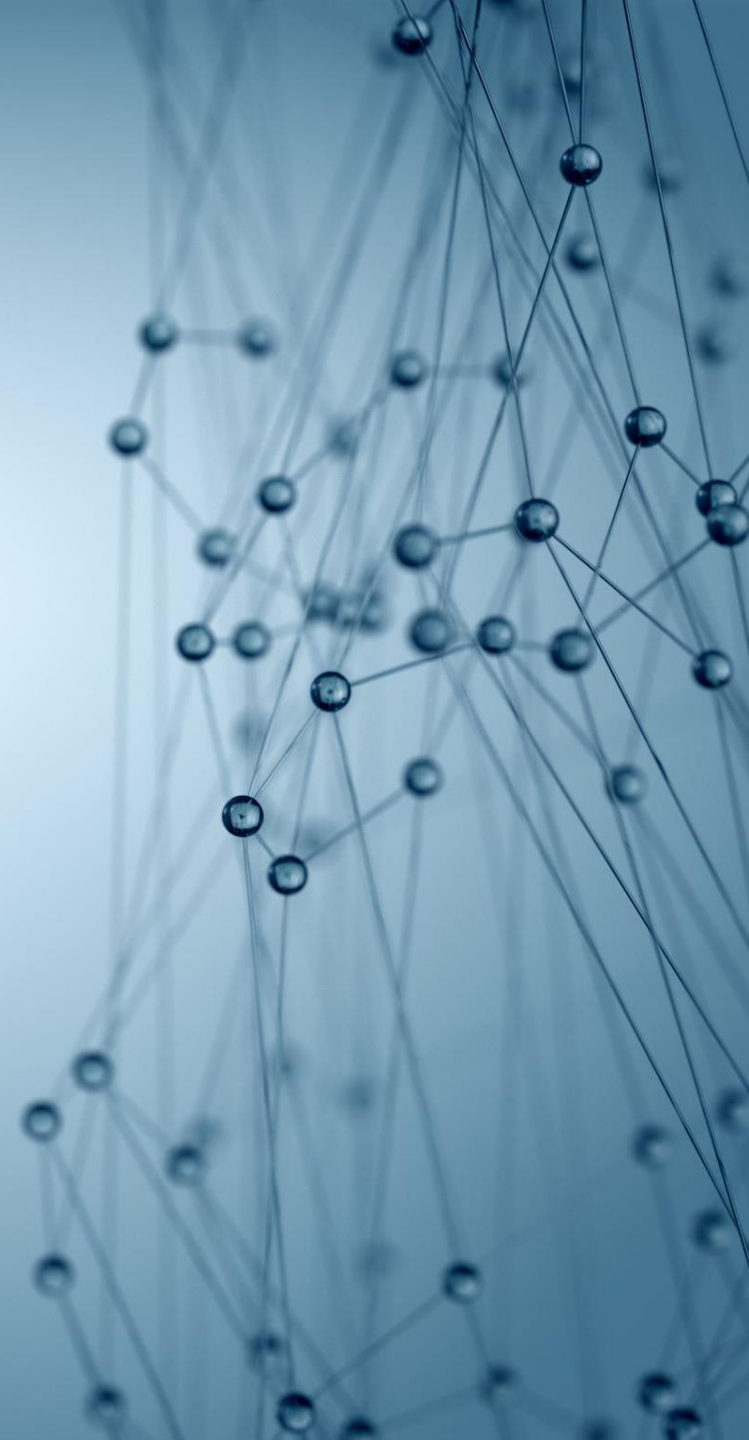


# BAYESIAN NETWORKS

**CS221 Fall 2019**

Dhruv Kedia

Jon Kotker



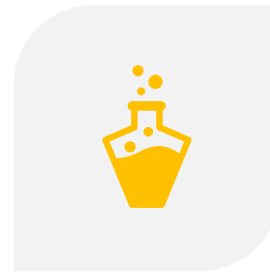
# What are we covering today?



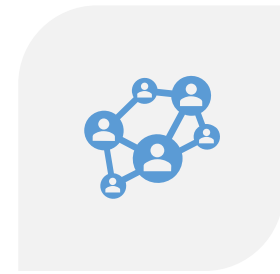
INTRODUCTION TO  
BAYESIAN NETWORKS



PROBABILISTIC  
INFERENCE COOKBOOK

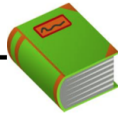
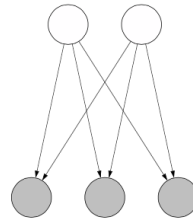


SAMPLE PROBLEMS



CONDITIONAL  
INDEPENDENCE

# Bayesian Networks



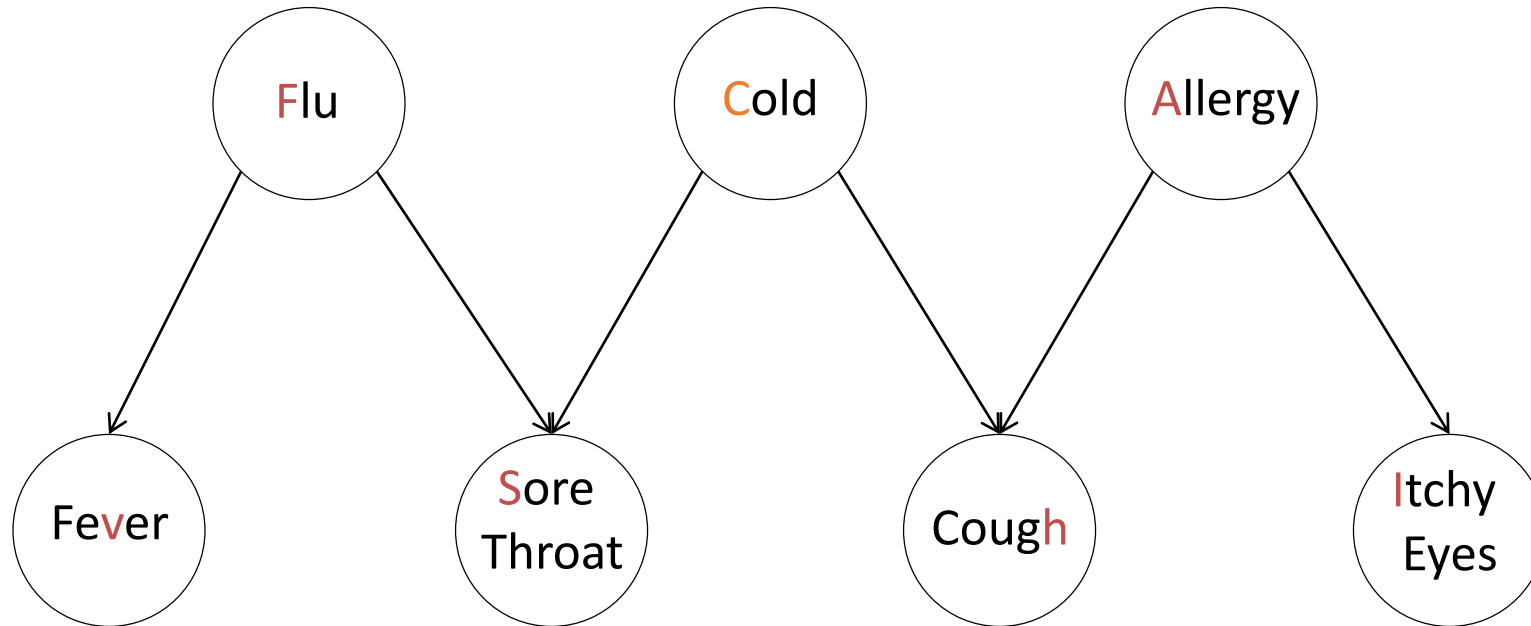
## Definition: Bayesian network

Let  $X = (X_1, \dots, X_n)$  be random variables.

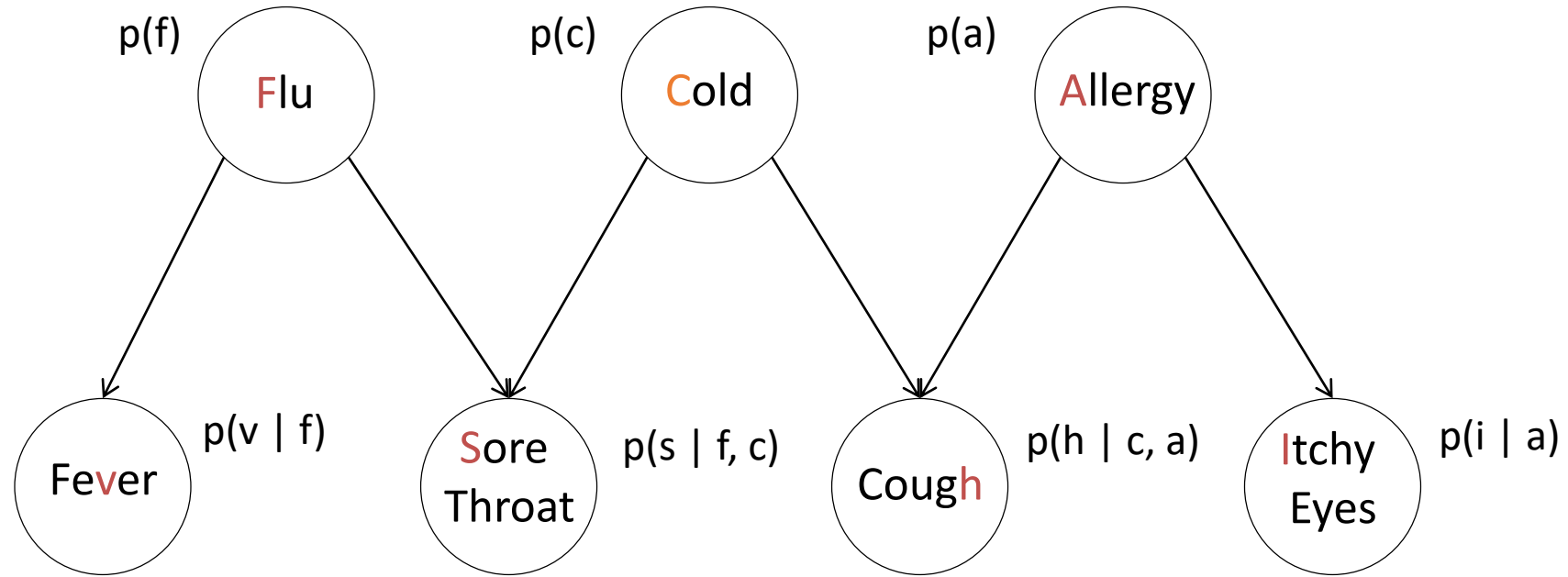
A **Bayesian network** is a directed acyclic graph (DAG) that specifies a **joint distribution** over  $X$  as a product of **local conditional distributions**, one for each node:

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n p(x_i \mid x_{\text{Parents}(i)})$$

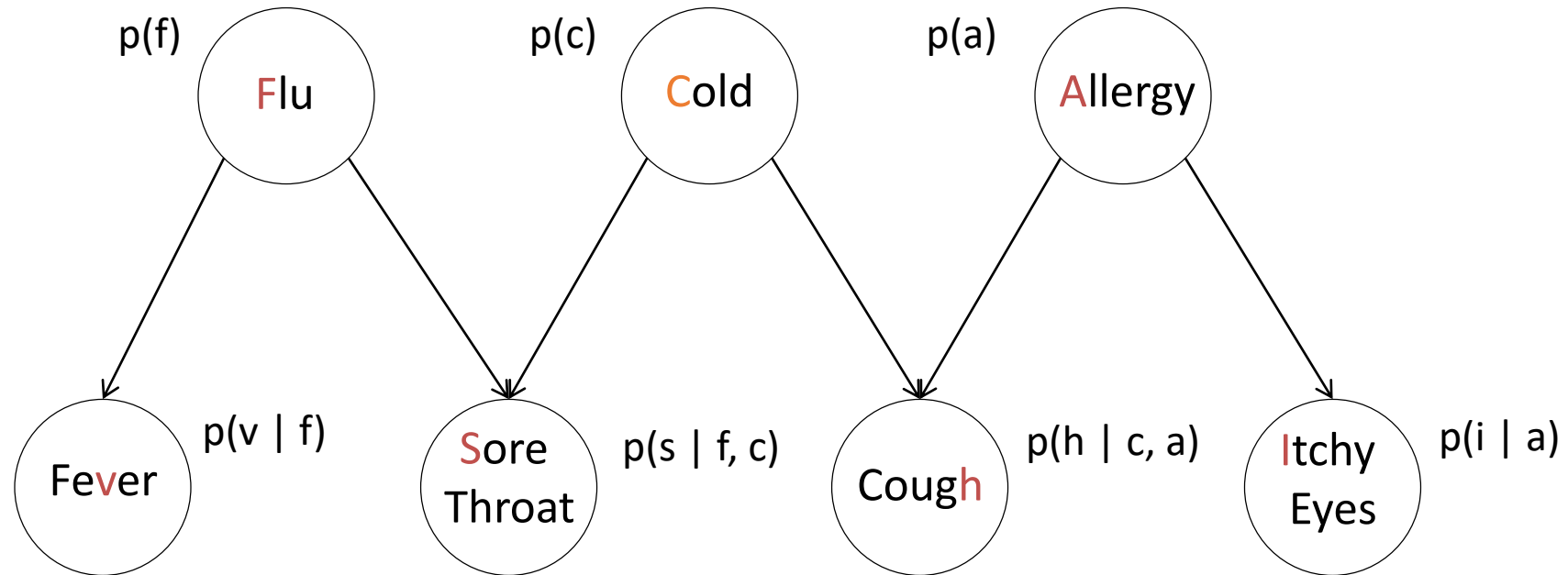
# Bayesian Networks



# A Bayesian network represents a joint probability distribution.



# A Bayesian network represents a joint probability distribution.



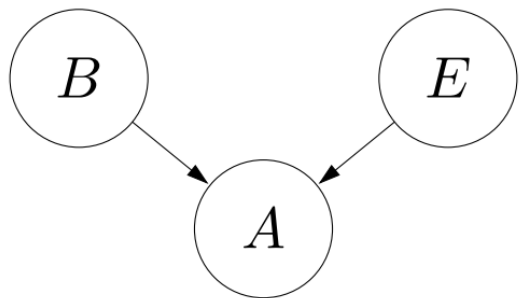
$$P(F=f, C=c, A=a, V=v, S=s, H=h, I=i) = p(f) p(c) p(a) p(v | f) p(s | f, c) p(h | c, a) p(i | a)$$

# Probabilistic Inference Cookbook

Given a query  $P(Q \mid E = e)$ :

1. Remove (marginalize) variables not ancestors of  $Q$  or  $E$ .
2. Convert Bayesian network to factor graph.
3. Condition (shade nodes / disconnect) on  $E = e$ .
4. Remove (marginalize) nodes disconnected from  $Q$ .
5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

# Example: alarm



$b$	$p(b)$
1	$\epsilon$
0	$1 - \epsilon$

$e$	$p(e)$
1	$\epsilon$
0	$1 - \epsilon$

$b$	$e$	$a$	$p(a   b, e)$
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

[whiteboard]

Query:  $\mathbb{P}(B)$

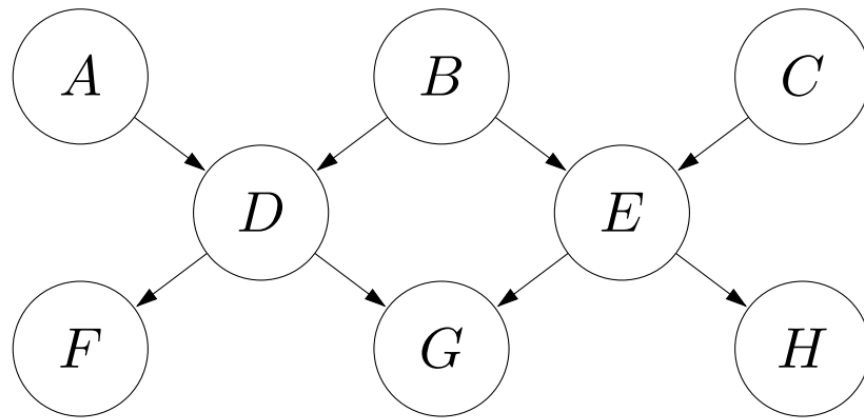
- Marginalize out  $A, E$

Query:  $\mathbb{P}(B | A = 1)$

- Condition on  $A = 1$



## Example: A-H (section)



[whiteboard]

Query:  $\mathbb{P}(C \mid B = b)$

- Marginalize out everything else, note  $C \perp\!\!\!\perp B$

Query:  $\mathbb{P}(C, H \mid E = e)$

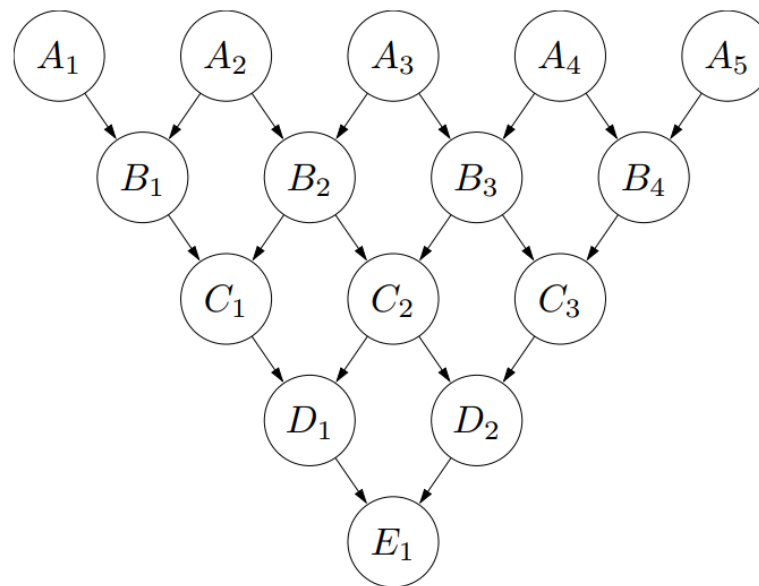
- Marginalize out  $A, D, F, G$ , note  $C \perp\!\!\!\perp H \mid E$

# Bayesian Lights

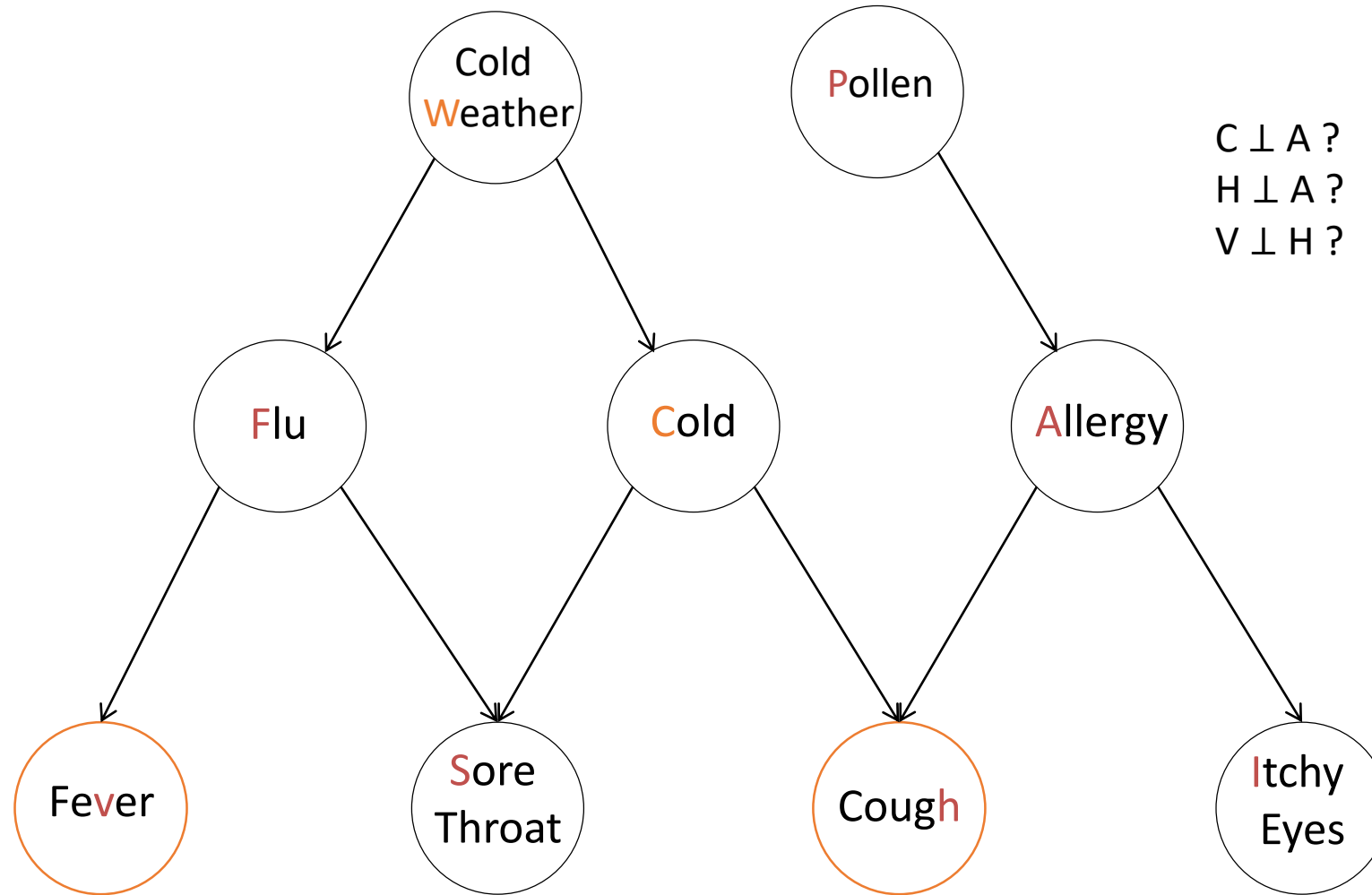
(Fall 2016 Midterm)

This holiday season, you decide to put your knowledge of Bayesian networks to good use. You decide to create Bayesian Lights<sup>TM</sup>, an arrangement of lights that turn on and off randomly according to the joint distribution of a Bayesian network.

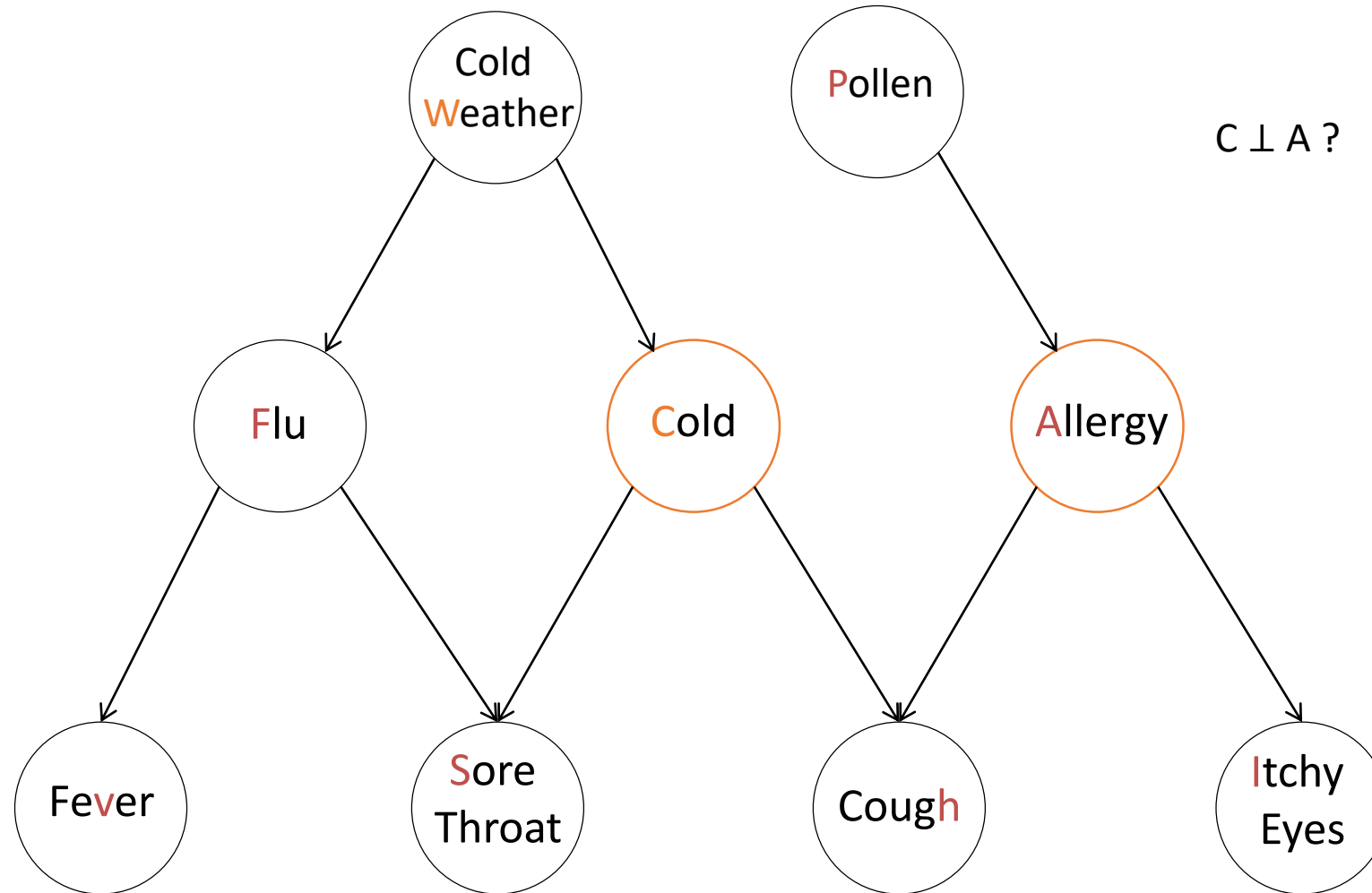
Figure 1 shows the Bayesian network corresponding to the lights. Each light is associated with a variable which takes on values in  $\{0, 1\}$  (off: 0, on: 1). For example,  $A_1$  is the light in the upper-left corner and  $E_1$  is the light at the very bottom. A light in the top row is on with probability  $\alpha$ . The status of a light in subsequent rows is governed by the two parent lights directly above it, and it is on with probability  $\beta$  when the two parent lights above it have different statuses (on-off or off-on), and off with probability  $\beta$  when the two parent lights above it have the same status. In other words, each light is the result of applying a noisy XOR function.



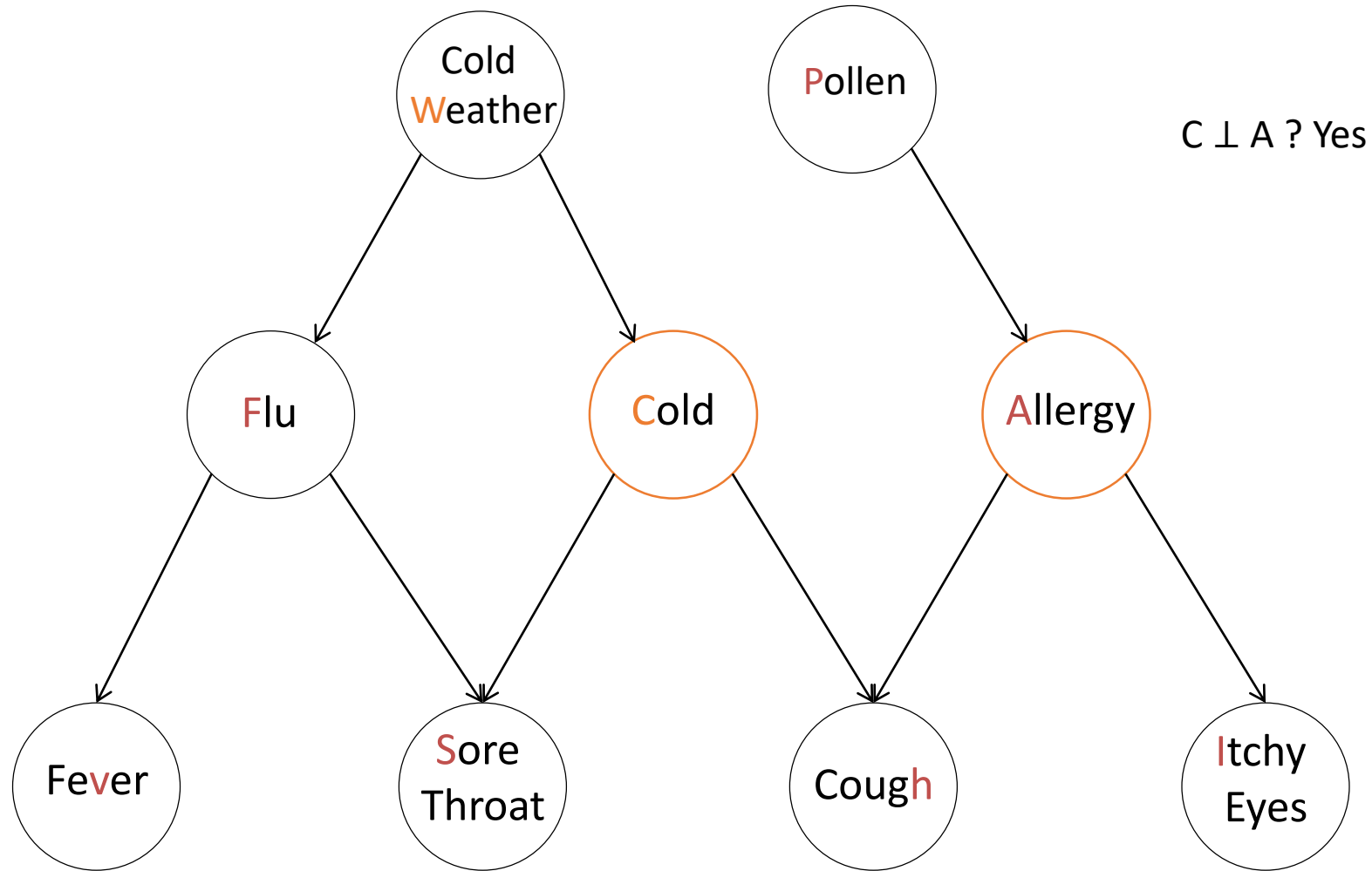
# Conditional Independence



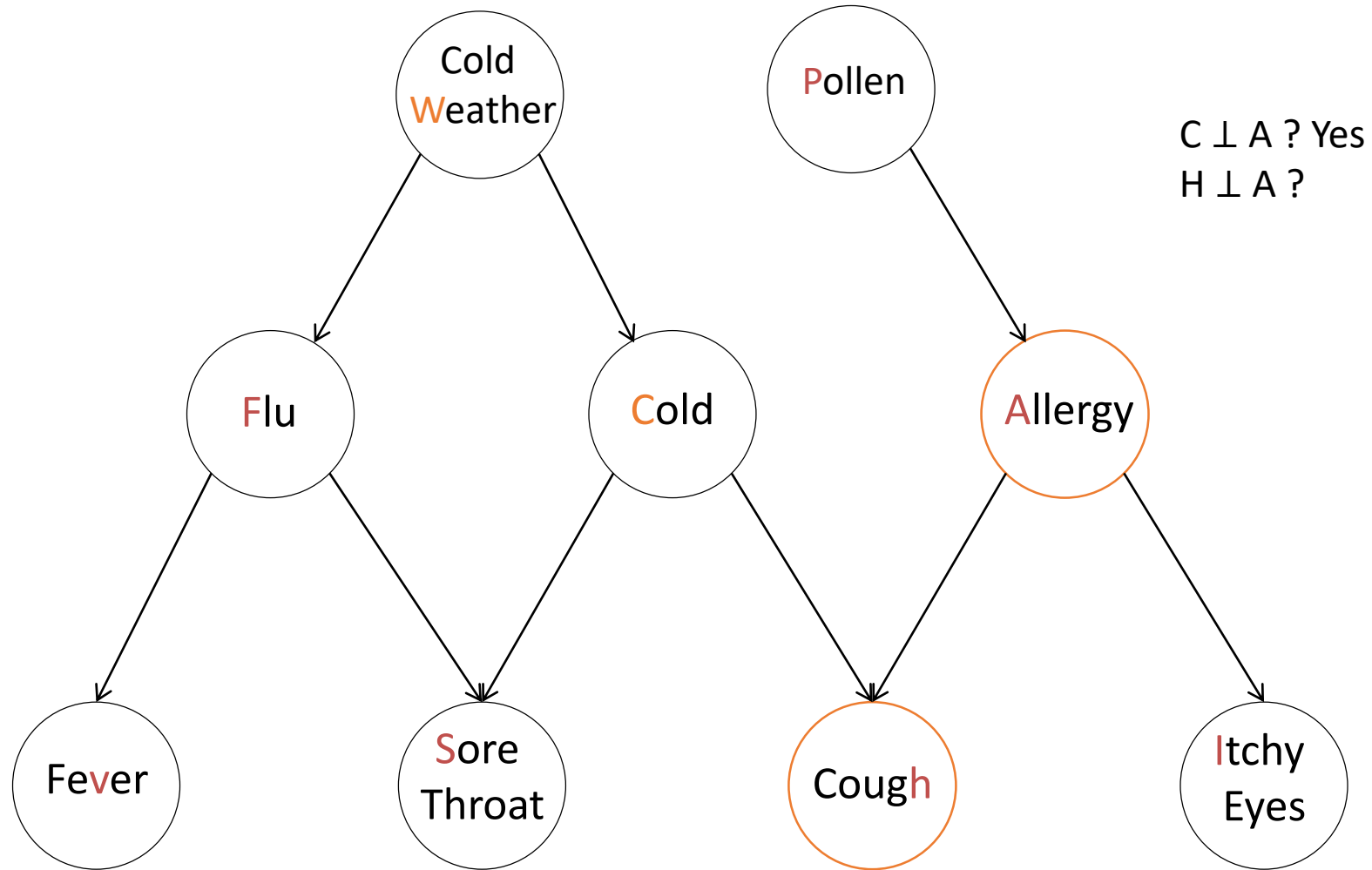
# Conditional Independence



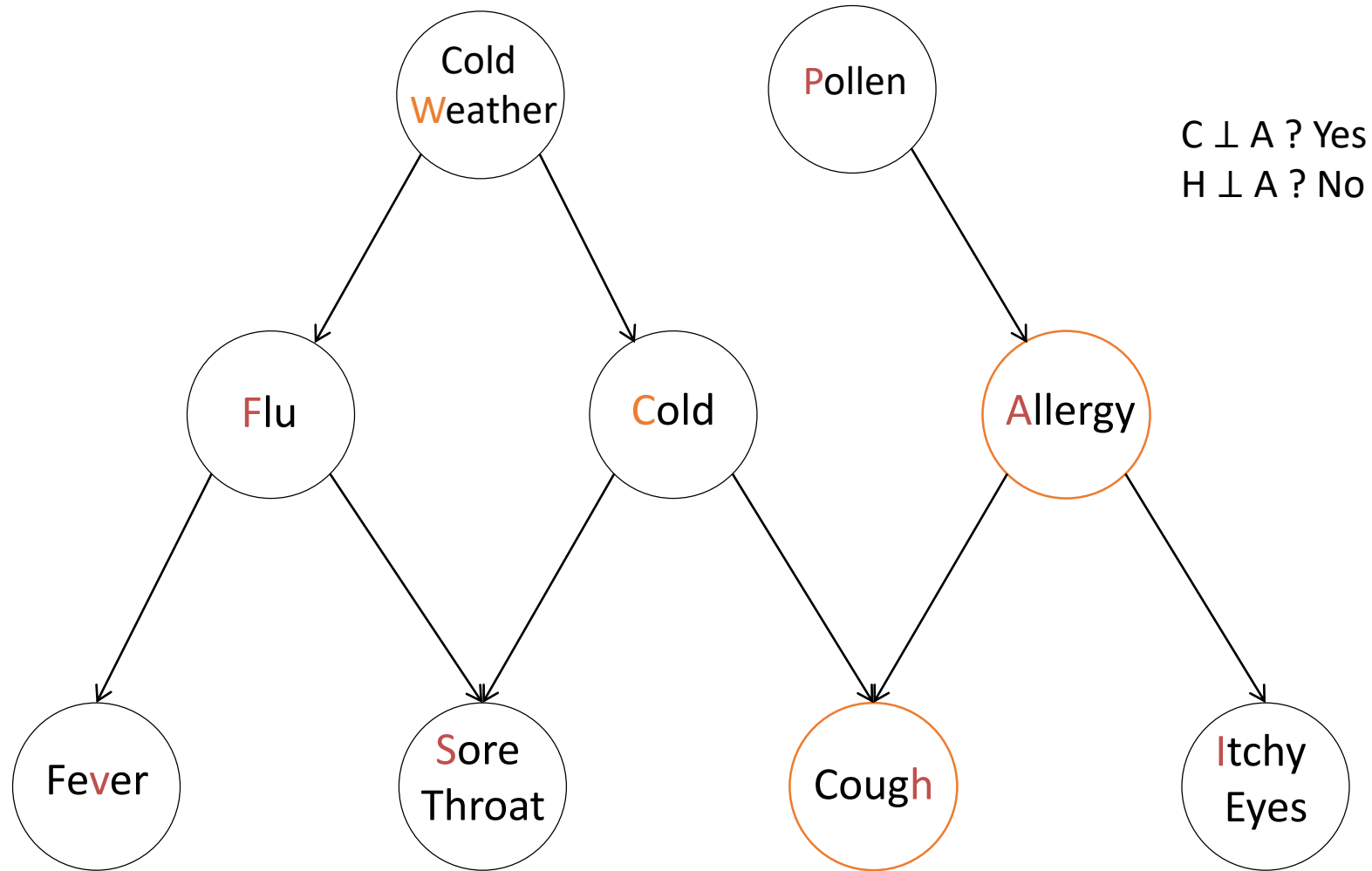
# Conditional Independence



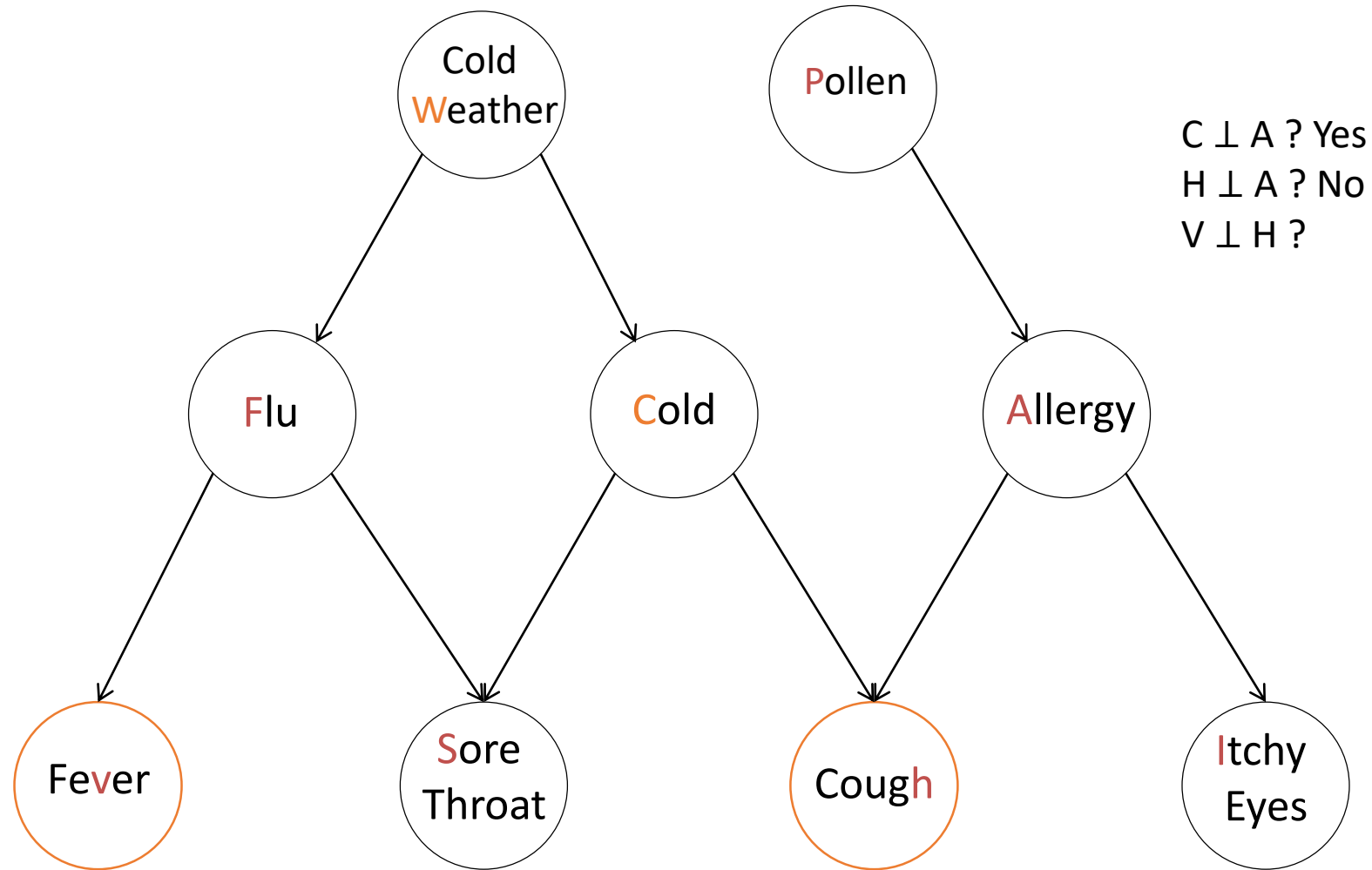
# Conditional Independence



# Conditional Independence

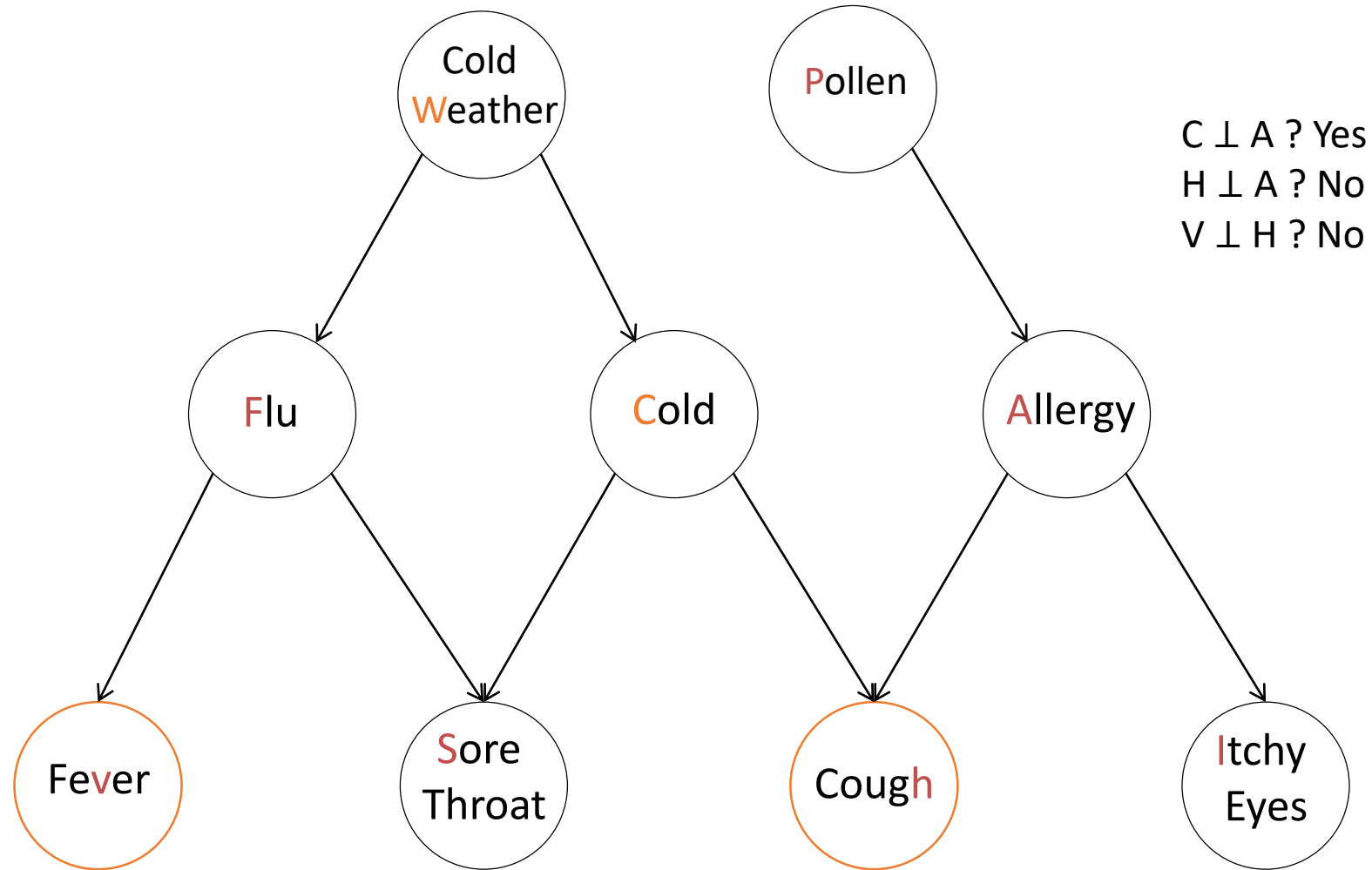


# Conditional Independence

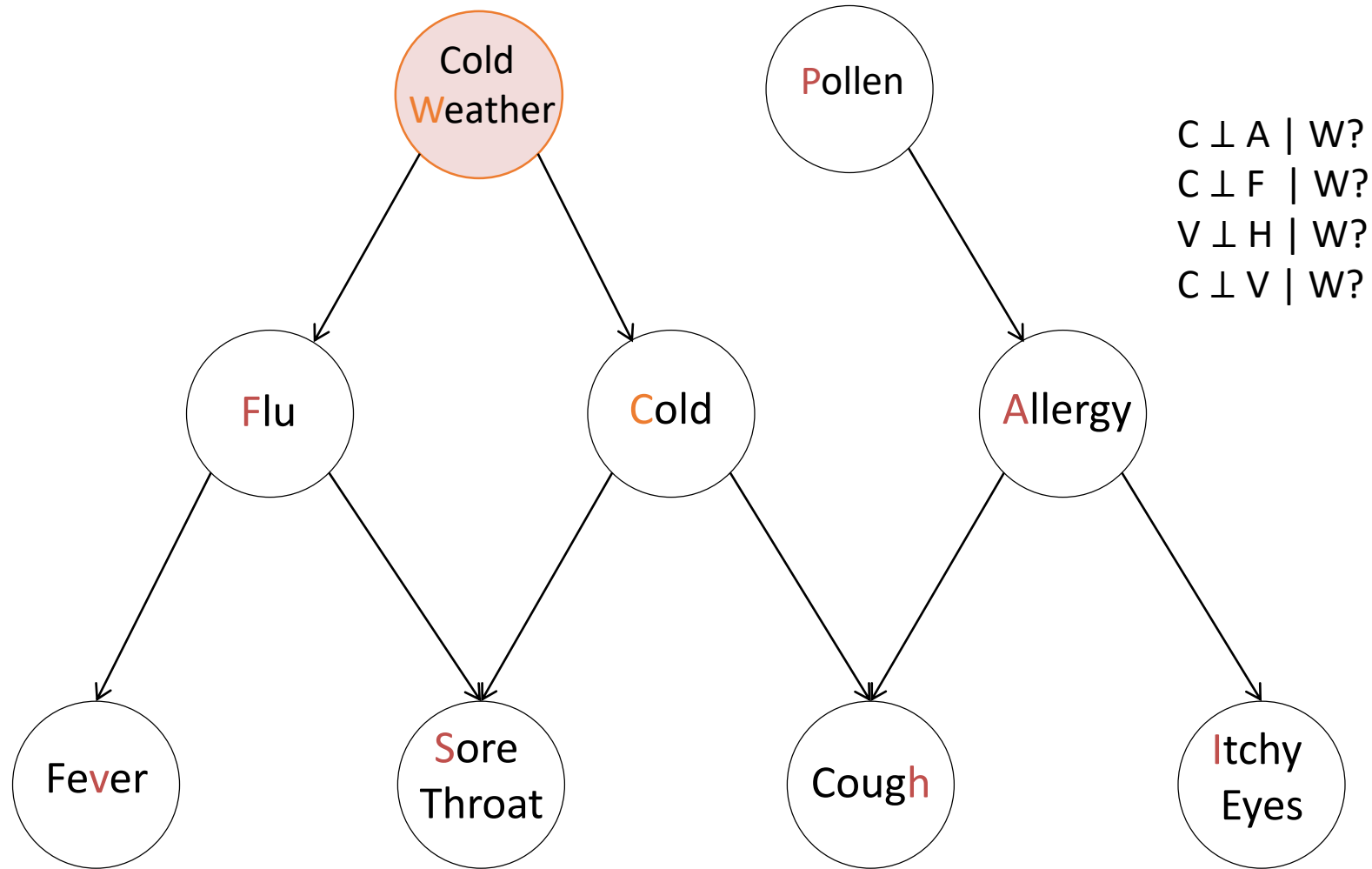




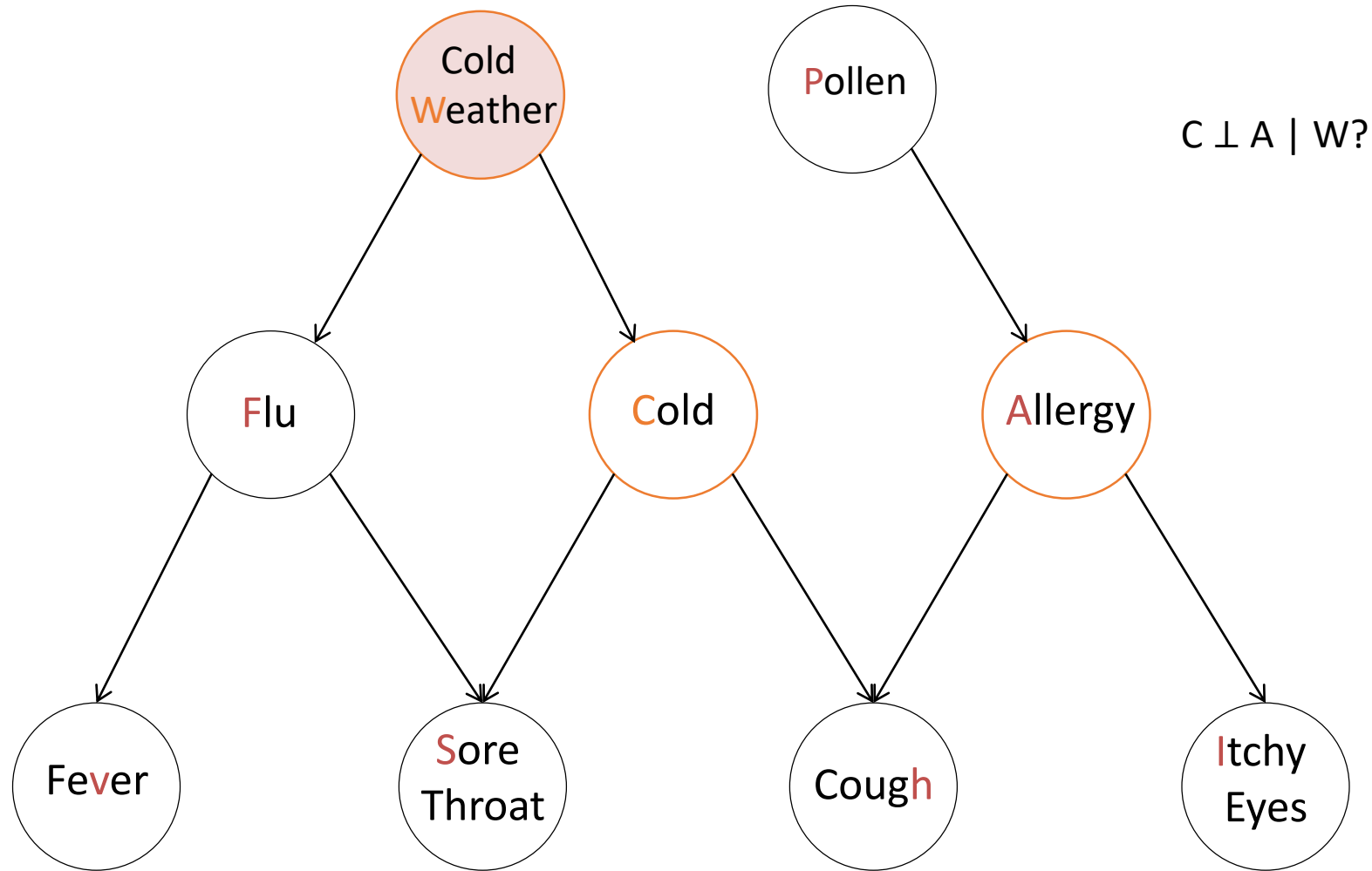
# Conditional Independence



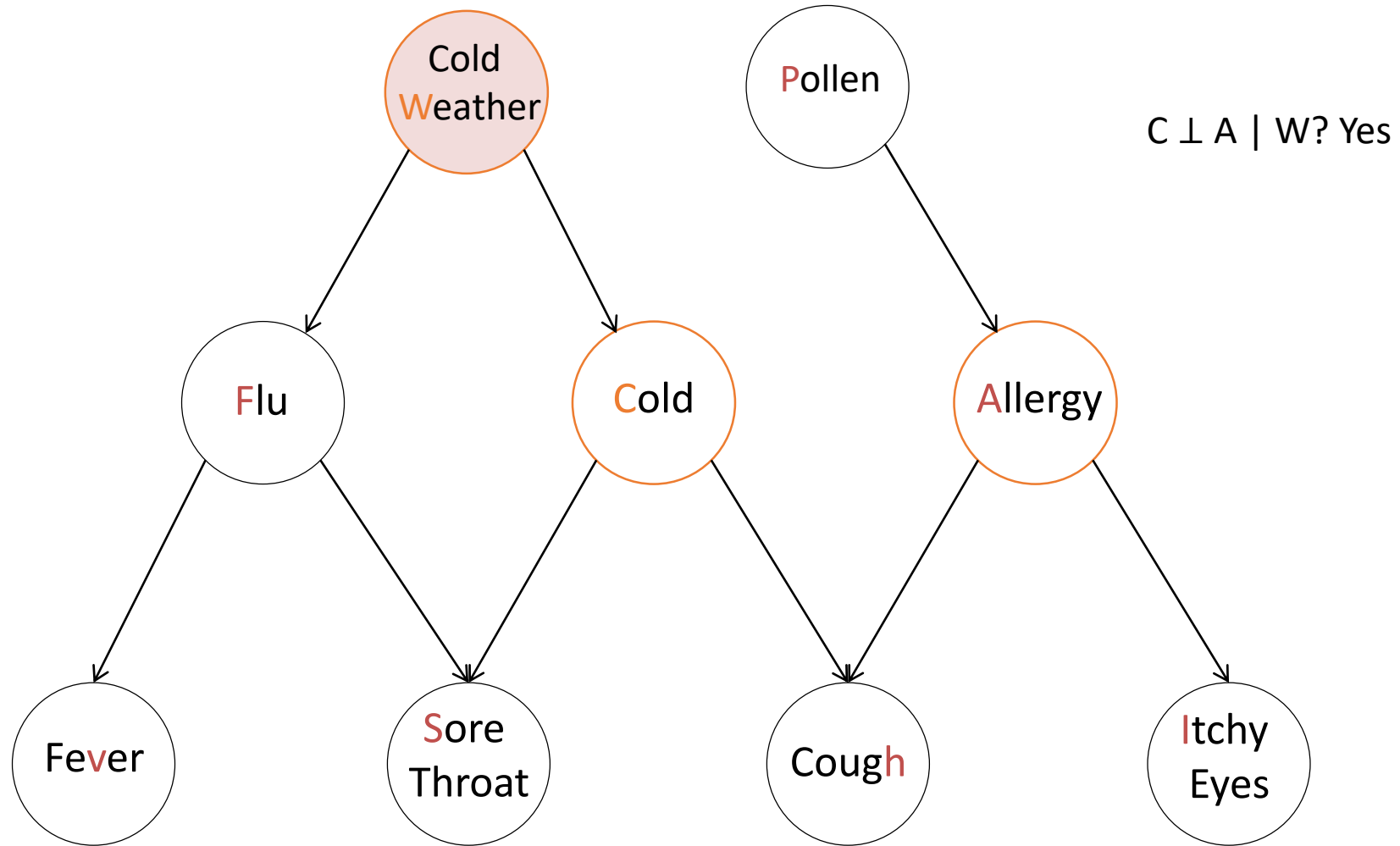
# Conditional Independence



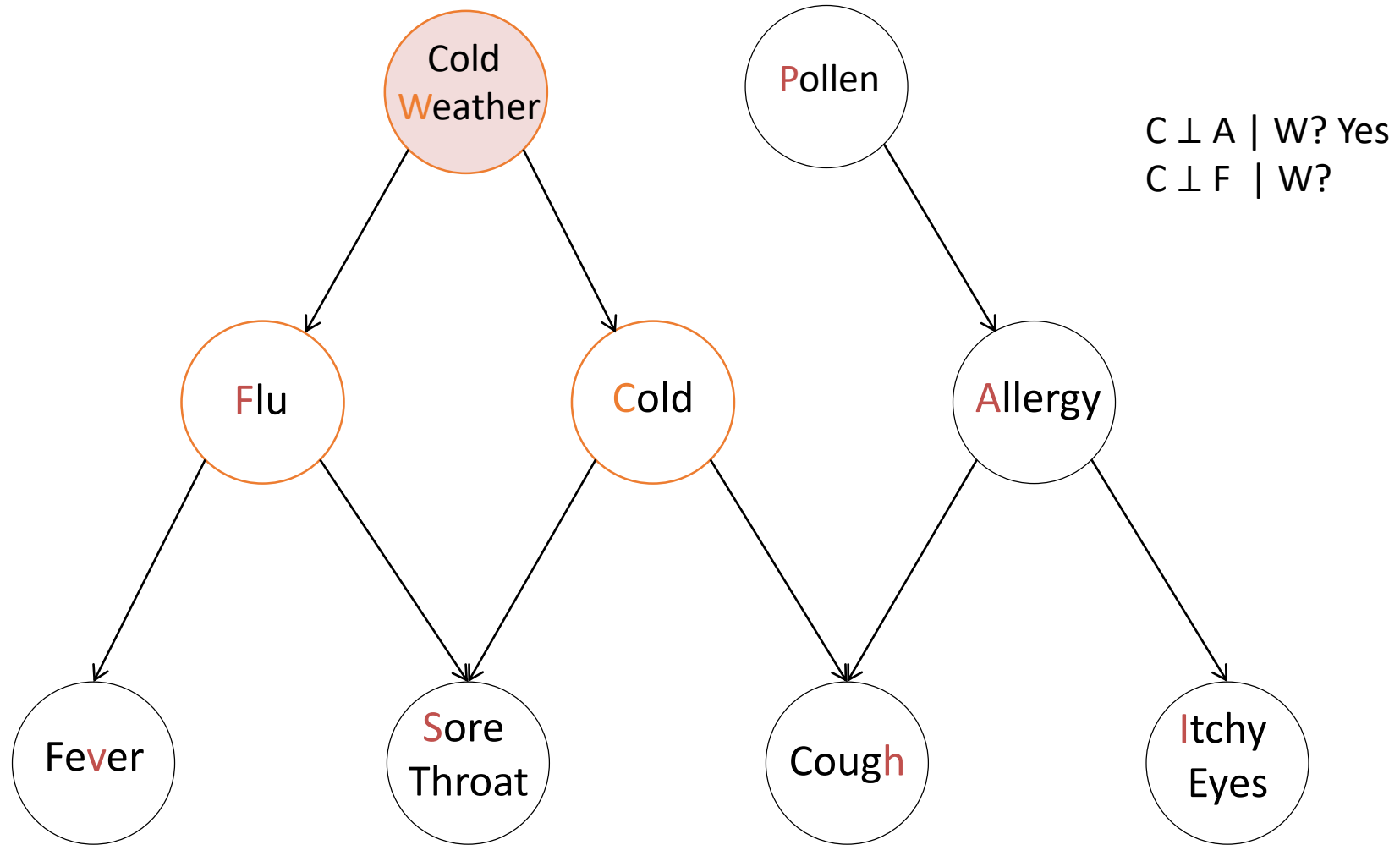
# Conditional Independence



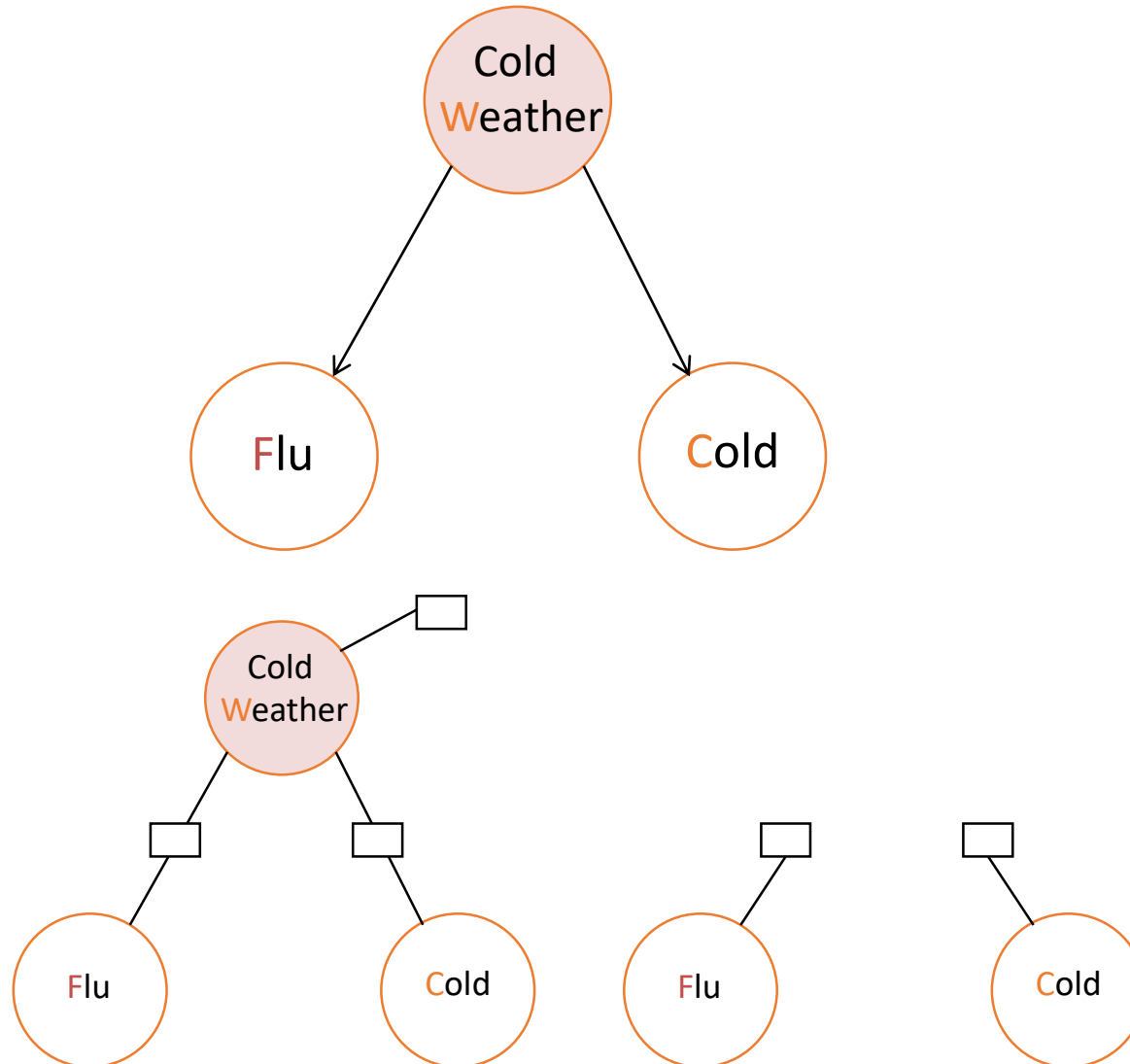
# Conditional Independence



# Conditional Independence

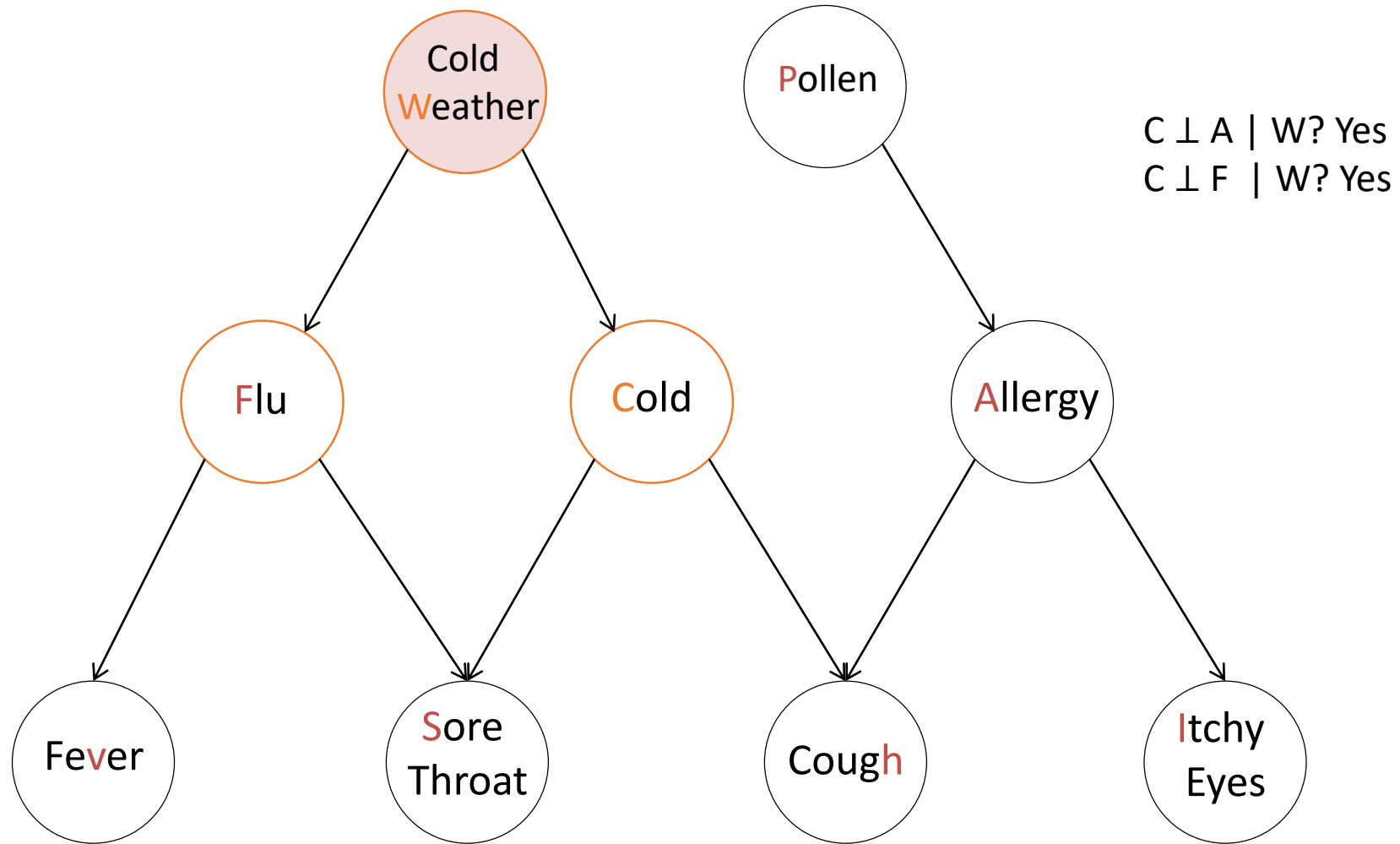


# Conditional Independence

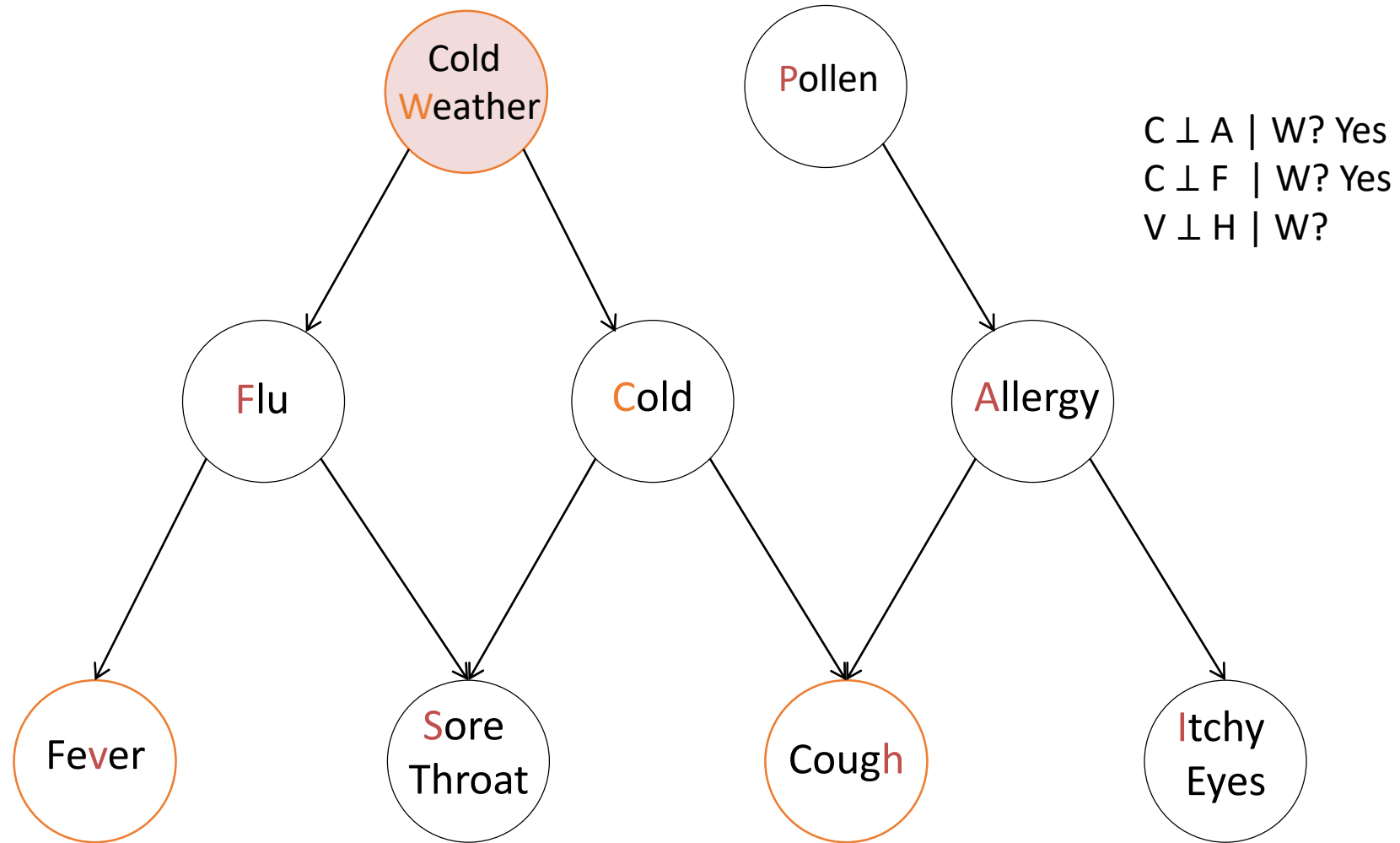


$C \perp A \mid W?$  Yes  
 $C \perp F \mid W?$

# Conditional Independence

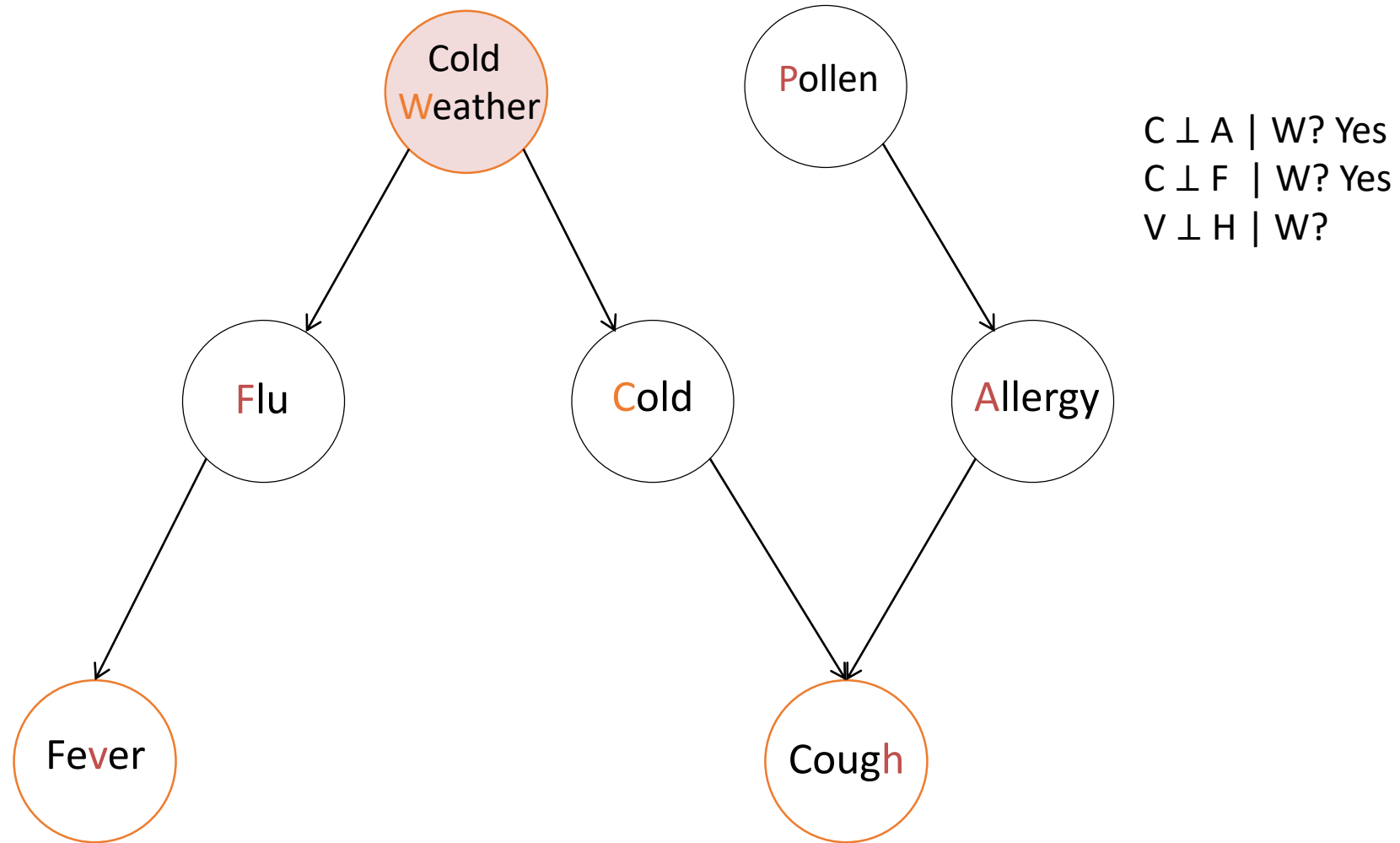


# Conditional Independence

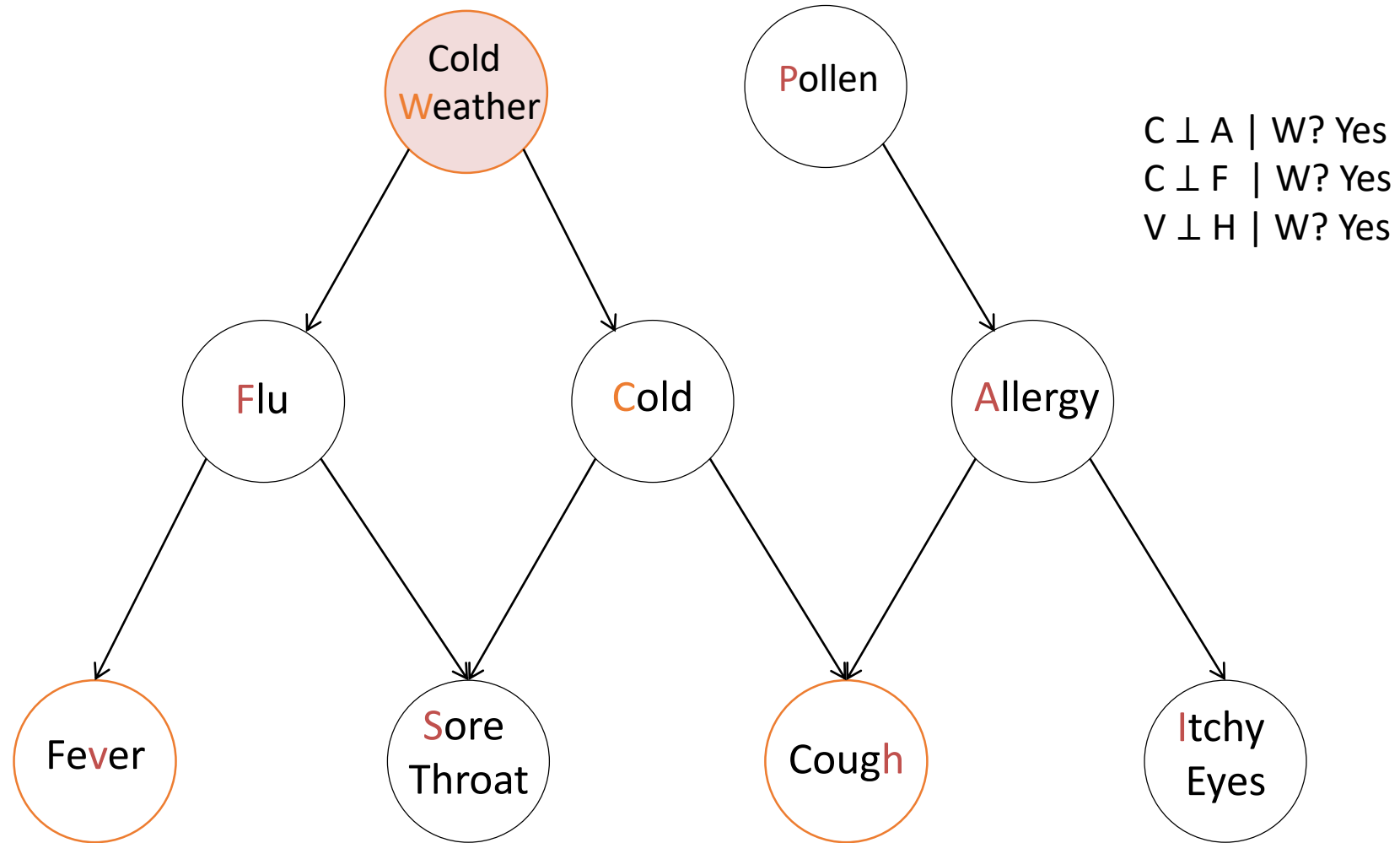




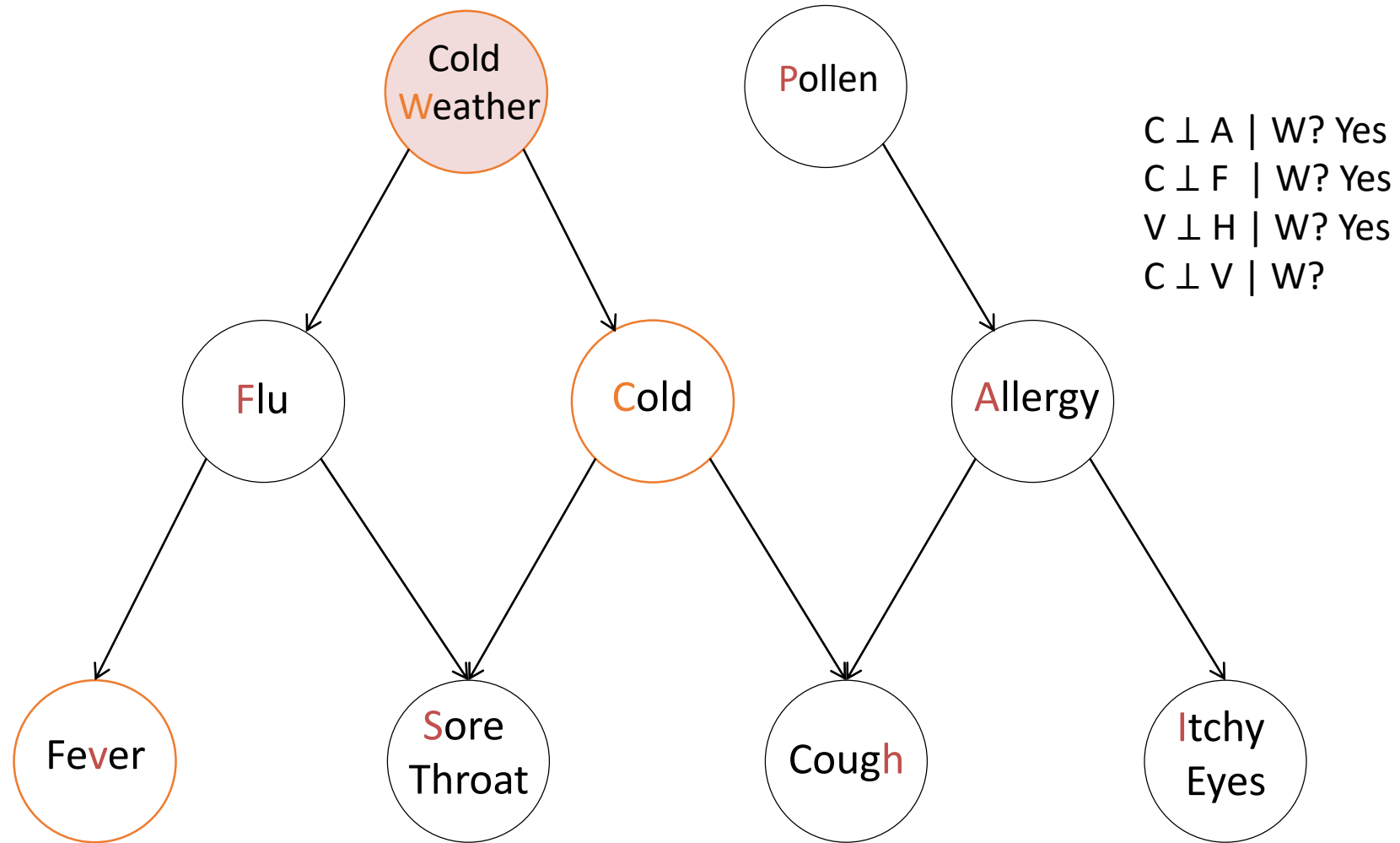
# Conditional Independence



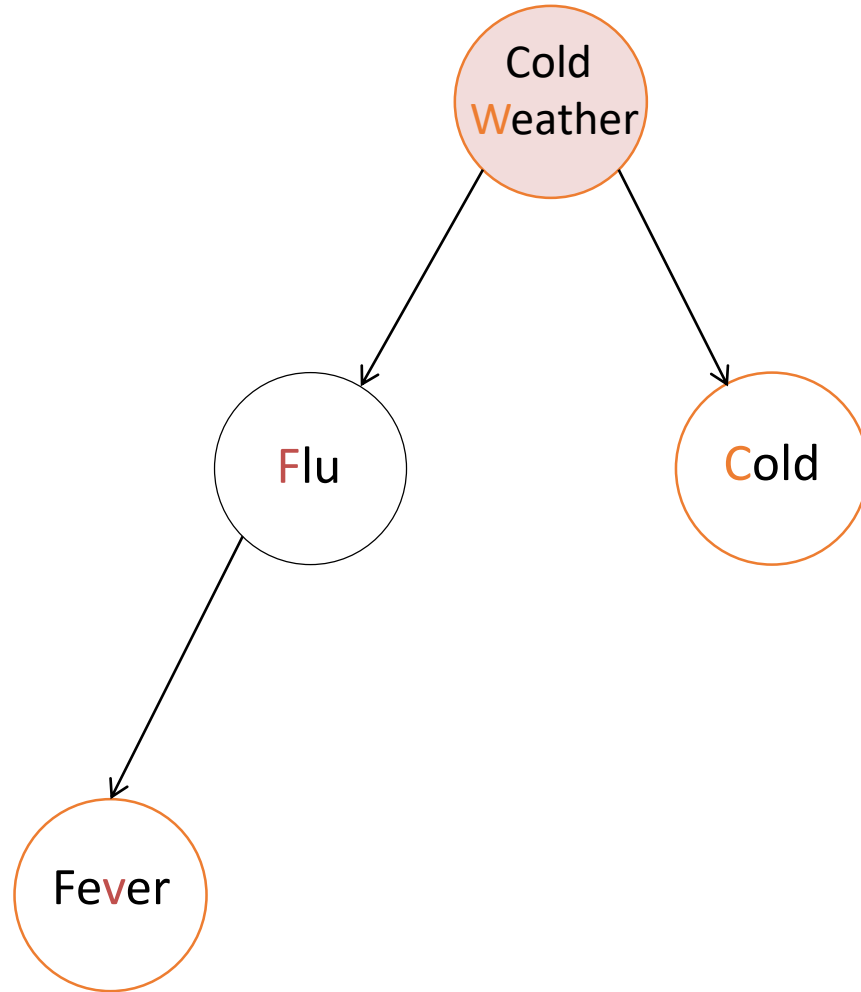
# Conditional Independence



# Conditional Independence



# Conditional Independence



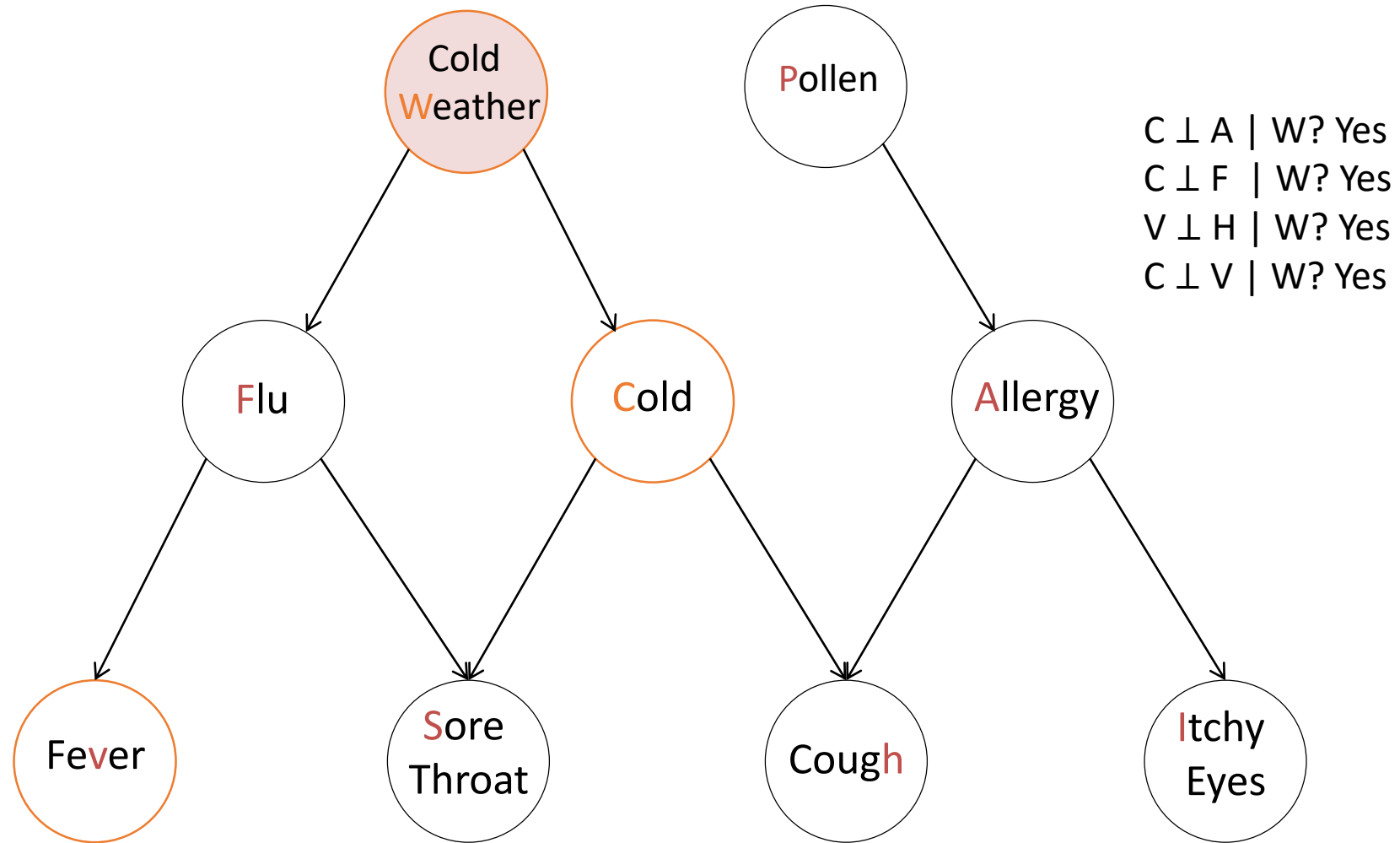
$C \perp A \mid W?$  Yes

$C \perp F \mid W?$  Yes

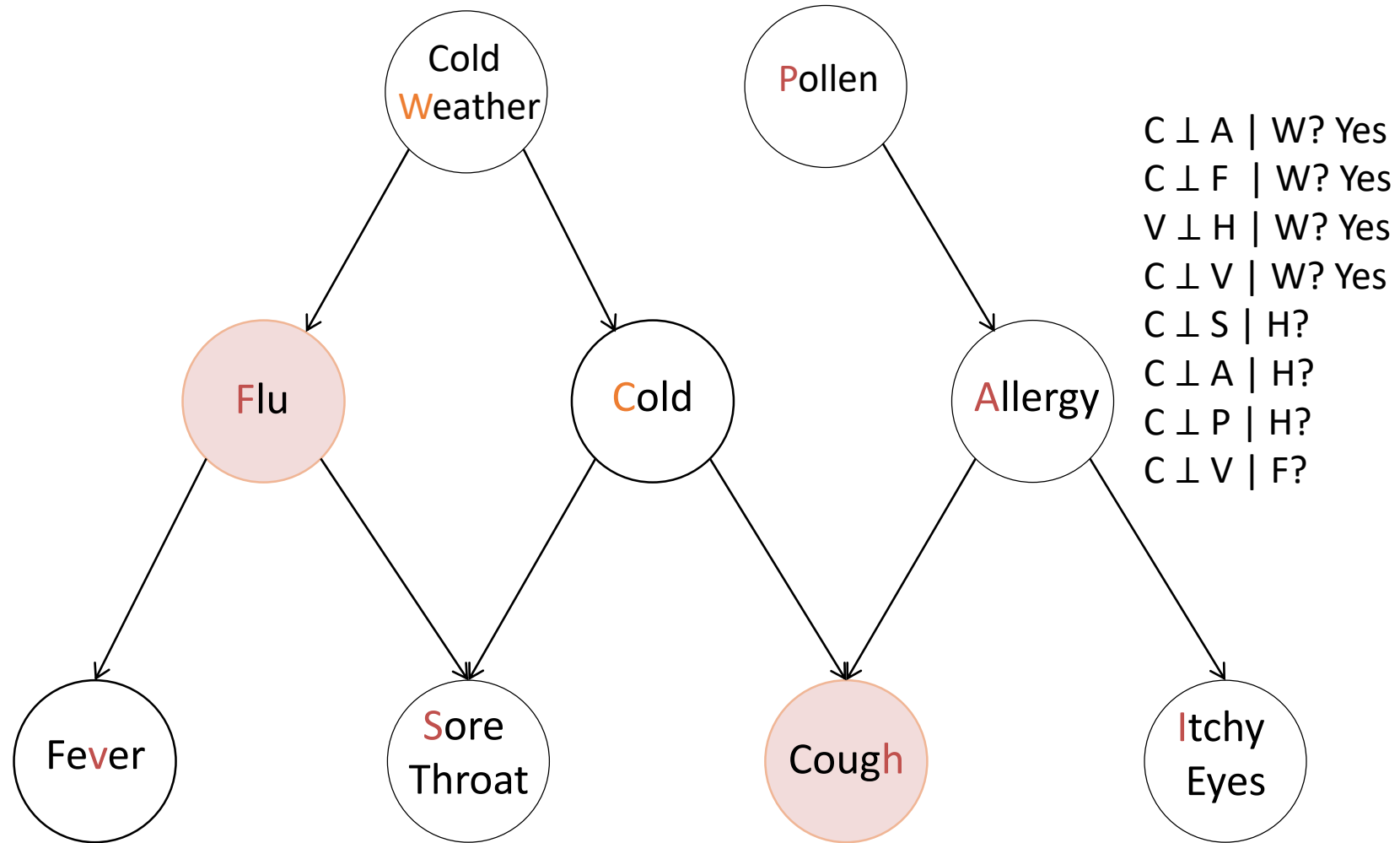
$V \perp H \mid W?$  Yes

$C \perp V \mid W?$

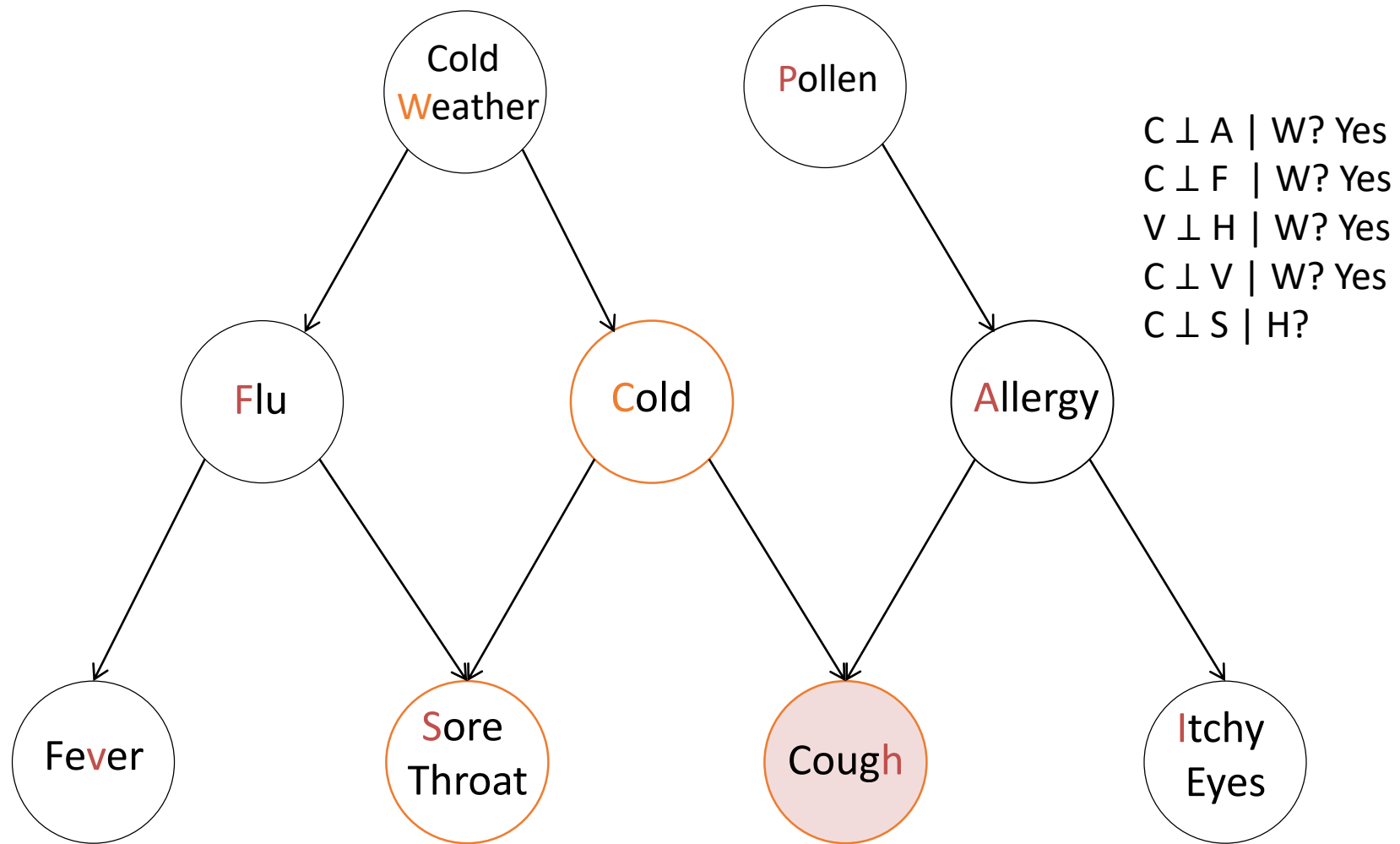
# Conditional Independence



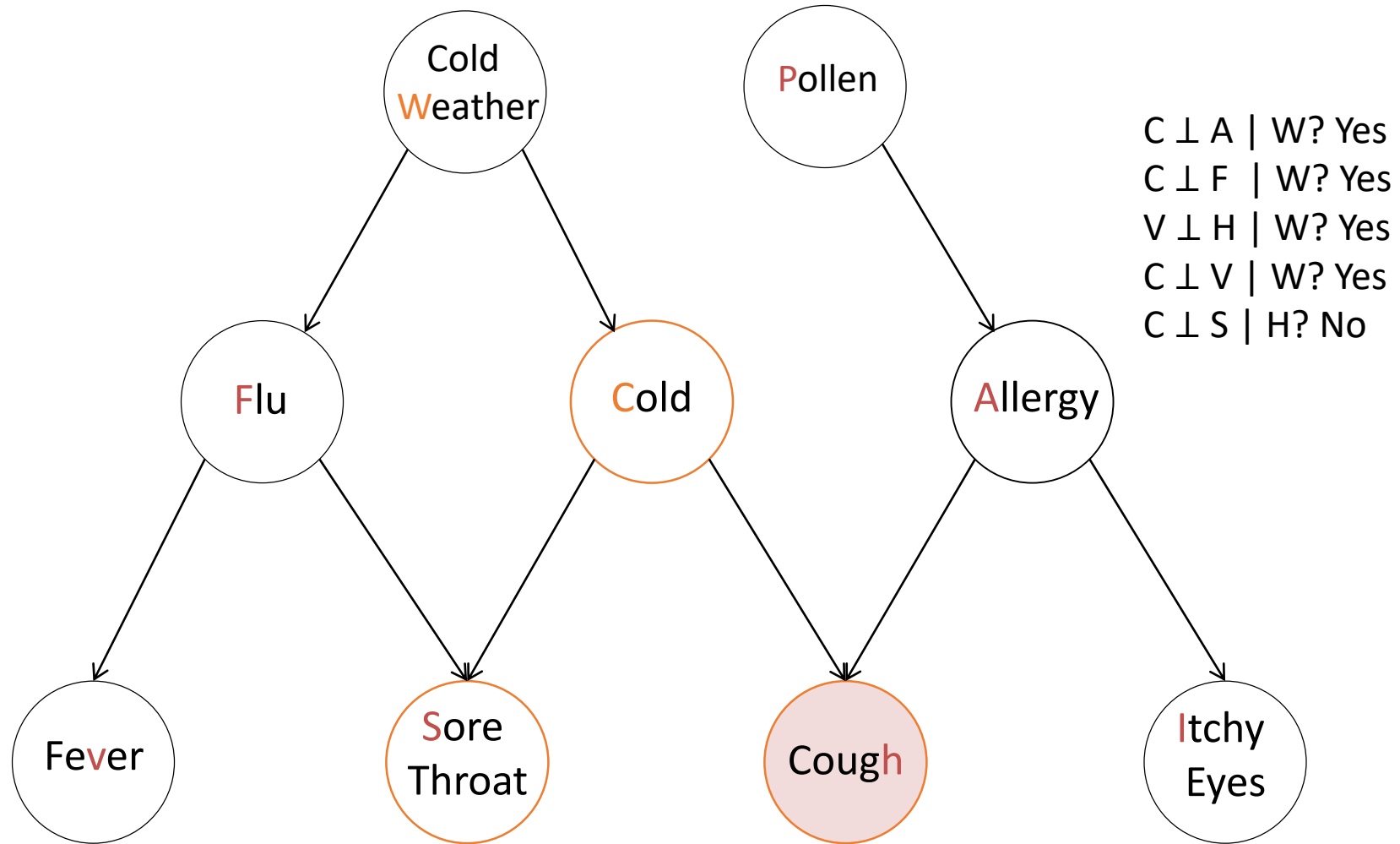
# Conditional Independence



# Conditional Independence

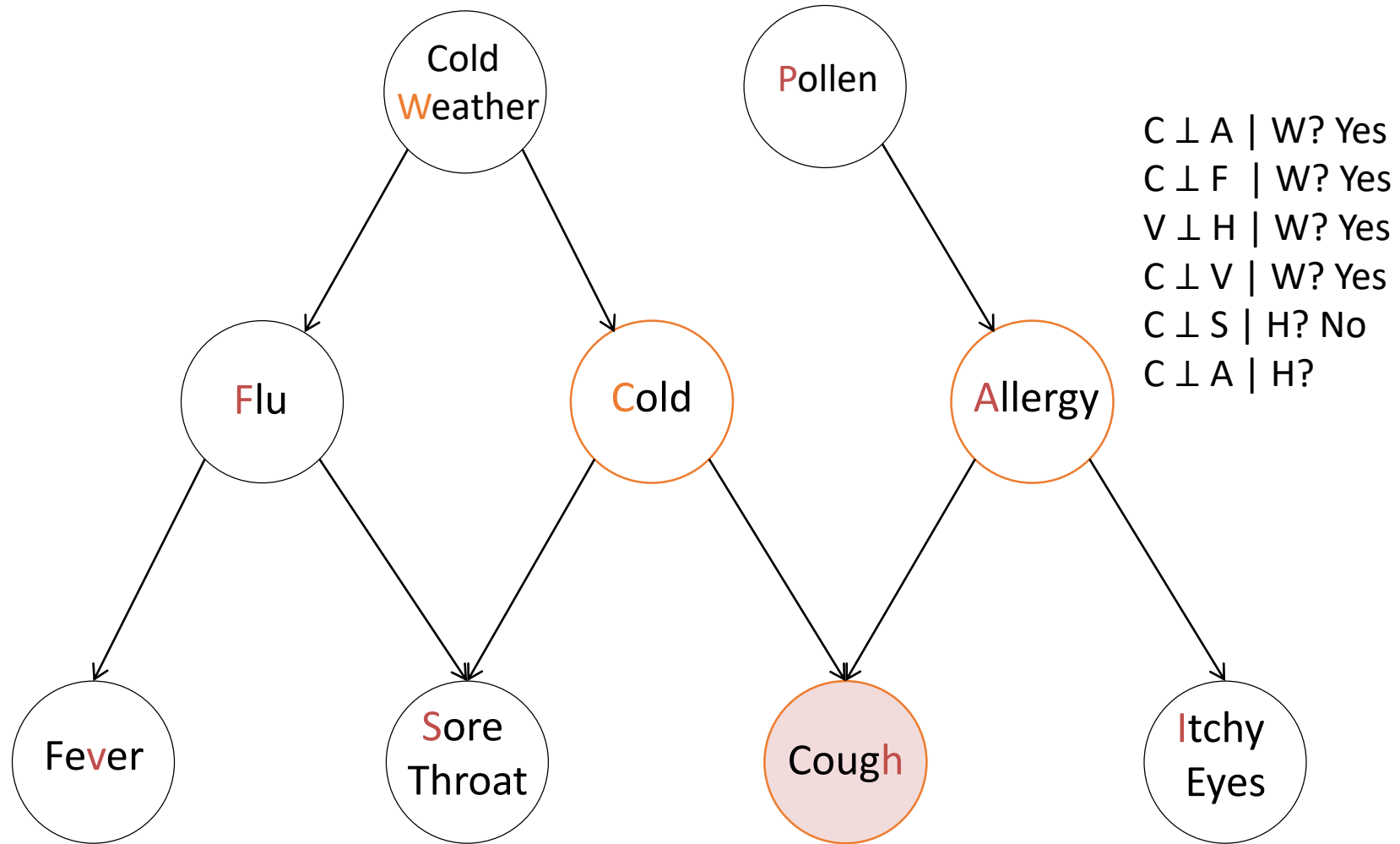


# Conditional Independence

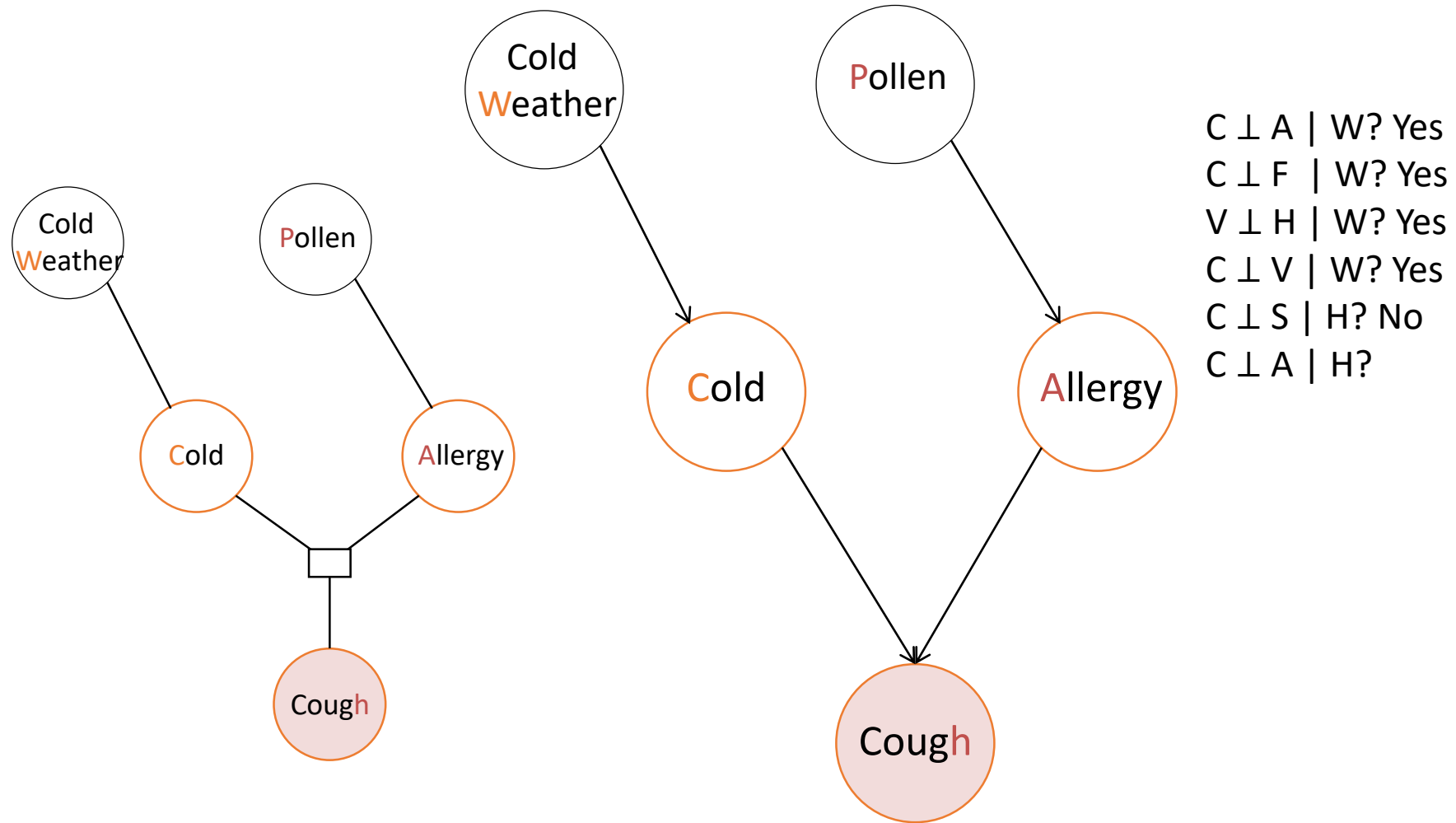




# Conditional Independence

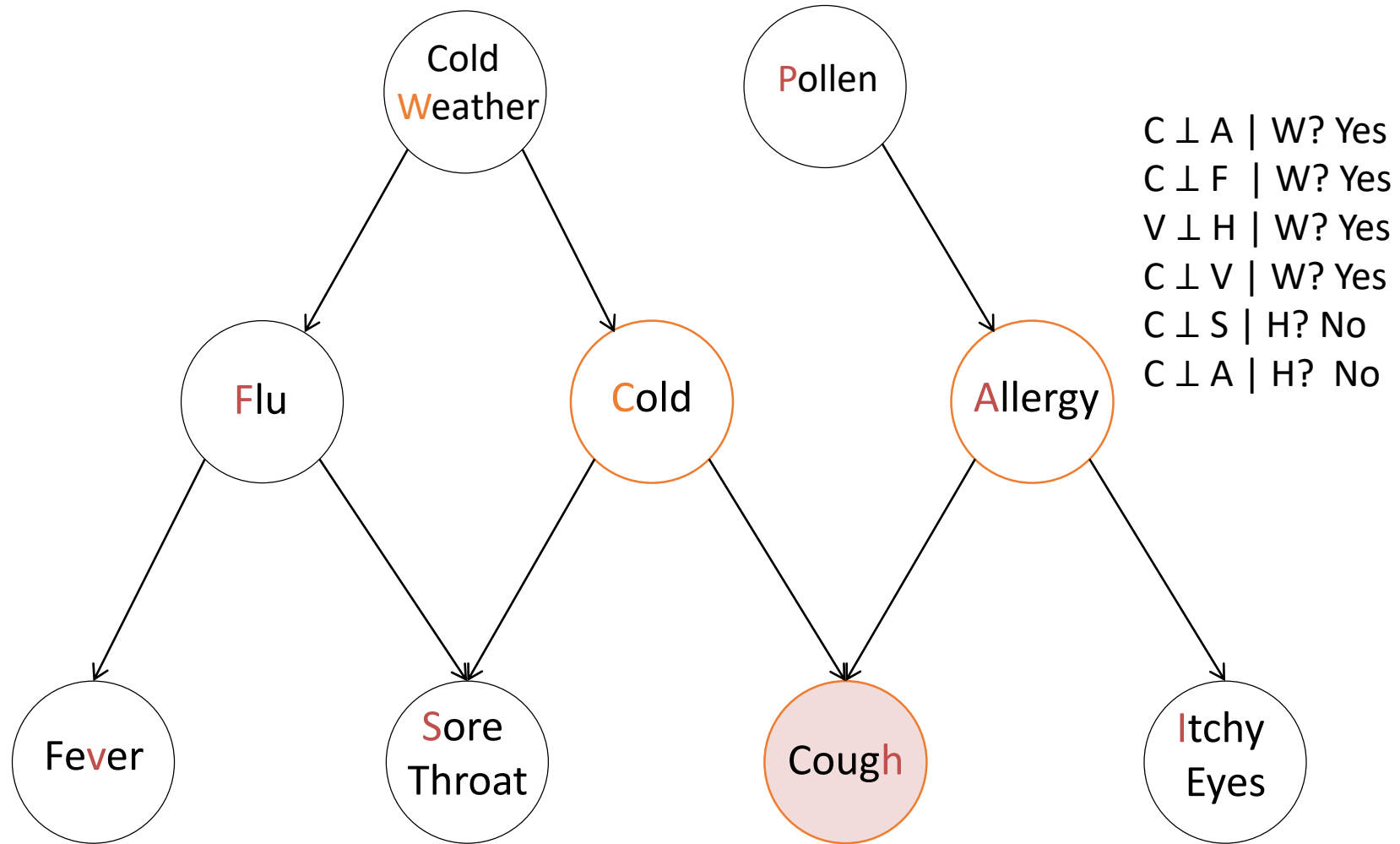


# Conditional Independence

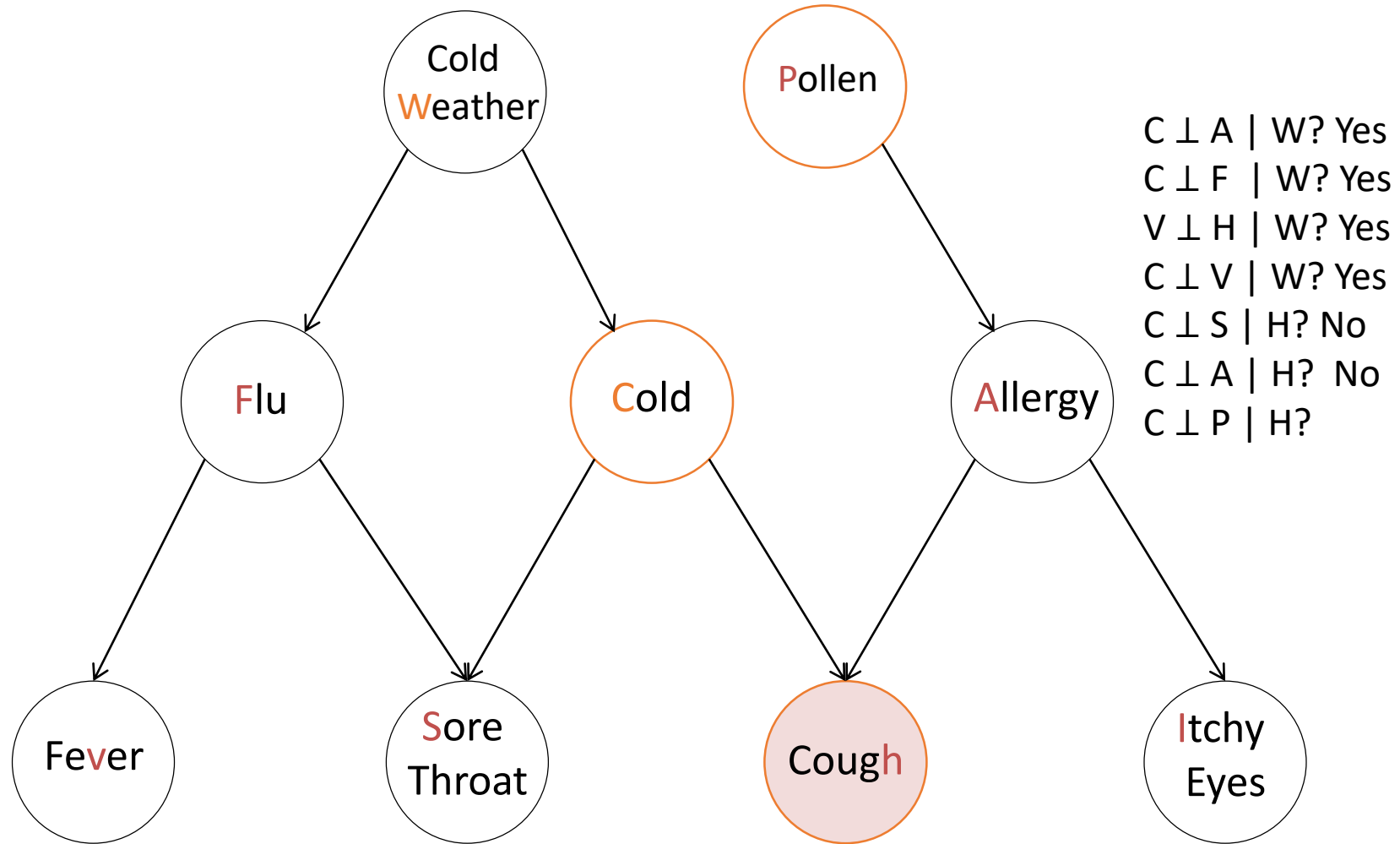


Explaining Away!

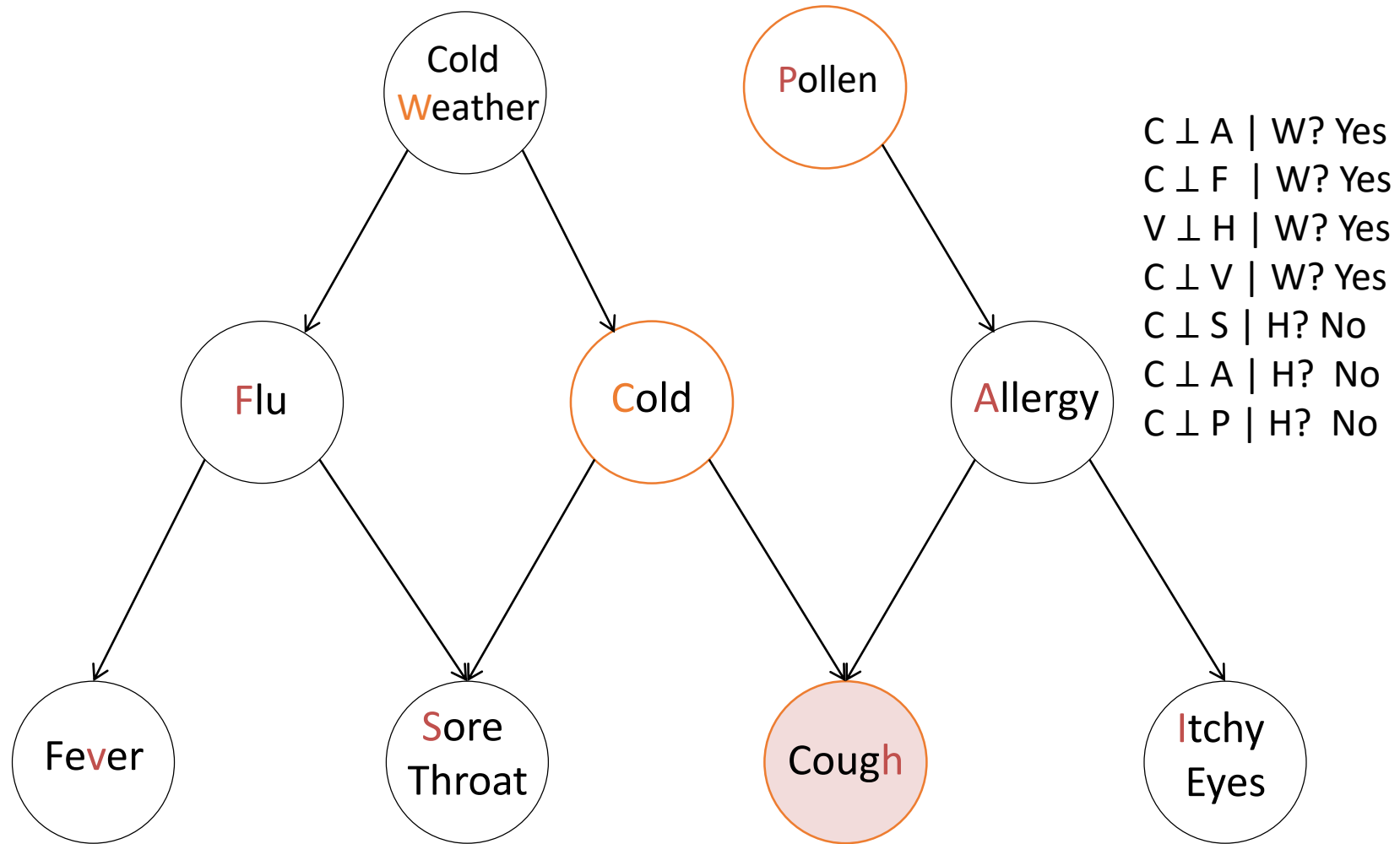
# Conditional Independence



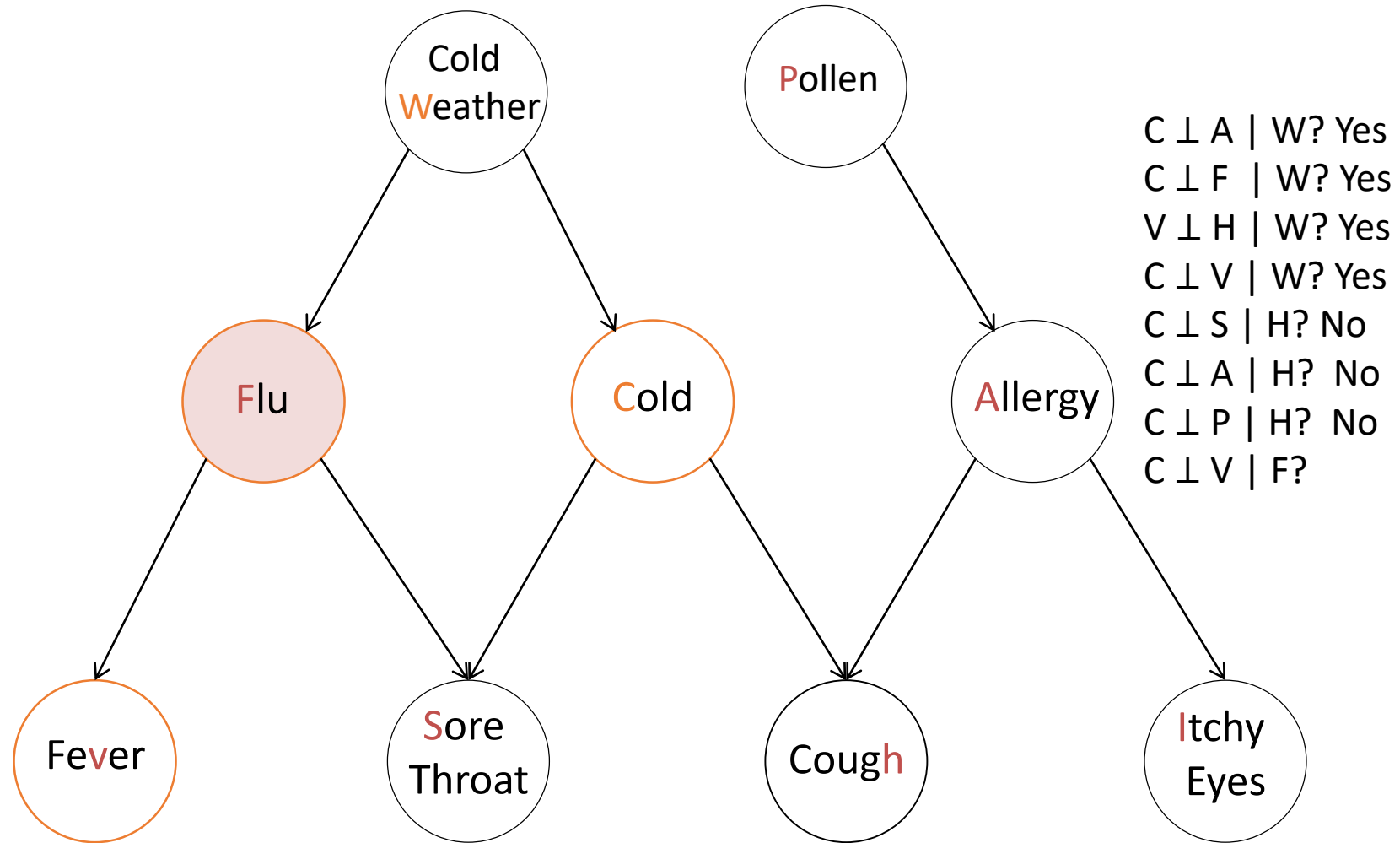
# Conditional Independence



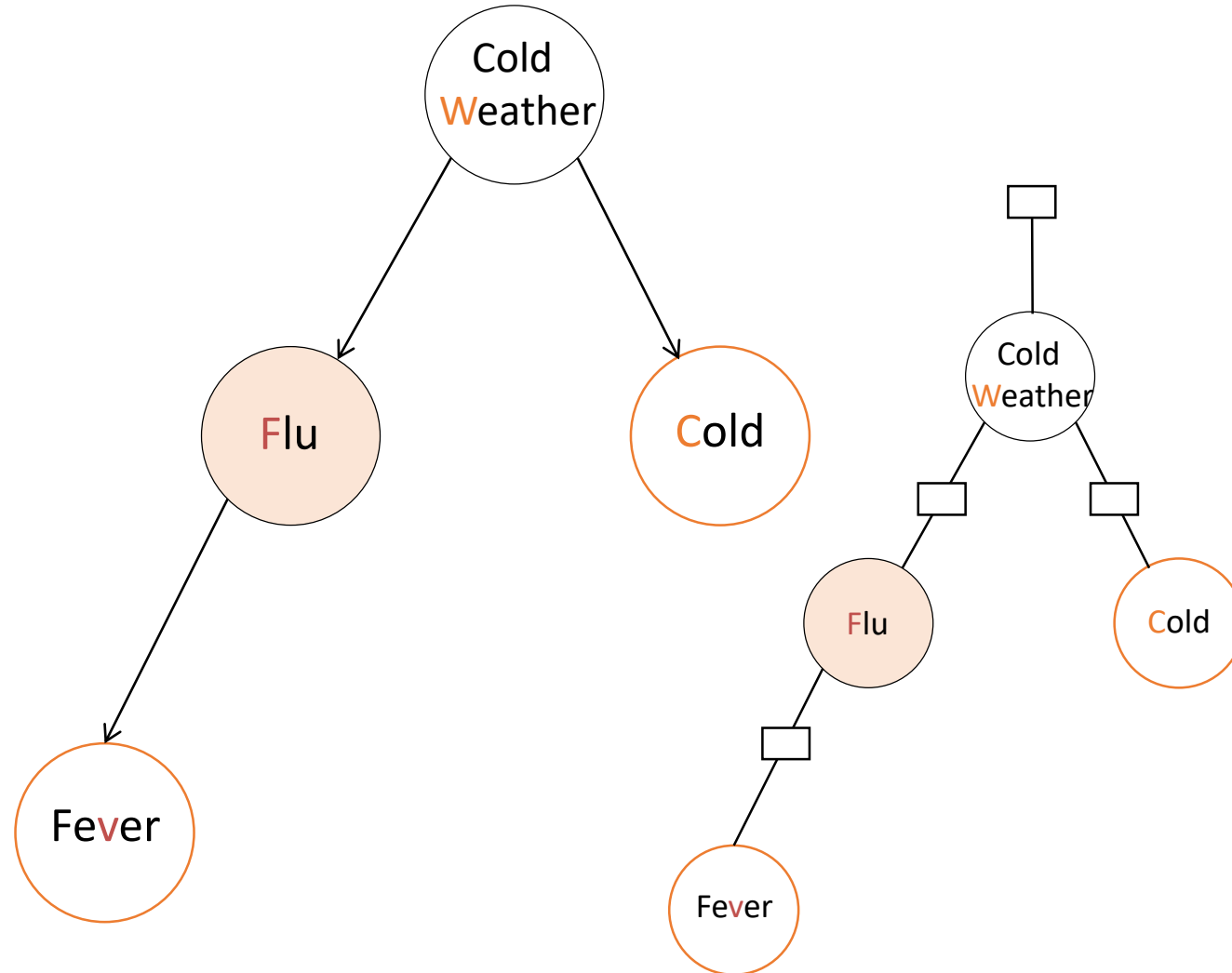
# Conditional Independence



# Conditional Independence

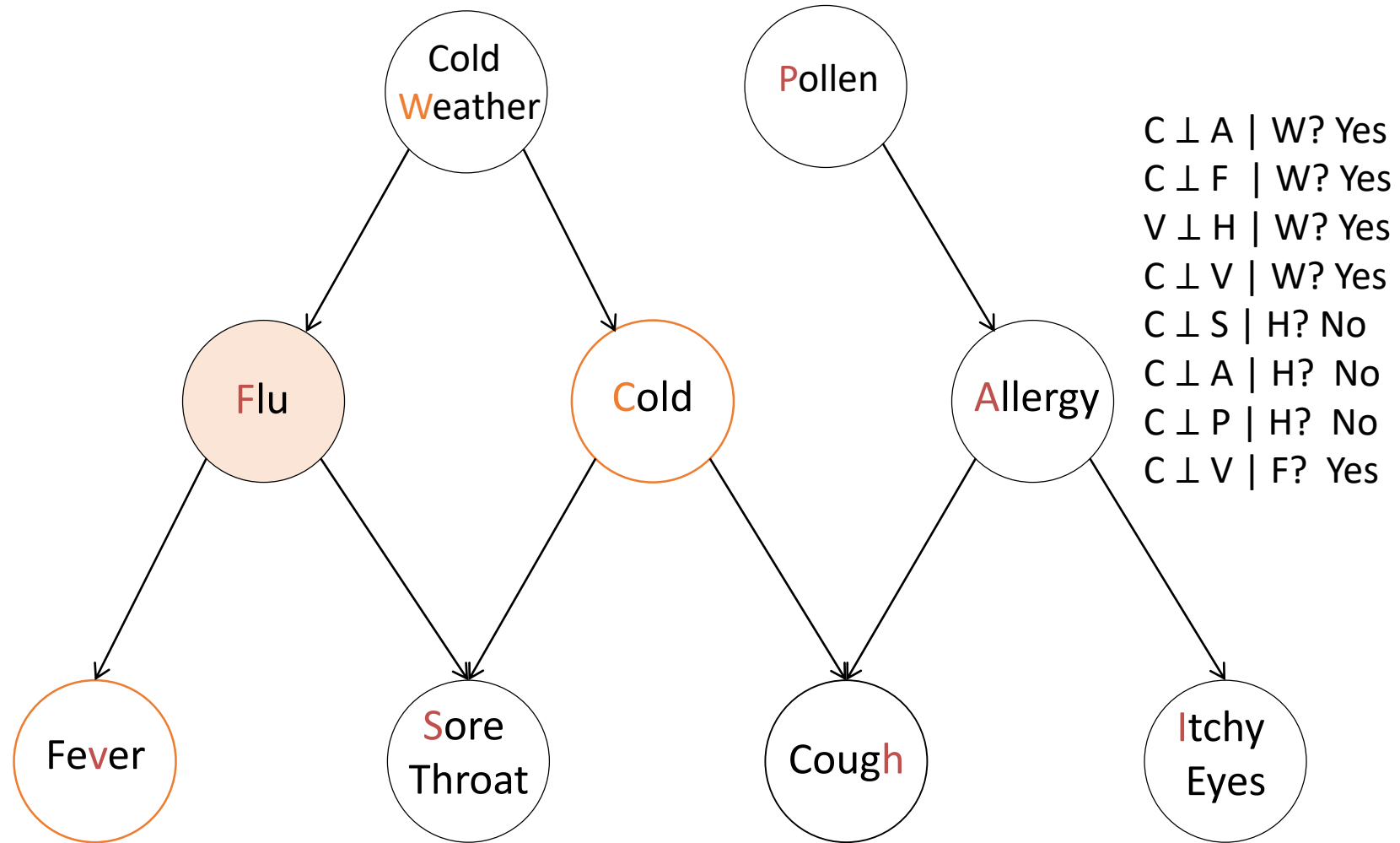


# Conditional Independence



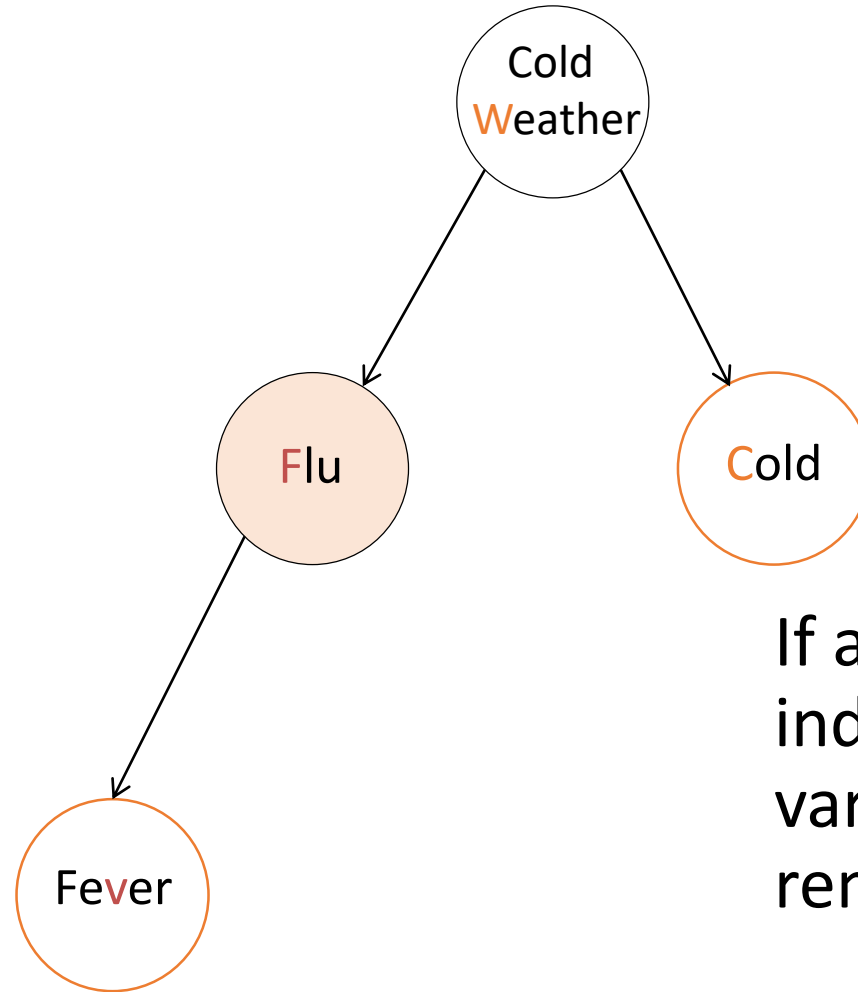
$C \perp A \mid W?$  Yes  
 $C \perp F \mid W?$  Yes  
 $V \perp H \mid W?$  Yes  
 $C \perp V \mid W?$  Yes  
 $C \perp S \mid H?$  No  
 $C \perp A \mid H?$  No  
 $C \perp P \mid H?$  No  
 $C \perp V \mid F?$

# Conditional Independence





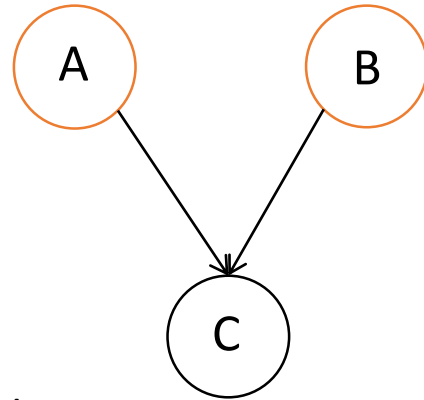
# Conditional Independence



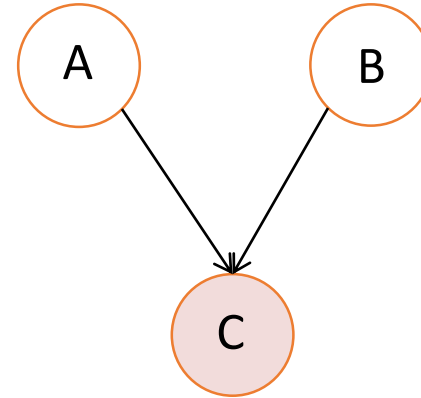
$$P(C=c \mid F=f) = ?$$

If a variable (Fever) is independent of the Query variable Q (Cold), we can remove (marginalize) it.

# Patterns



Independent



Dependent

