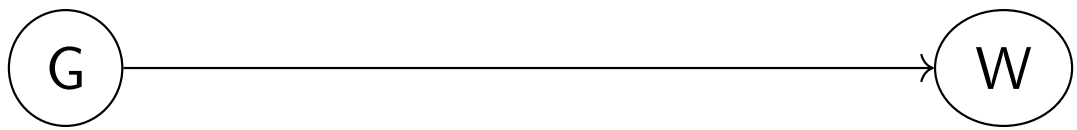
# Example 3: Reconstruct the Holmes Scenario Network

Construct a new Bayesian network from the Holmes scenario, using the following order for adding variables: $G, W, E, B, A, R$.

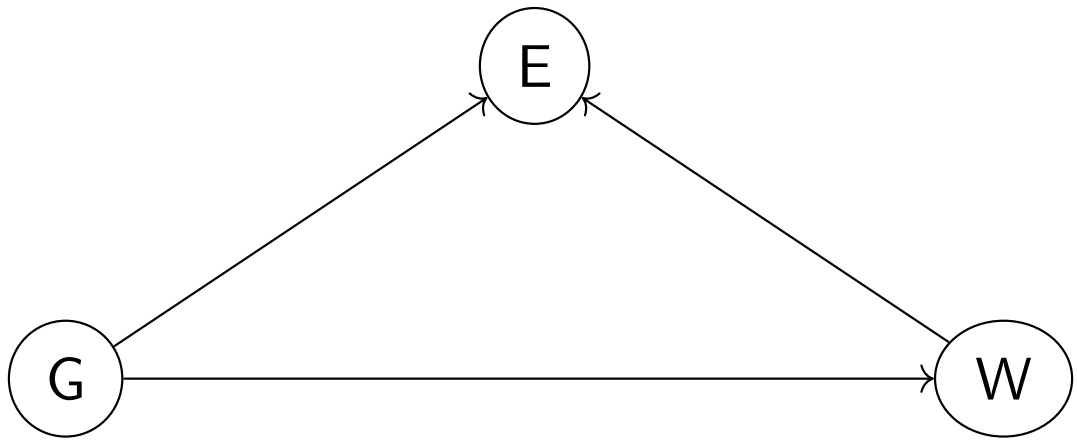


# Example 3: Reconstruct the Holmes Scenario Network

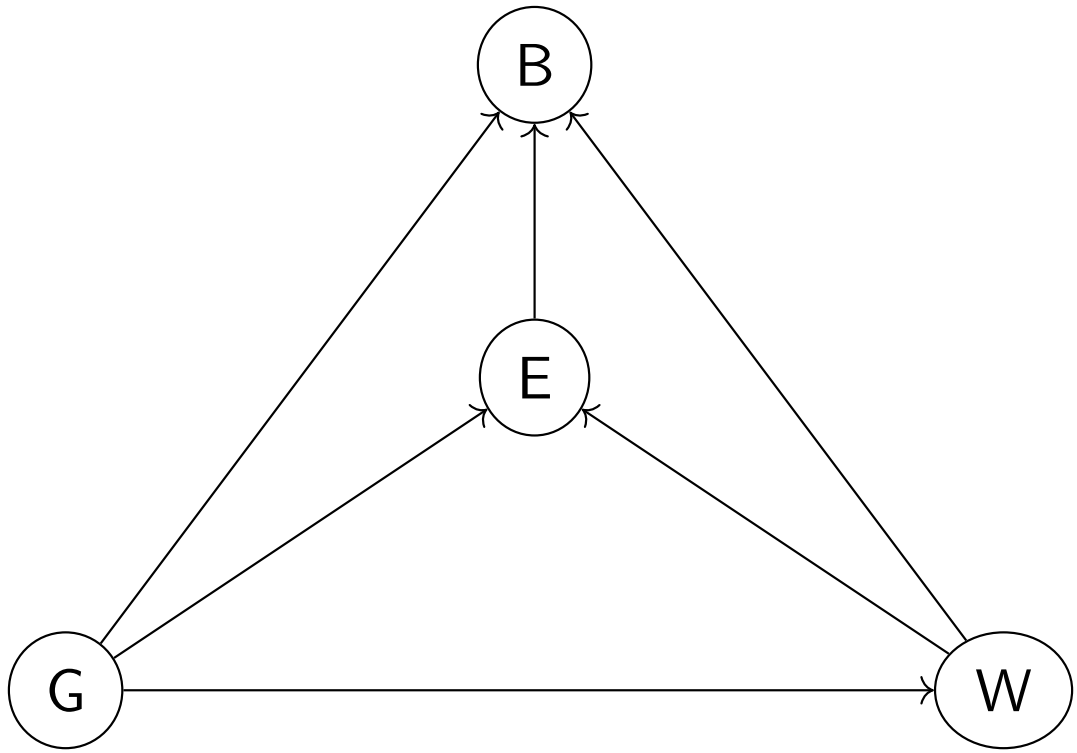
# Example 3: Reconstruct the Holmes Scenario Network



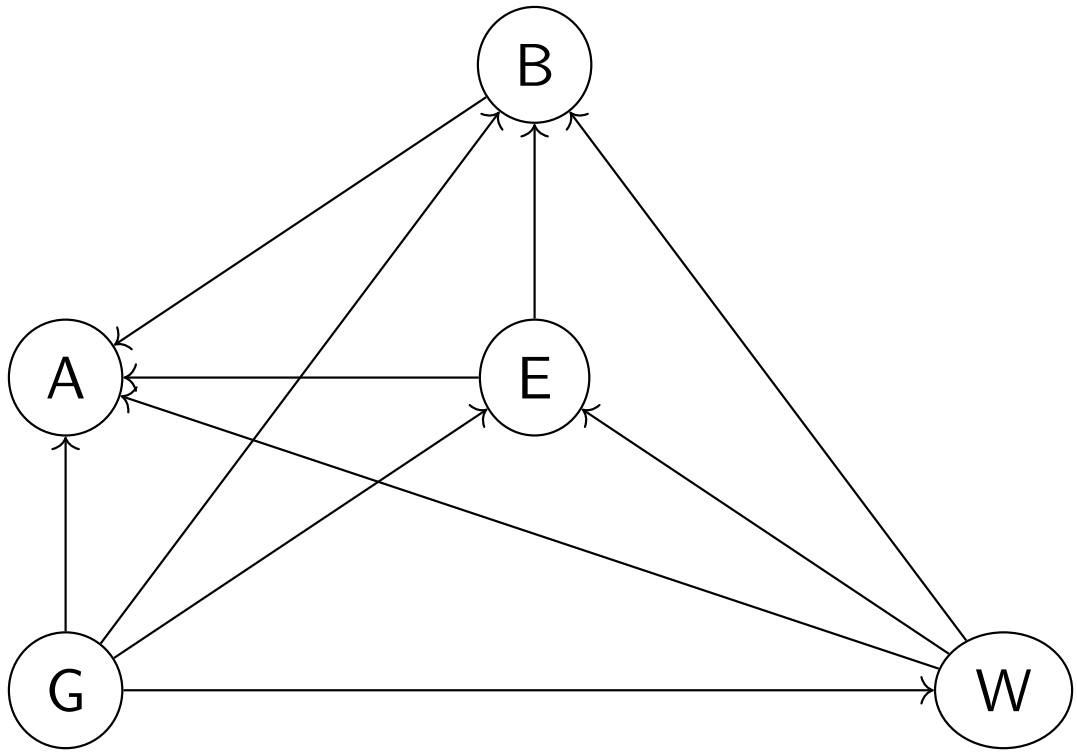
# Example 3: Reconstruct the Holmes Scenario Network



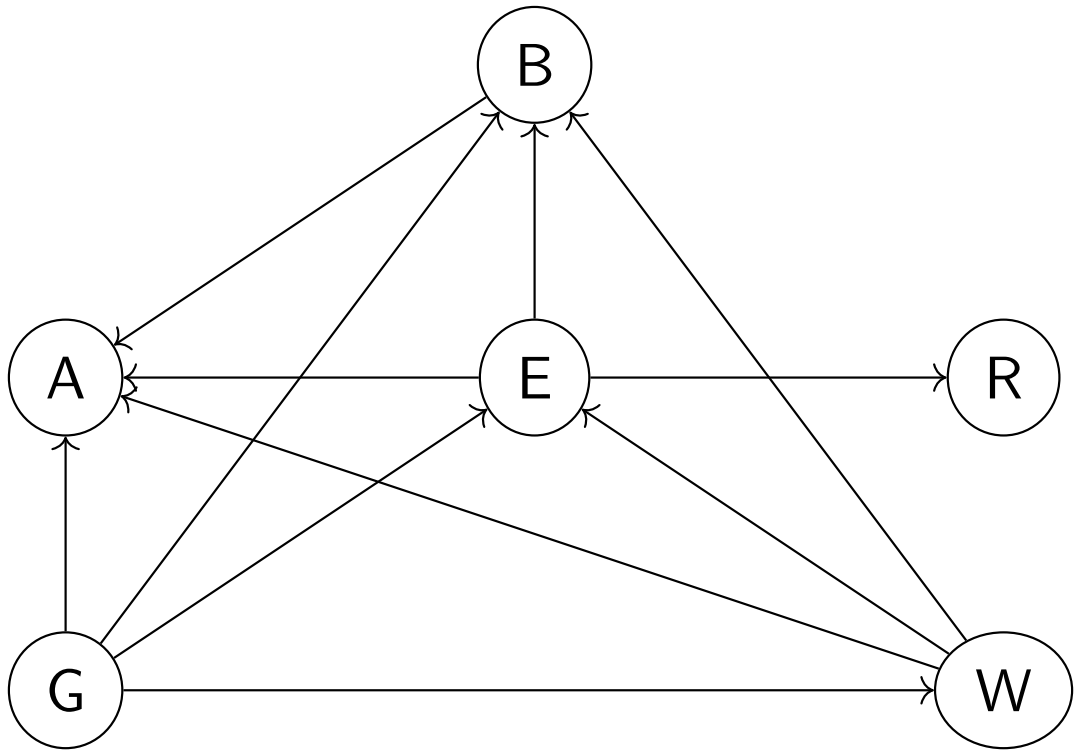
# Example 3: Reconstruct the Holmes Scenario Network



# Example 3: Reconstruct the Holmes Scenario Network



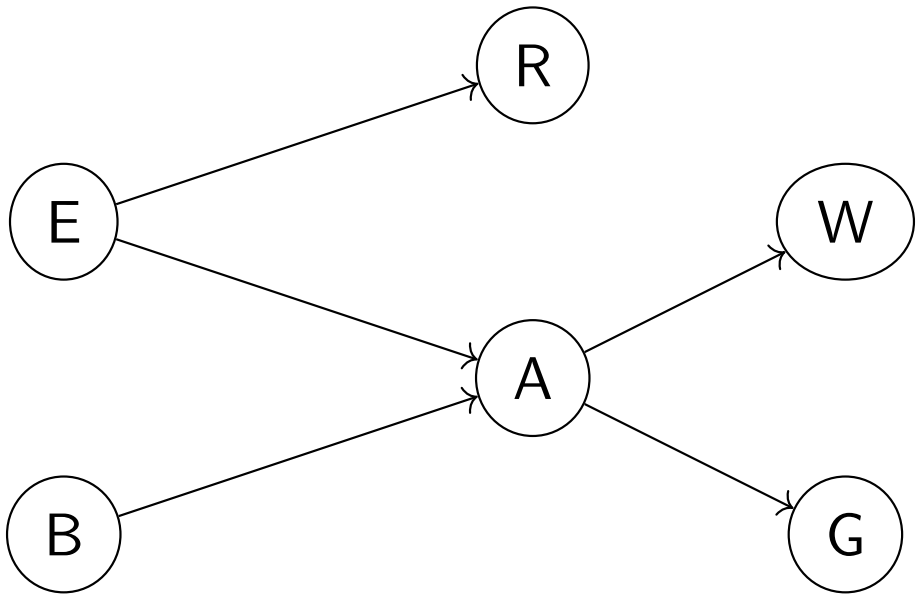
# Example 3: Reconstruct the Holmes Scenario Network



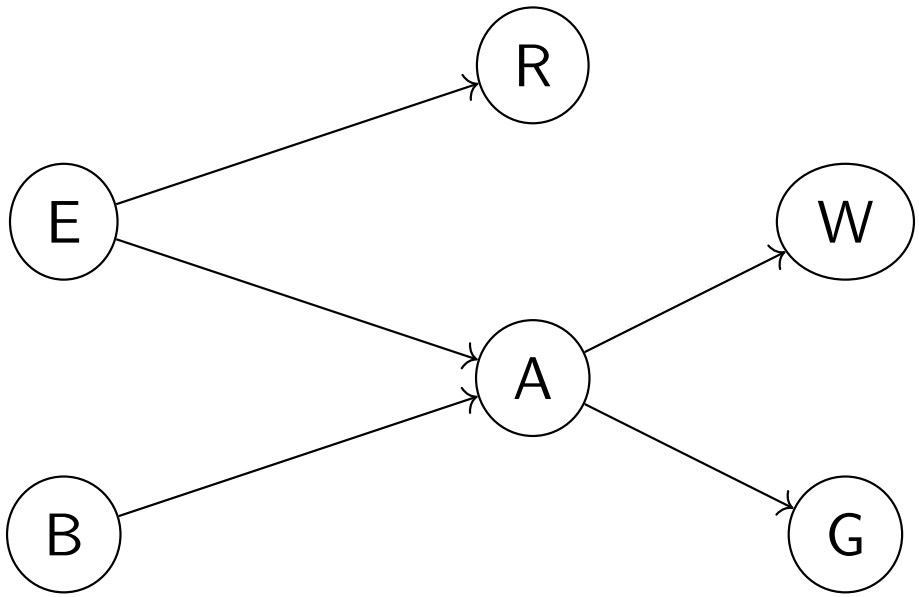
# Example 3: Reconstruct the Holmes Scenario Network

- Gibbon and Watson can both cause Earthquake.
- Gibbon and Watson and Earthquake can cause Burglary.
- Number of probabilities: $1 + 2 + 4 + 8 + 16 + 2 = 33$.
- Previous, we only need 12 probabilities.

# What is the correct order of correct Reconstruction



# What is the correct order of correct Reconstruction




- Correct Order: E, R, B, A, W, G, Let's validate it!
- Finding the most compact Bayesian Network is NP-hard!

# Constructing a Compact Bayesian Network

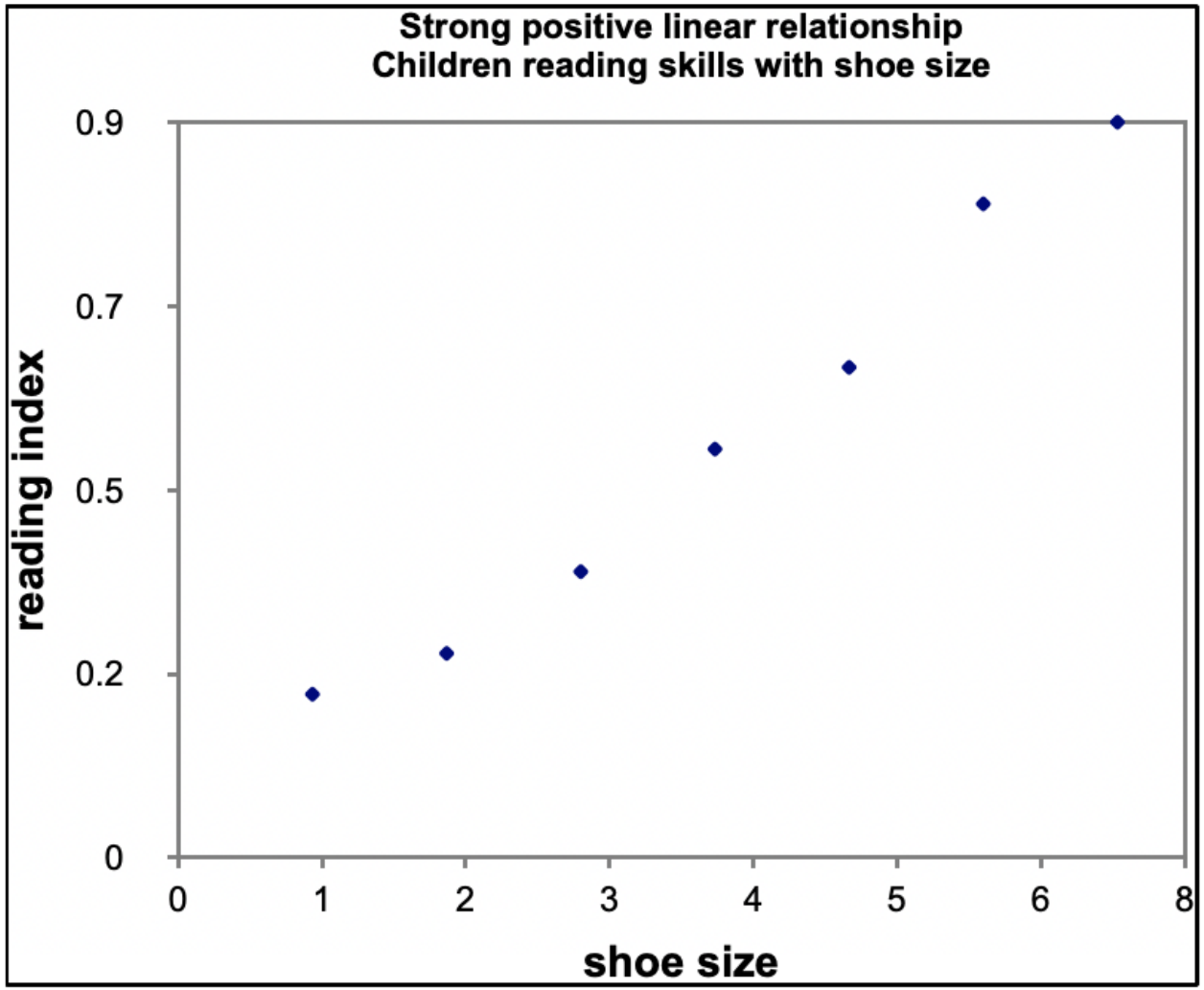
- What does an edge mean?
  Does an edge always represent a causal relationship?

  $\rightarrow$ An edge indicates an associational relationship that is not necessarily causal.

- How can we construct a Bayesian network with the smallest number of edges?

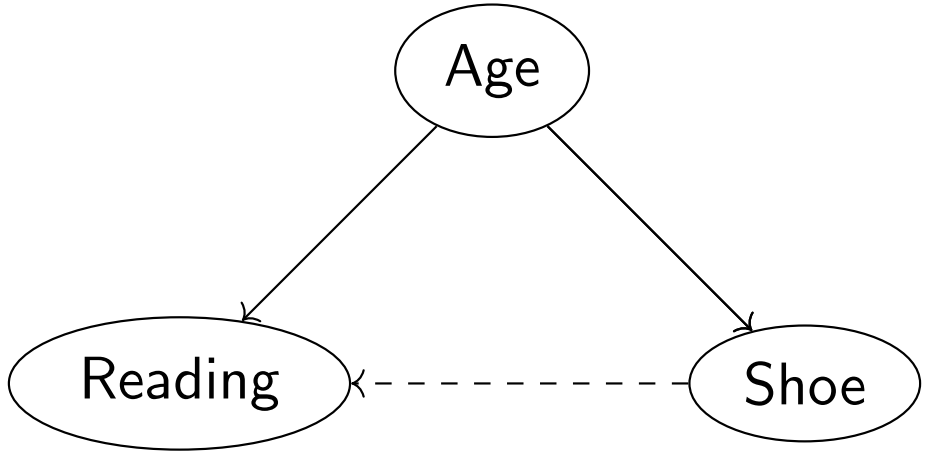
  $\rightarrow$ Cause precedes effect. So add causes first, then effects.

# Are Correlation and Causation the same?



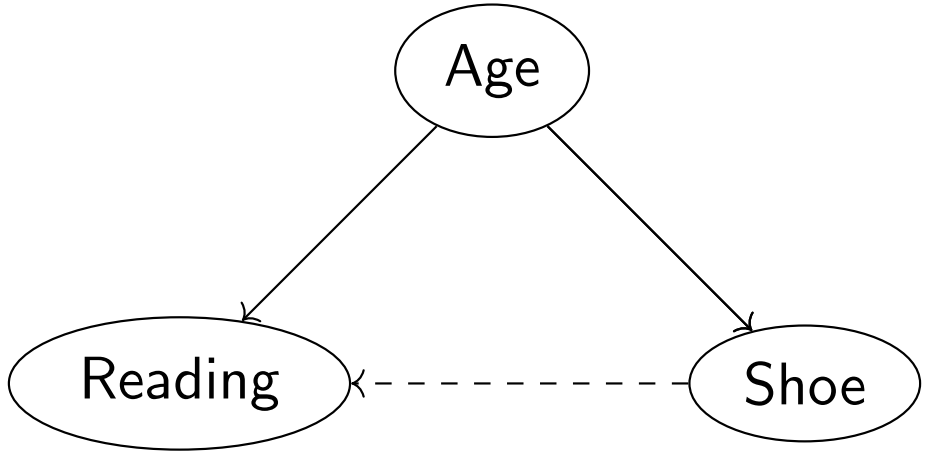

Why are these two factors so highly correlated?

# Causality vs Correlation




- There is a confounding variable 'Age', which we did not take into account.
- The hidden variable confounds the relationship between Shoe Size and Reading.
- Randomized Experiments (Causal Intervention).

# Causal Intervention




- Intervention: $p(R|do(Shoe) = 1) - p(R|do(Shoe) = 0)$.
- Average Treatment Effect:
  $ATE = \sum_A p(R|S = 1, A)p(A) - \sum_A p(R|S = 0, A)p(A)$
- $ATE \approx 0$, which means no causal relation.

Reference: https://www.bradyneal.com/slides/1%20-%20A%20Brief%20Introduction%20to%20Causal%20Inference.pdf

# Revisiting Learning Goals

- Determine whether an independence relationship holds by applying d-separation.
- Given a Bayesian network and an order of the variables, construct a Bayesian network that correctly represents the independence relationships among the variables.
- Understand the difference between correlation and causality.