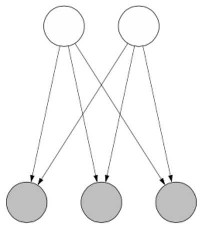
CS221 Section 7
Bayesian Networks

Nov 9th 2018

Roadmap

- Bayesian Networks Introduction
- Probabilistic Queries
- Conditional Independence
- Gibbs Sampling

Bayesian Networks



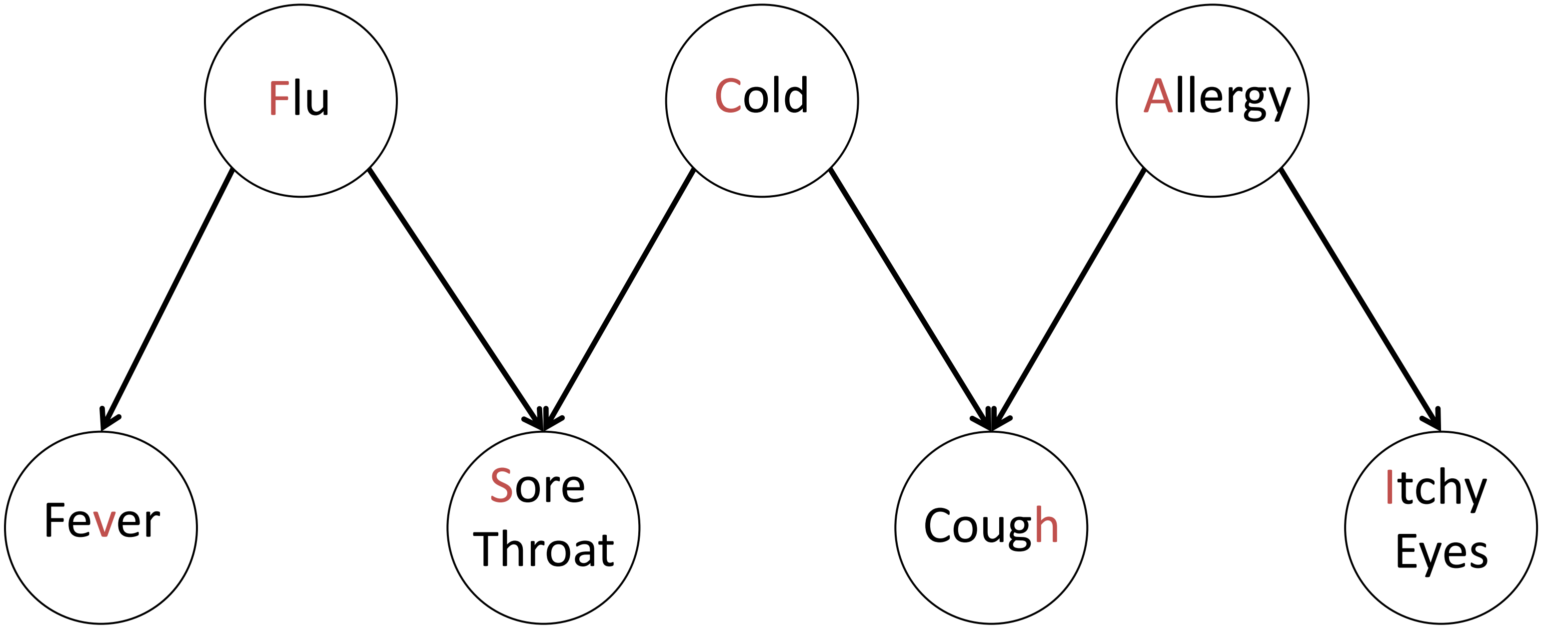
Definition: Bayesian network

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

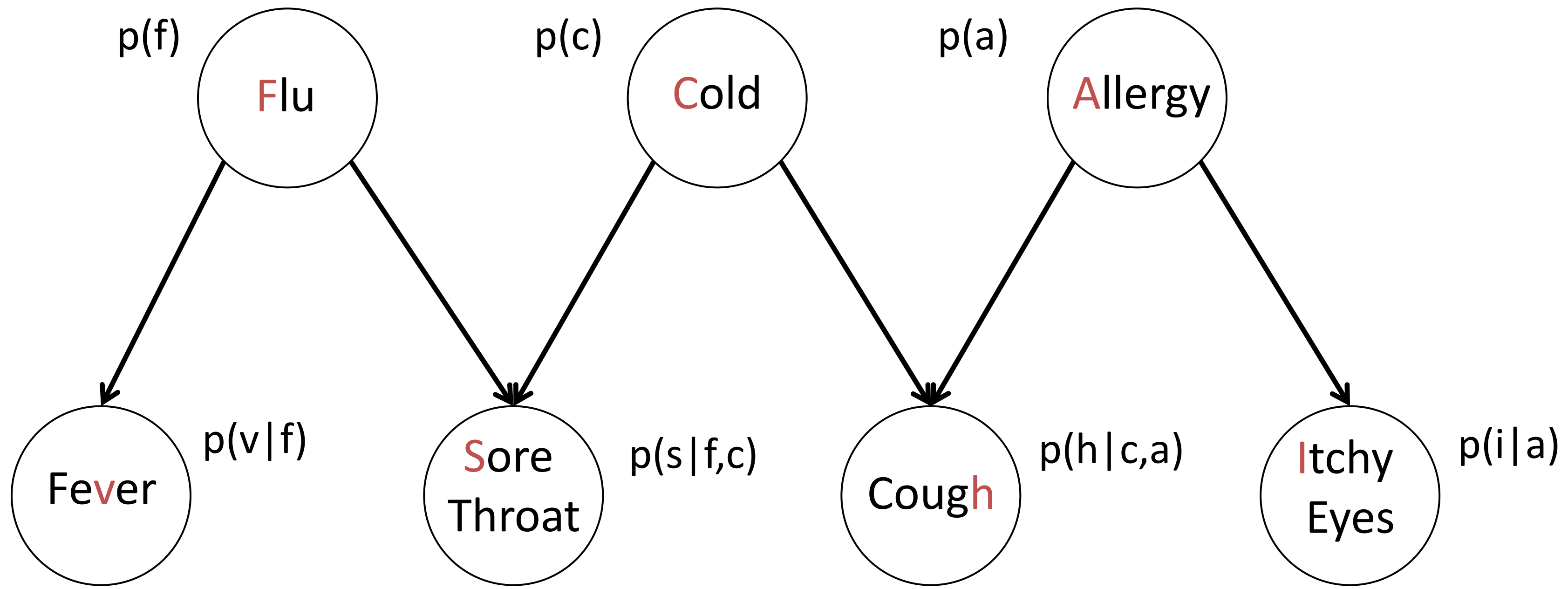
A **Bayesian network** is a directed acyclic graph (DAG) that specifies a joint distribution over X as a product of local conditional distributions, one for each node:

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

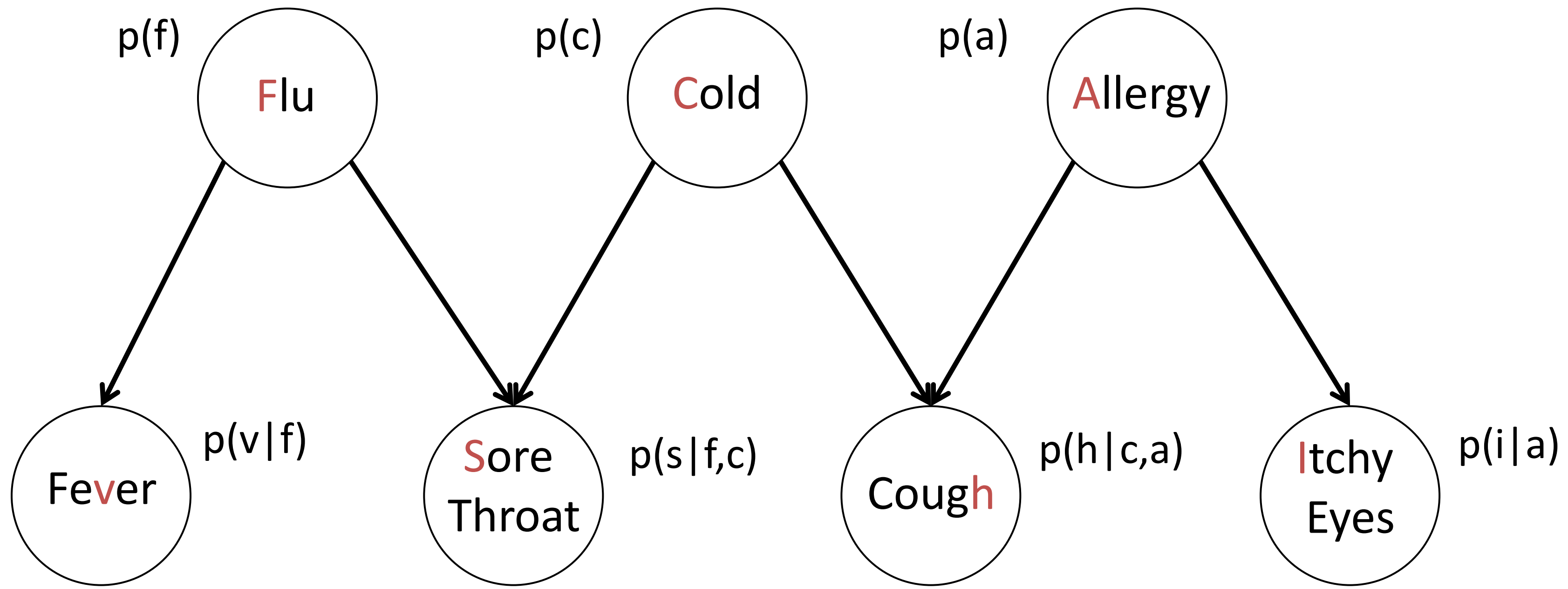
Bayesian Networks



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, C=c, I=i) = p(f)p(c)p(a)p(v|a)p(s|f, c)p(h|c,a)p(i|a)

Roadmap

- Bayesian Networks Introduction
- Probabilistic Queries
- Conditional Independence
- Gibbs Sampling

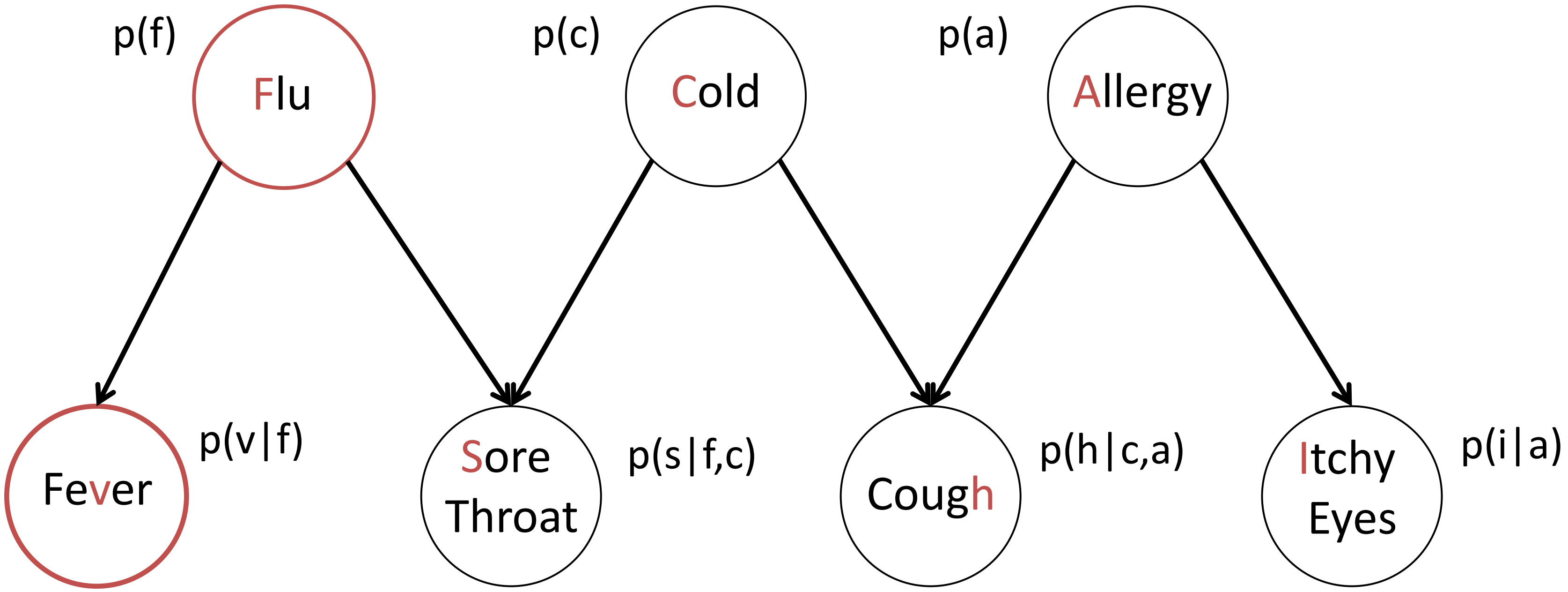
Probabilistic Queries – Cookbook

Given a query $P(Q|E=e)$

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).

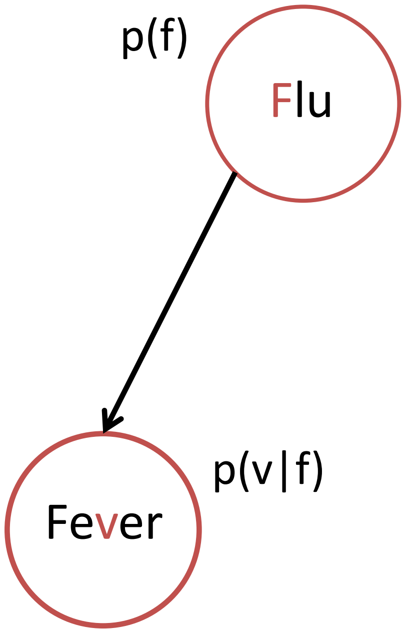
Probabilistic Queries - Examples

P(F=1|V=1) = ?



Probabilistic Queries - Examples

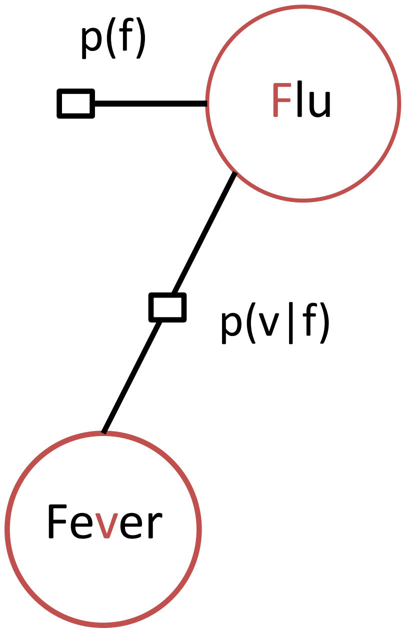
P(F=1|V=1) = ?



1. Remove (marginalize) variables not ancestors of Q or E.

Probabilistic Queries - Examples

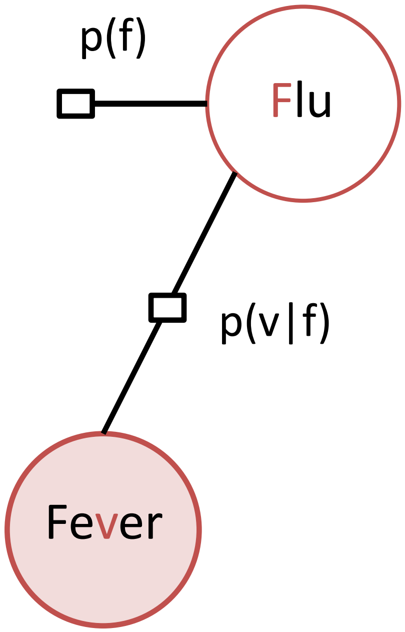
P(F=1|V=1) = ?



2. Convert Bayesian network to factor graph.

Probabilistic Queries - Examples

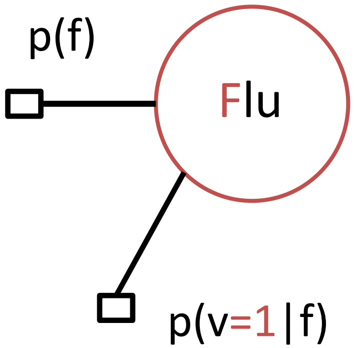
$P(F=1|V=1) = ?$



3. Condition on E = e.
    3.1 shade nodes

Probabilistic Queries - Examples

$P(F=1|V=1) = ?$

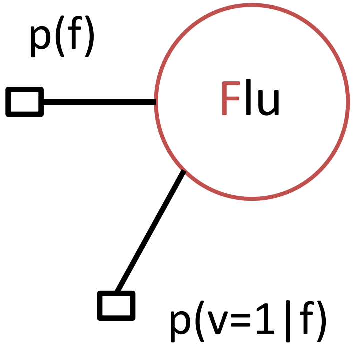


3. Condition on E = e.
   3.2 disconnect

Probabilistic Queries - Examples

$P(F=1|V=1) = ?$

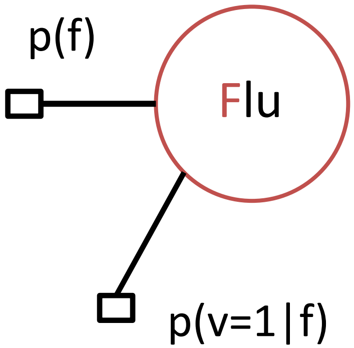
4. Remove (marginalize) nodes disconnected from Q.



Probabilistic Queries - Examples

$P(F=1|V=1) = ?$

5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).



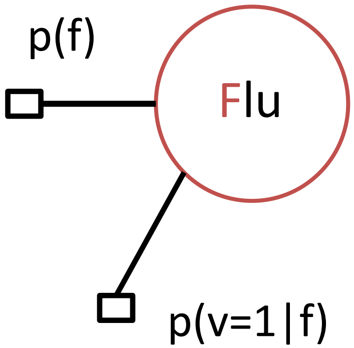
f	p(f)
0	1-α
1	α

f	v	p(v|f)
0	0	0.70
0	1	0.30
1	0	0.20
1	1	0.80

$$P(F=f|V=1)\ \mu\ \ p(f)\, p(v=1|f)$$

Probabilistic Queries - Examples

$P(F{=}1 \mid V{=}1) = ?$



5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

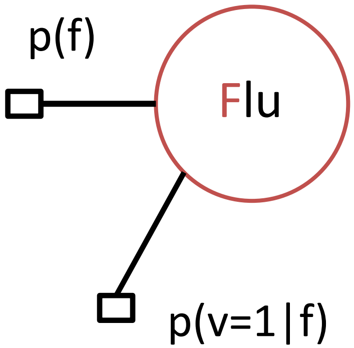
f	p(f)
0	1-α
1	α

f	v	p(v|f)
0	0	0.70
0	1	0.30
1	0	0.20
1	1	0.80

$$P(F{=}f \mid V{=}1) \propto p(f)\, p(v{=}1 \mid f) = (1 \qquad) * 0.30, \quad f = 0$$

Probabilistic Queries - Examples

P(F=1|V=1) = ?



5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

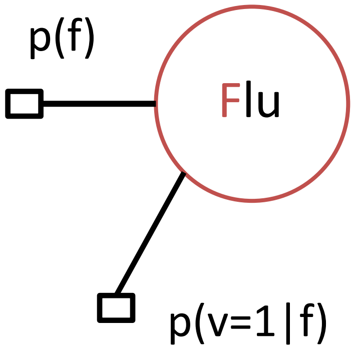
f	p(f)
0	1-α
1	α

f	v	p(v|f)
0	0	0.70
0	1	0.30
1	0	0.20
1	1	0.80

$$P(F=f|V=1) \propto p(f)\, p(v=1|f) = \begin{cases} (1 \quad\quad) * 0.30, & f = 0 \\ \quad * 0.80, & f = 1 \end{cases}$$

Probabilistic Queries - Examples

P(F=1|V=1) = ?



5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

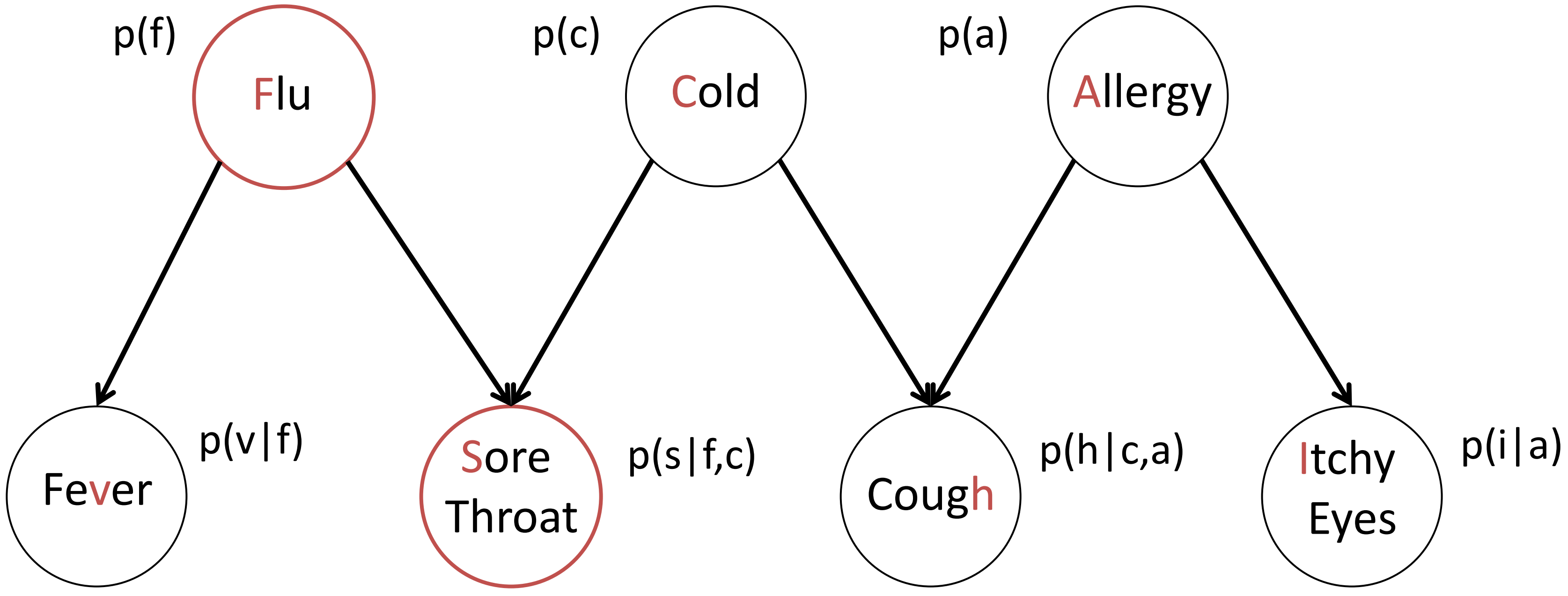
f	v	p(v|f)
0	0	0.70
0	1	0.30
1	0	0.20
1	1	0.80

$$P(F=f|V=1)\,\mu\;\; p(f)\,p(v=1|f) = \begin{cases} (1 \quad\;)*0.30, & f=0 \\ \quad *0.80, & f=1 \end{cases}$$

$$P(F=1|V=1) = \frac{*0.80}{*0.80+(1 \quad\;)*0.30}$$

Probabilistic Queries - Examples

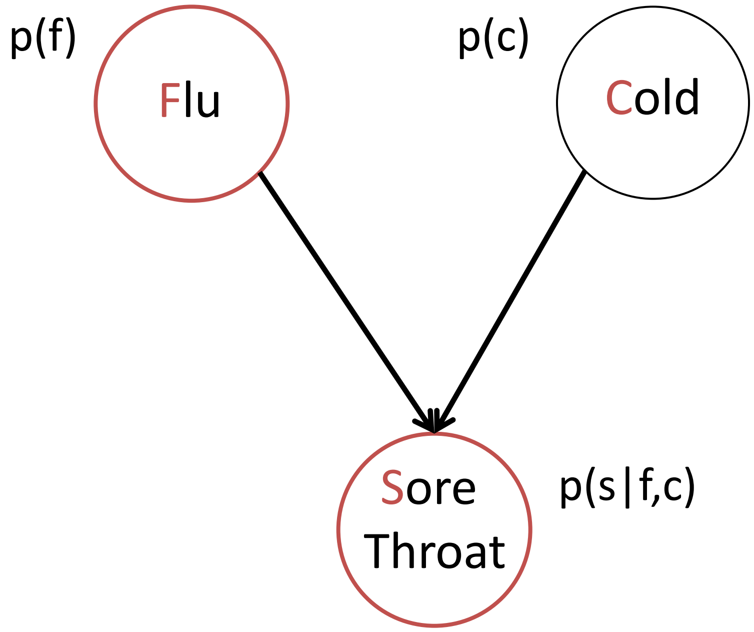
P(F=1|S=1) = ?



Probabilistic Queries - Examples

$P(F=1|S=1) = ?$

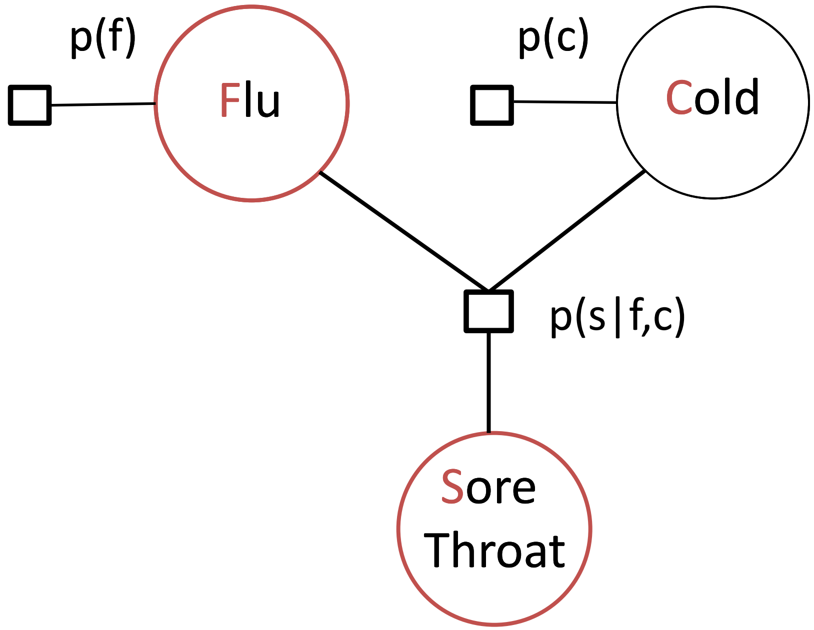
1. Remove (marginalize) variables not ancestors of Q or E.



Probabilistic Queries - Examples

$P(F=1|S=1) = ?$

2. Convert Bayesian network to factor graph.

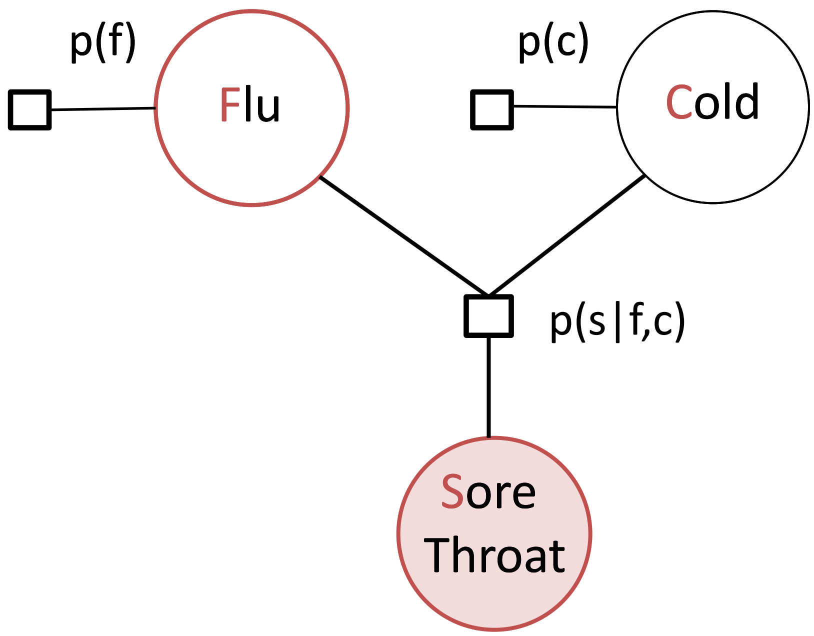


★ One factor per variable!

Probabilistic Queries - Examples

$P(F=1|S=1) = ?$

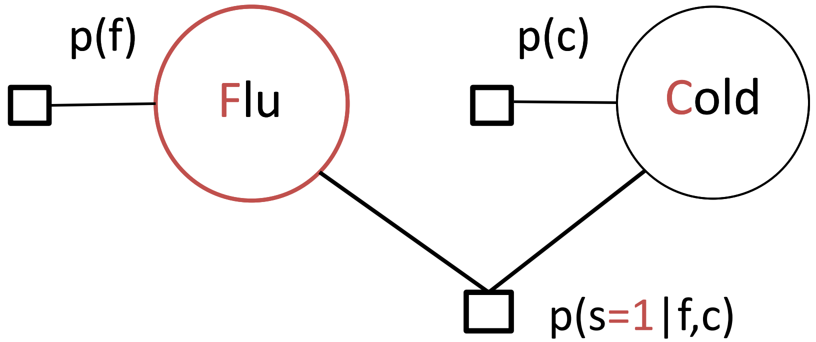
3. Condition on E = e.
   3.1 shade nodes



Probabilistic Queries - Examples

P(F=1|S=1) = ?

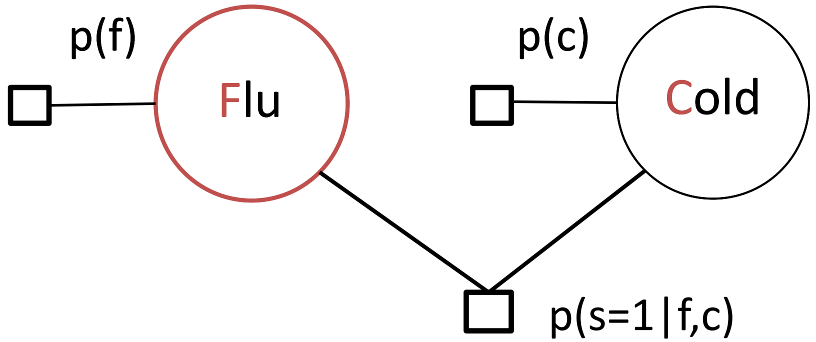
3. Condition on E = e.
   3.2 disconnect



Probabilistic Queries - Examples

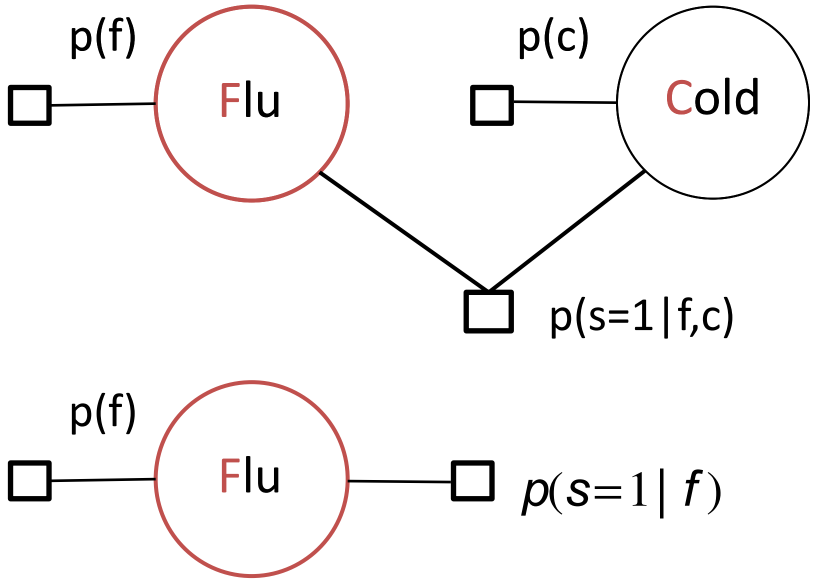
$P(F=1|S=1) = ?$

4. Remove (marginalize) nodes disconnected from Q.



Probabilistic Queries - Examples

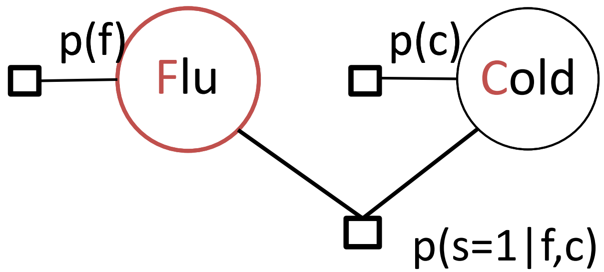
$P(F=1|S=1) = ?$

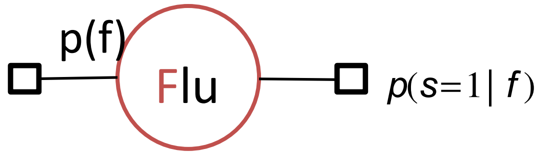


5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

Probabilistic Queries - Examples

P(F=1|S=1) = ?





5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

$$p(s=1|f)$$

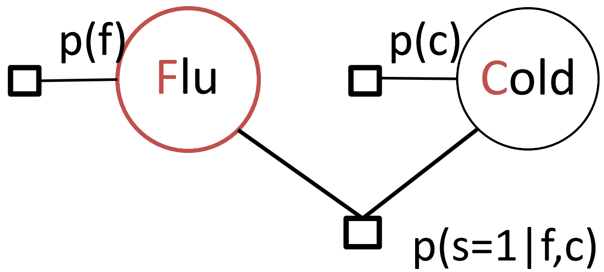
$$= \sum_c p(c)p(s=1|f,c)$$

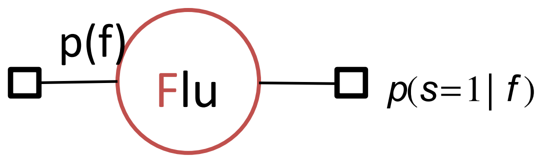
$$= p(c=0)p(s=1|f,c=0)+p(c=1)p(s=1|f,c=1)$$

f	p(s=1,f)
0	?
1	?

Probabilistic Queries - Examples

$P(F=1|S=1) = ?$





$$p(s=1|f)$$

$$= {}_{c}\, p(c)p(s=1|f,c)$$

$$= p(c=0)p(s=1|f,c=0) + p(c=1)p(s=1|f,c=1)$$

$$= (1 \quad)*0 + \quad *0.75, \quad f=0$$

5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

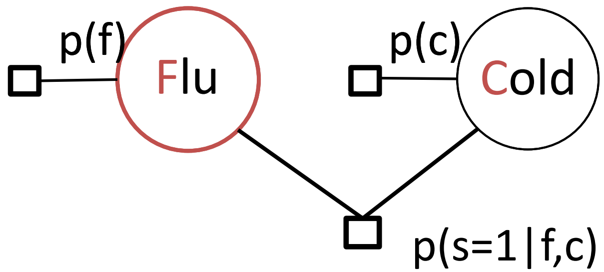
c	p(c)
0	1-β
1	β

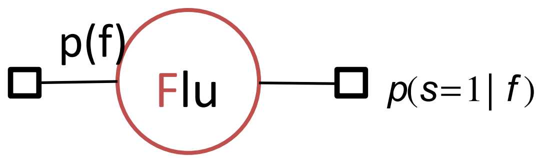
s	f	c	p(s|f,c)
0	0	0	1.00
1	0	0	0
0	1	0	0.30
1	1	0	0.70
0	0	1	0.25
1	0	1	0.75
0	1	1	0.10
1	1	1	0.90

f	p(s=1,f)
0	β*0.75
1	?

Probabilistic Queries - Examples

P(F=1|S=1) = ?





$$p(s=1|f)$$

$$= \quad {}_{c}\, p(c)p(s=1|f,c)$$

$$= p(c=0)p(s=1|f,c=0)+p(c=1)p(s=1|f,c=1)$$

$$= \begin{array}{ll} (1 \quad)*0+ \quad *0.75, & f=0 \\ (1 \quad)*0.70+ \quad *0.9, & f=1 \end{array}$$

5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

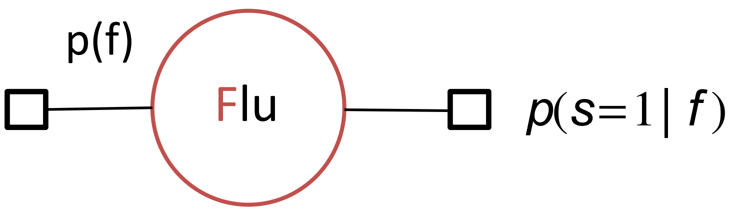
c	p(c)
0	1-β
1	β

s	f	c	p(s|f,c)
0	0	0	1.00
1	0	0	0
0	1	0	0.30
1	1	0	0.70
0	0	1	0.25
1	0	1	0.75
0	1	1	0.10
1	1	1	0.90

f	p(s=1,f)
0	β*0.75
1	((1-β)*0.7+β*0.9)

Probabilistic Queries - Examples

P(F=1|S=1) = ?



$$P(F=f\,|\,S=1)$$

$$\propto p(f)\,p(s=1\,|\,f)$$

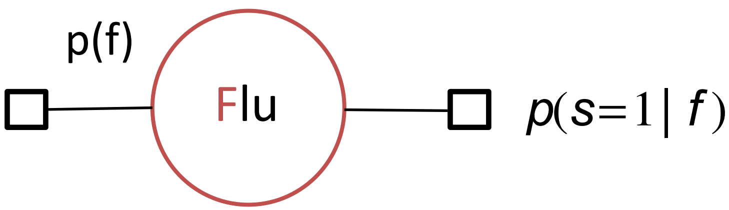
5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

f	p(s=1|f)
0	β*0.75
1	((1-β)*0.7+β*0.9)

Probabilistic Queries - Examples

P(F=1|S=1) = ?



$$P(F = f\,|\,S = 1)$$

$$\mu\ p(f)\,p(s=1\,|\,f)$$

$$= \quad (1 \qquad) \;*0.75, \qquad f = 0$$

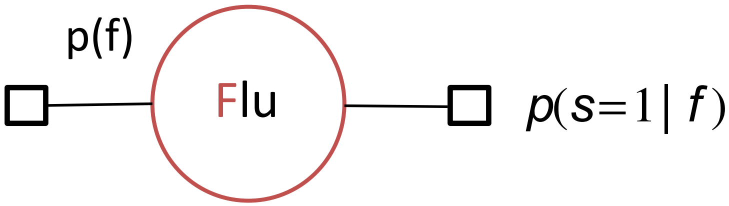
5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

f	p(s=1|f)
0	β*0.75
1	((1-β)*0.7+β*0.9)

Probabilistic Queries - Examples

P(F=1|S=1) = ?



$$P(F=f\mid S=1)$$

$$\mu\ p(f)p(s=1\mid f)$$

$$= \begin{cases} (1\quad)\ *0.75, & f=0 \\ ((1\quad)*0.70+\quad*0.9), & f=1 \end{cases}$$

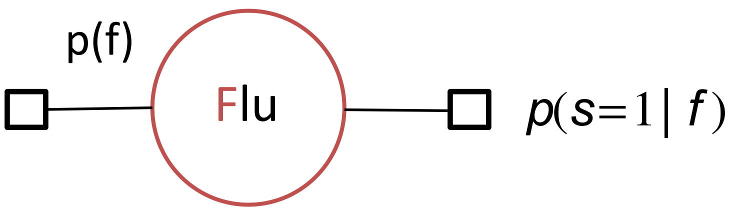
5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

f	p(s=1|f)
0	β*0.75
1	(1-β)*0.7+β*0.9

Probabilistic Queries - Examples

P(F=1|S=1) = ?



$$P(F=f\,|\,S=1)$$

$$\mu\ \ p(f)p(s=1\,|\,f)$$

$$=\begin{array}{ll}(1\qquad)\quad *0.75, & \qquad f=0\\ ((1\qquad)*0.70+\quad *0.9), & \qquad f=1\end{array}$$

$$P(F=1\,|\,S=1)=\frac{p(f=1)p(s=1\,|\,f=1)}{p(f=1)p(s=1\,|\,f=1)+p(f=1)p(s=1\,|\,f=1)}$$

$$=\frac{((1\qquad)*0.70+\quad *0.9)}{(1\qquad)\quad *0.75+\quad((1\qquad)*0.70+\quad *0.9),}$$

5. Run probabilistic inference algorithm (manual, variable elimination, Gibbs sampling, particle filtering).

f	p(f)
0	1-α
1	α

f	p(s=1|f)
0	β*0.75
1	(1-β)*0.7+β*0.9

Probabilistic Queries – Cookbook

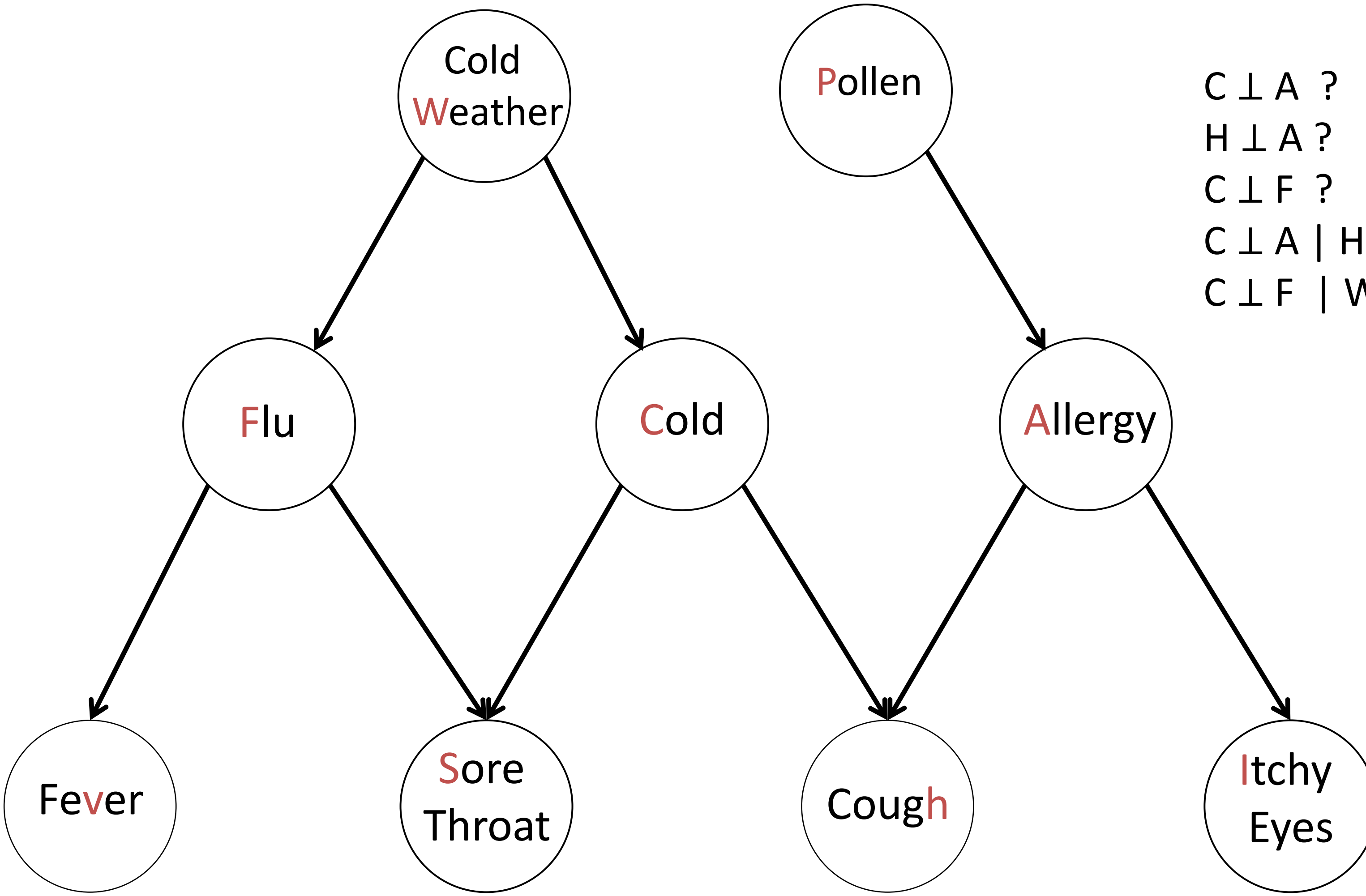
Given a query $P(Q|E=e)$

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).

Roadmap

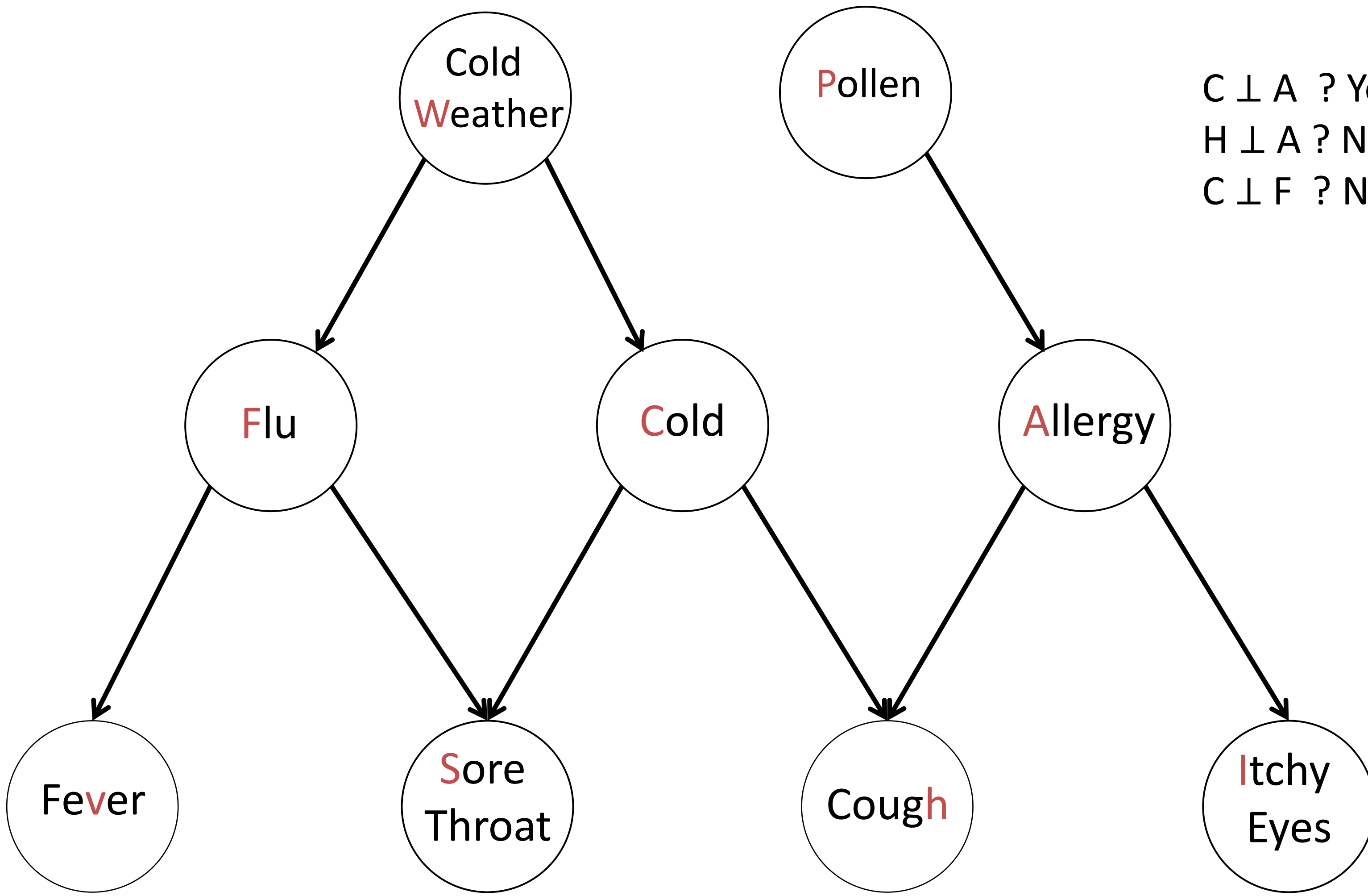
- Bayesian Networks Introduction
- Probabilistic Queries
- Conditional Independence
- Gibbs Sampling

Conditional Independence



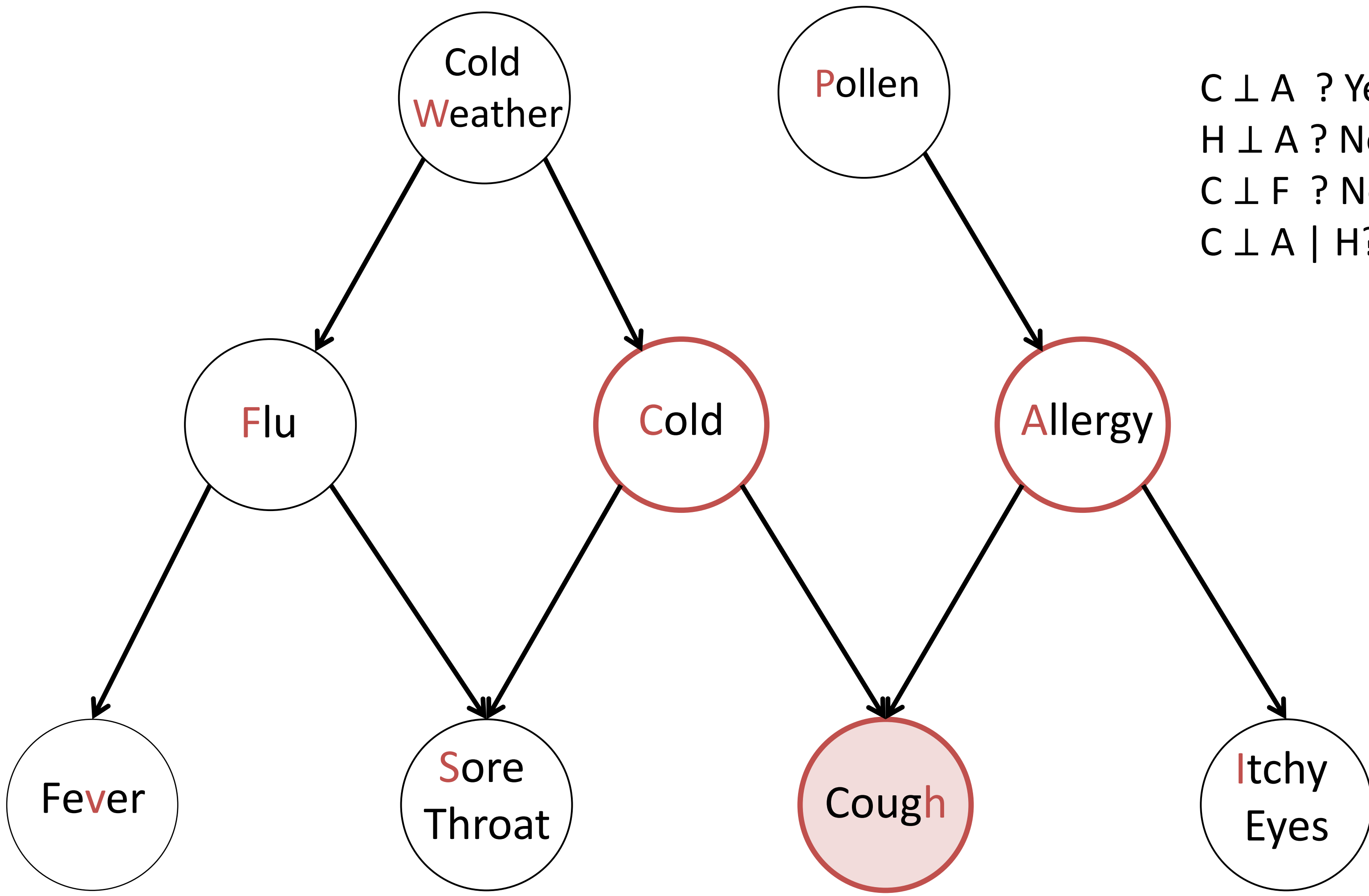
$C \perp A$?
$H \perp A$?
$C \perp F$?
$C \perp A \mid H$?
$C \perp F \mid W$?

Conditional Independence



$C \perp A$? Yes
$H \perp A$? No
$C \perp F$? No

Conditional Independence



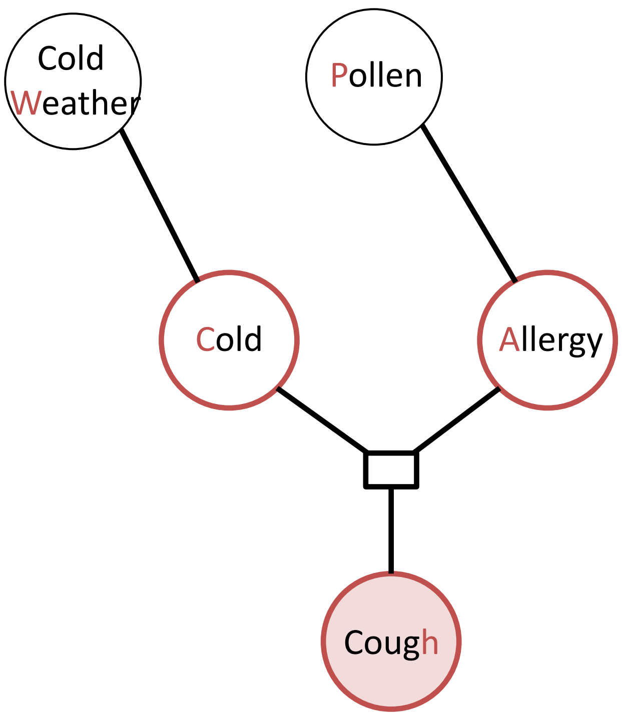
$C \perp A$? Yes

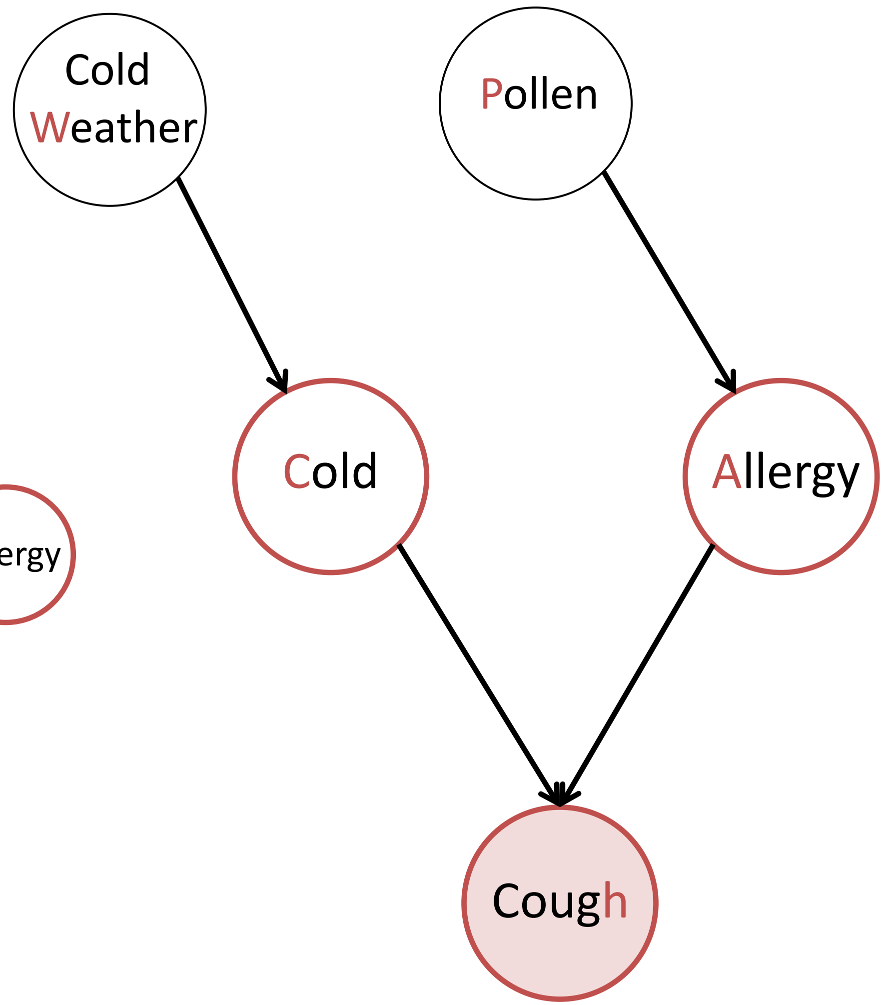
$H \perp A$? No

$C \perp F$? No

$C \perp A \mid H$?

Conditional Independence

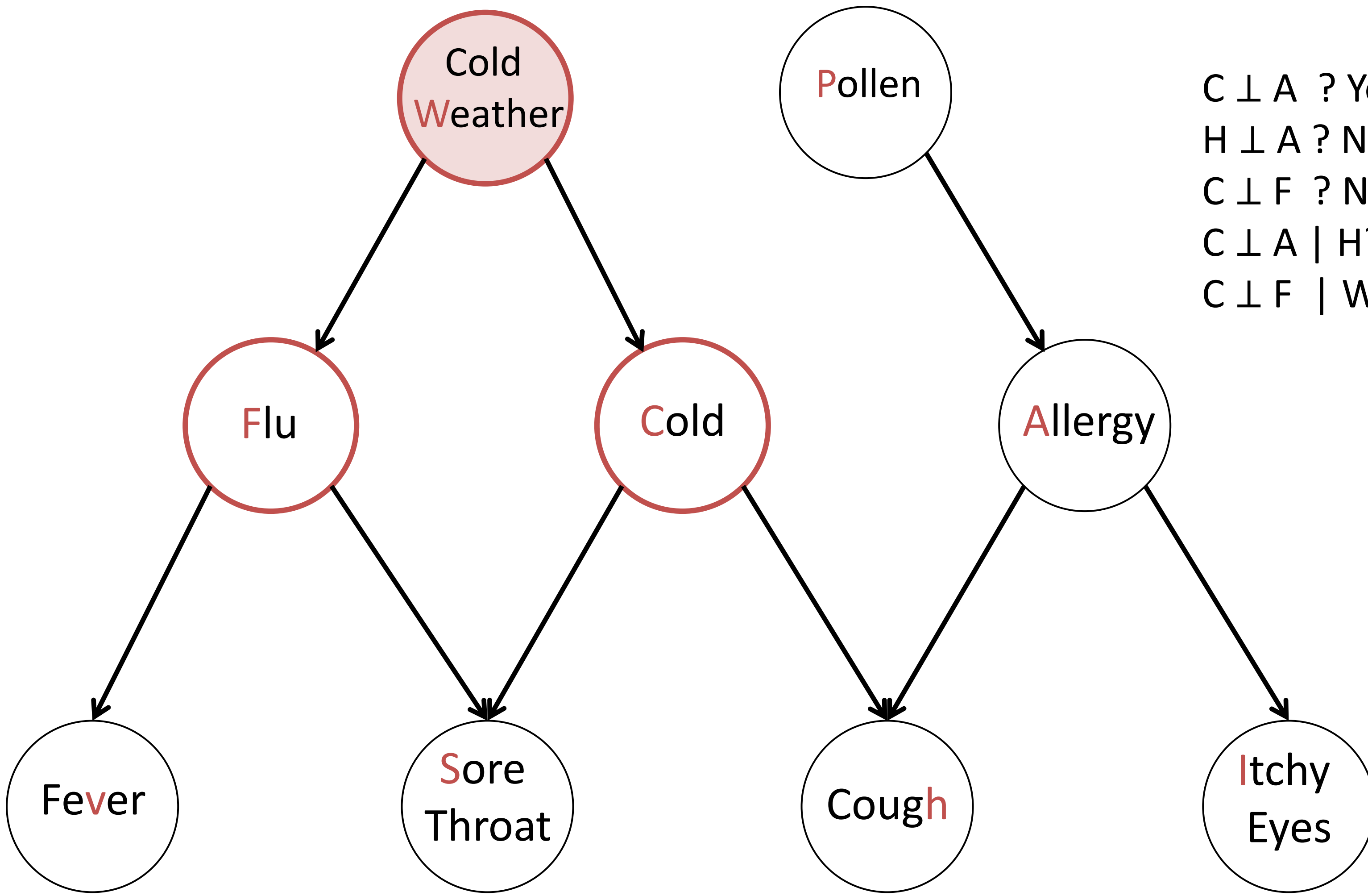




$C \perp A$? Yes
$H \perp A$? No
$C \perp F$? No
$C \perp A \mid H$?

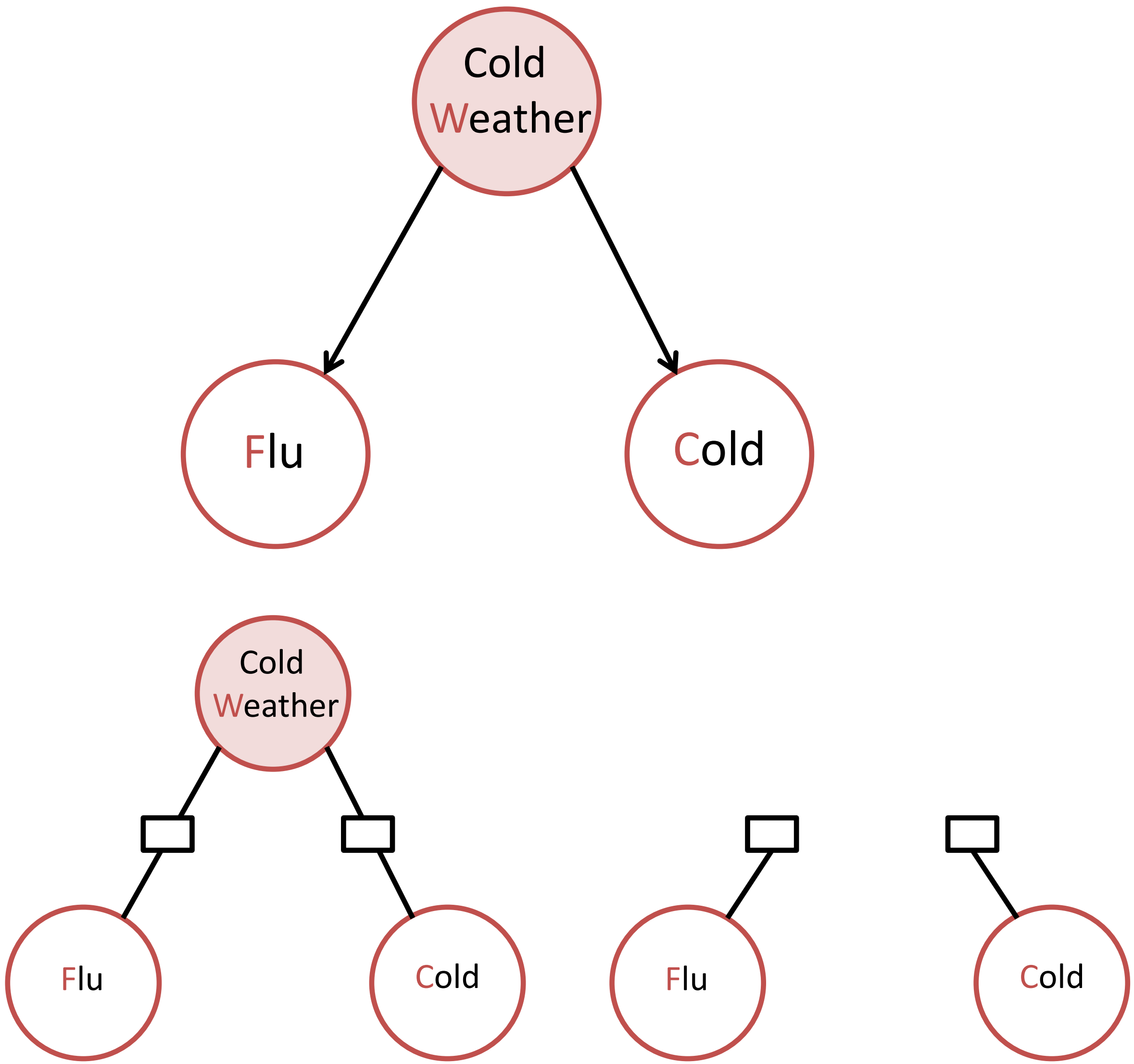
Explaining Away!

Conditional Independence



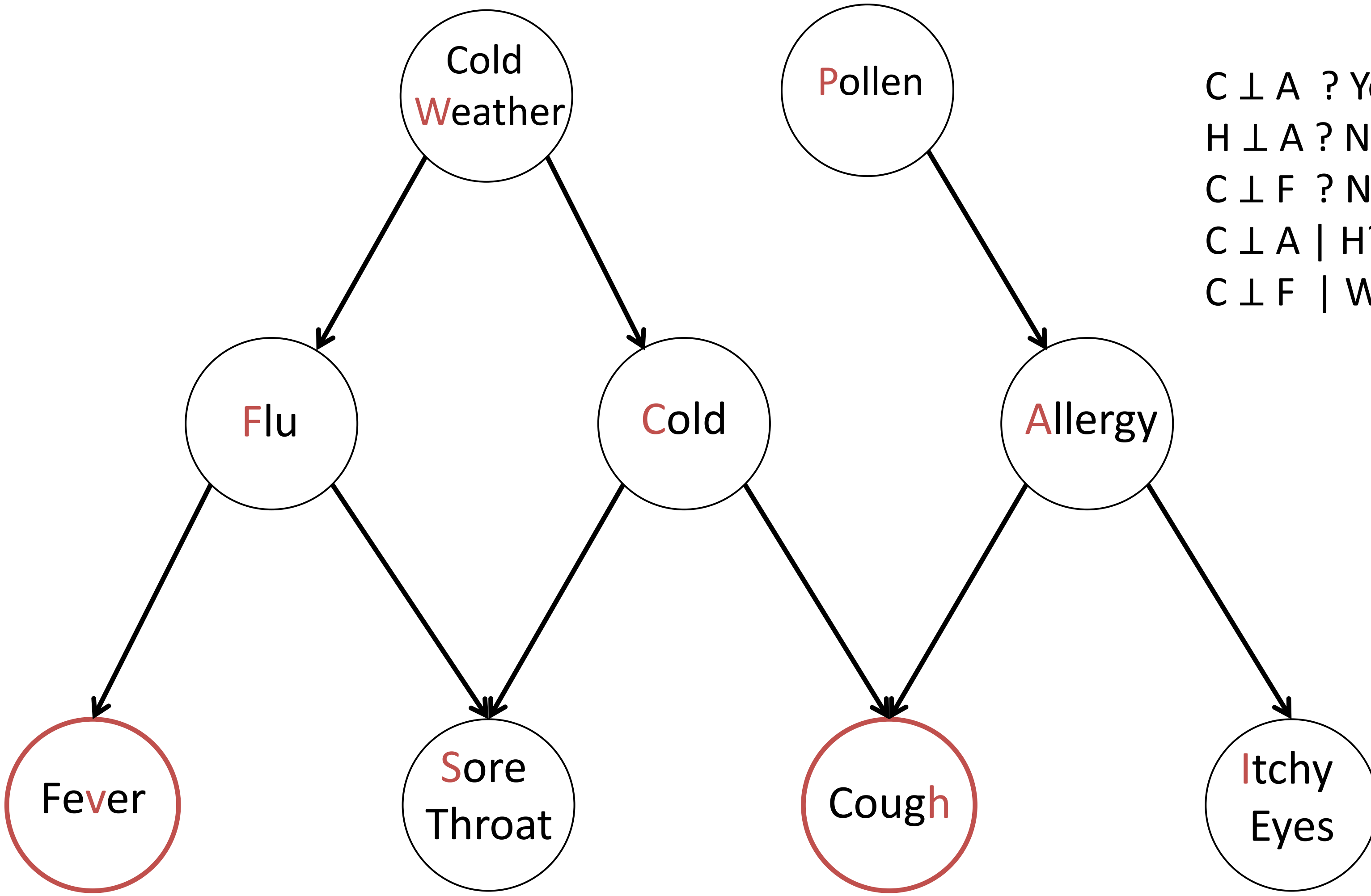
C ⊥ A ? Yes
H ⊥ A ? No
C ⊥ F ? No
C ⊥ A | H? No
C ⊥ F | W?

Conditional Independence



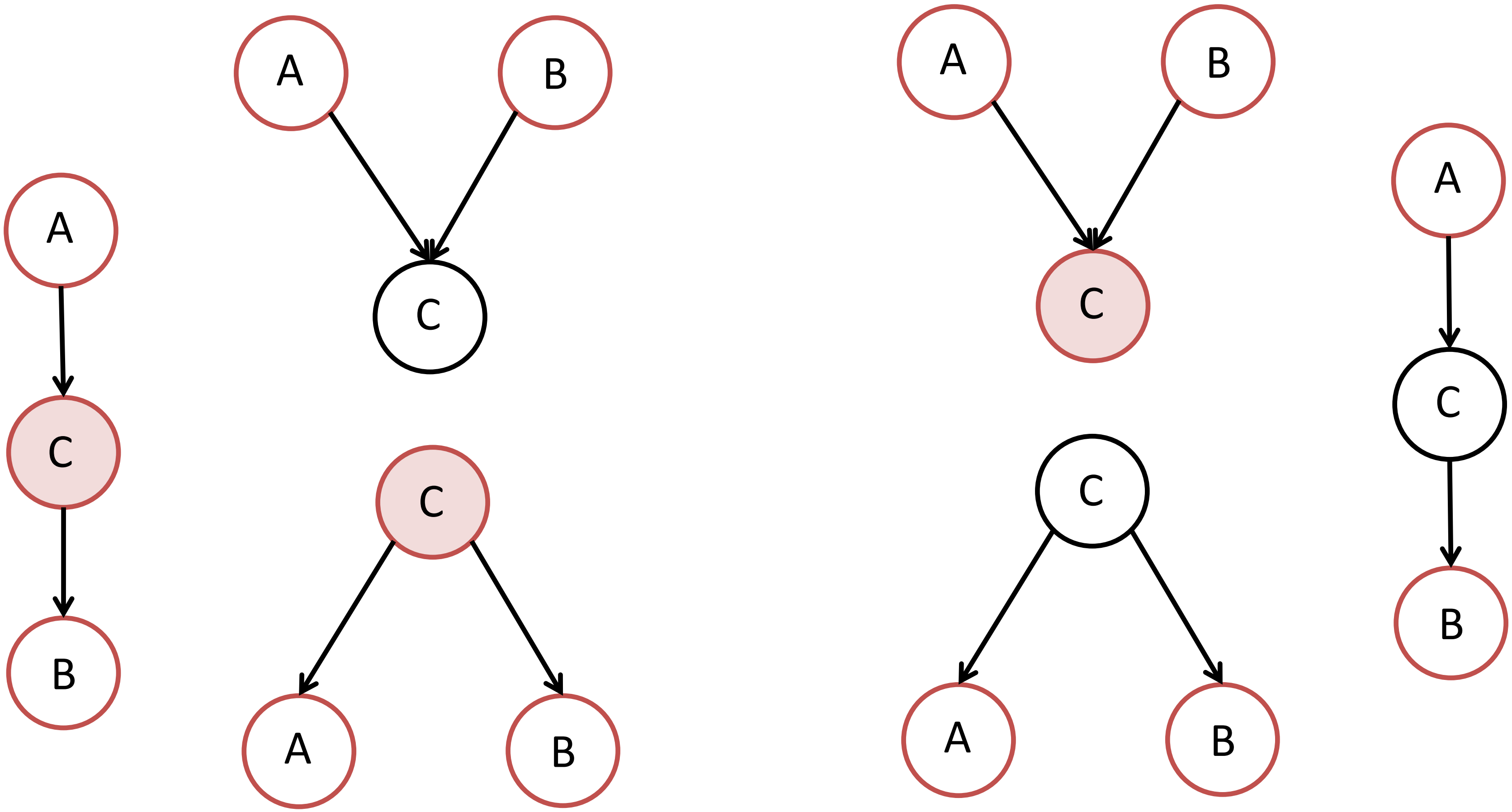
$C \perp A$? Yes
$H \perp A$? No
$C \perp F$? No
$C \perp A \mid H$? No
$C \perp F \mid W$?

Conditional Independence



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

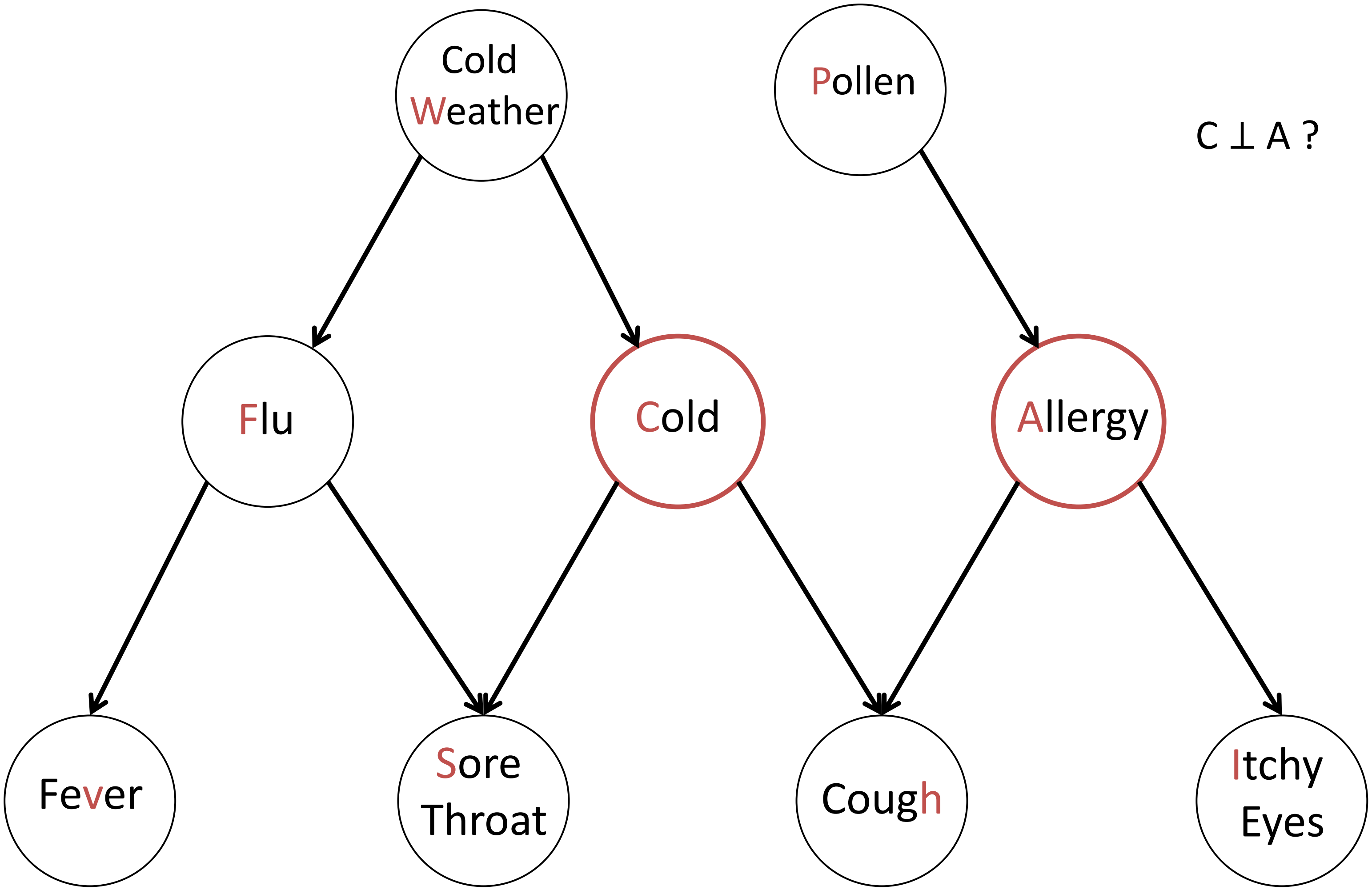
Patterns



Independent

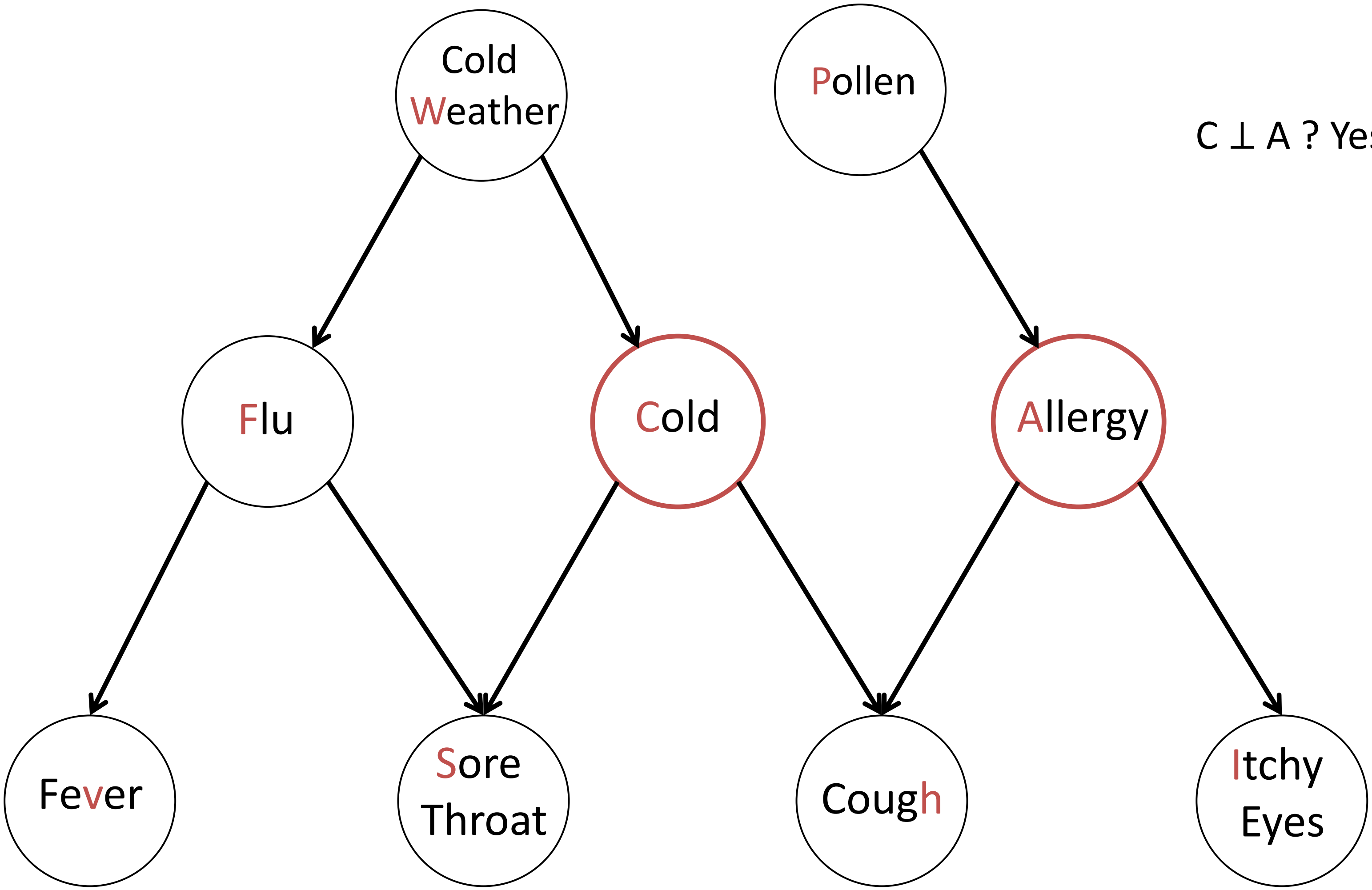
Dependent

Conditional Independence



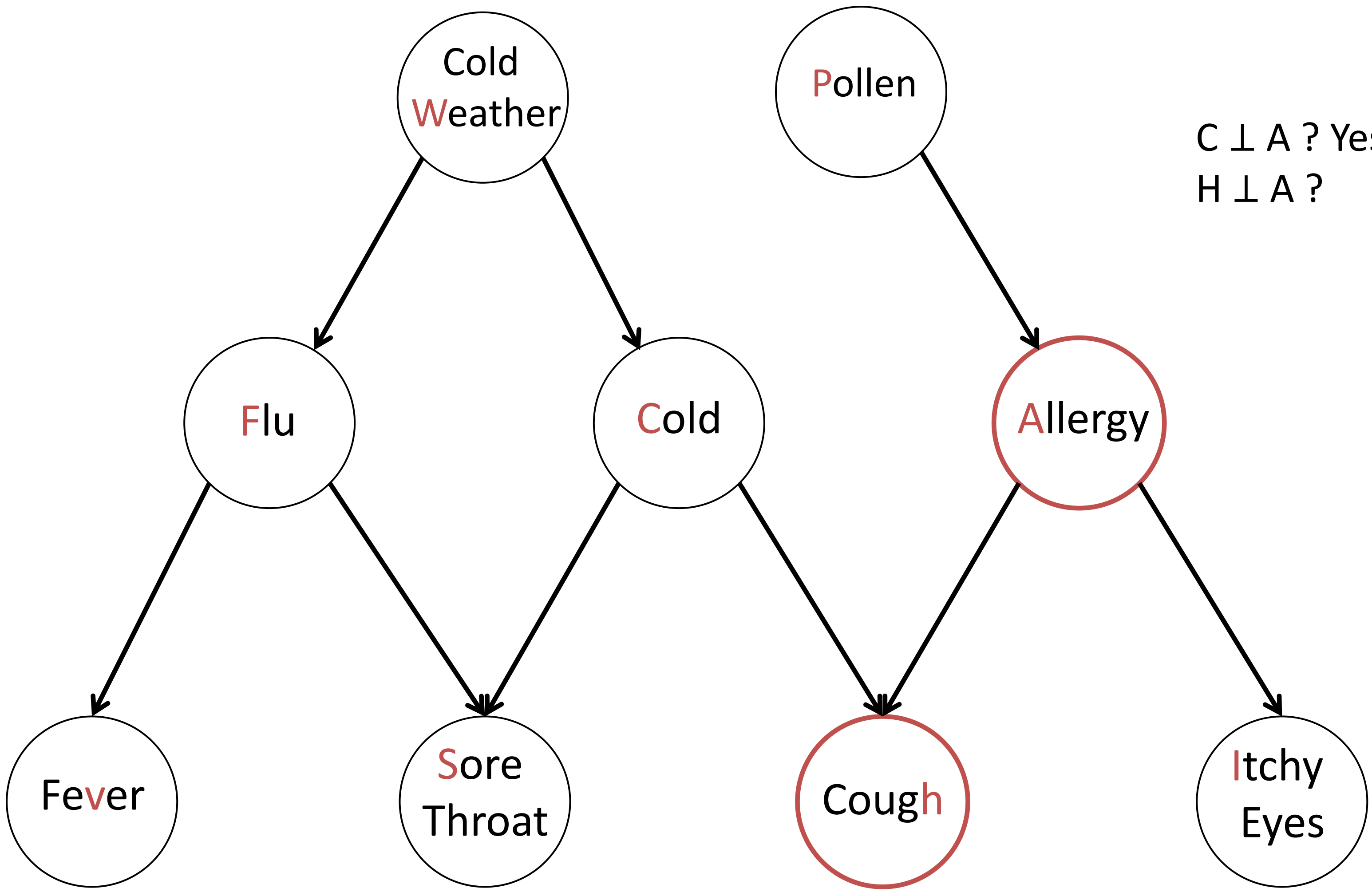
C ⊥ A ?

Conditional Independence



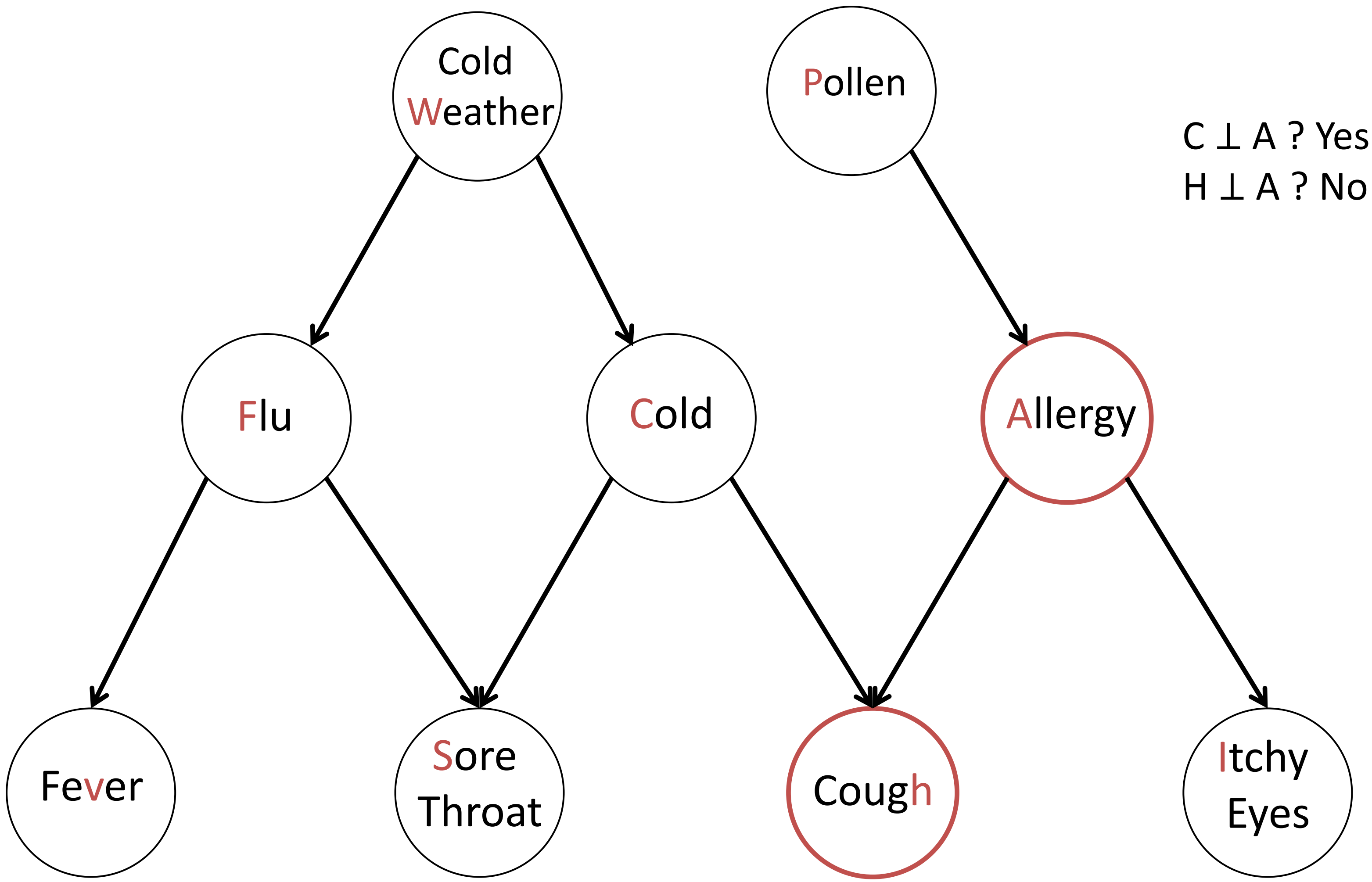
$C \perp A$? Yes

Conditional Independence



C ⊥ A ? Yes
H ⊥ A ?

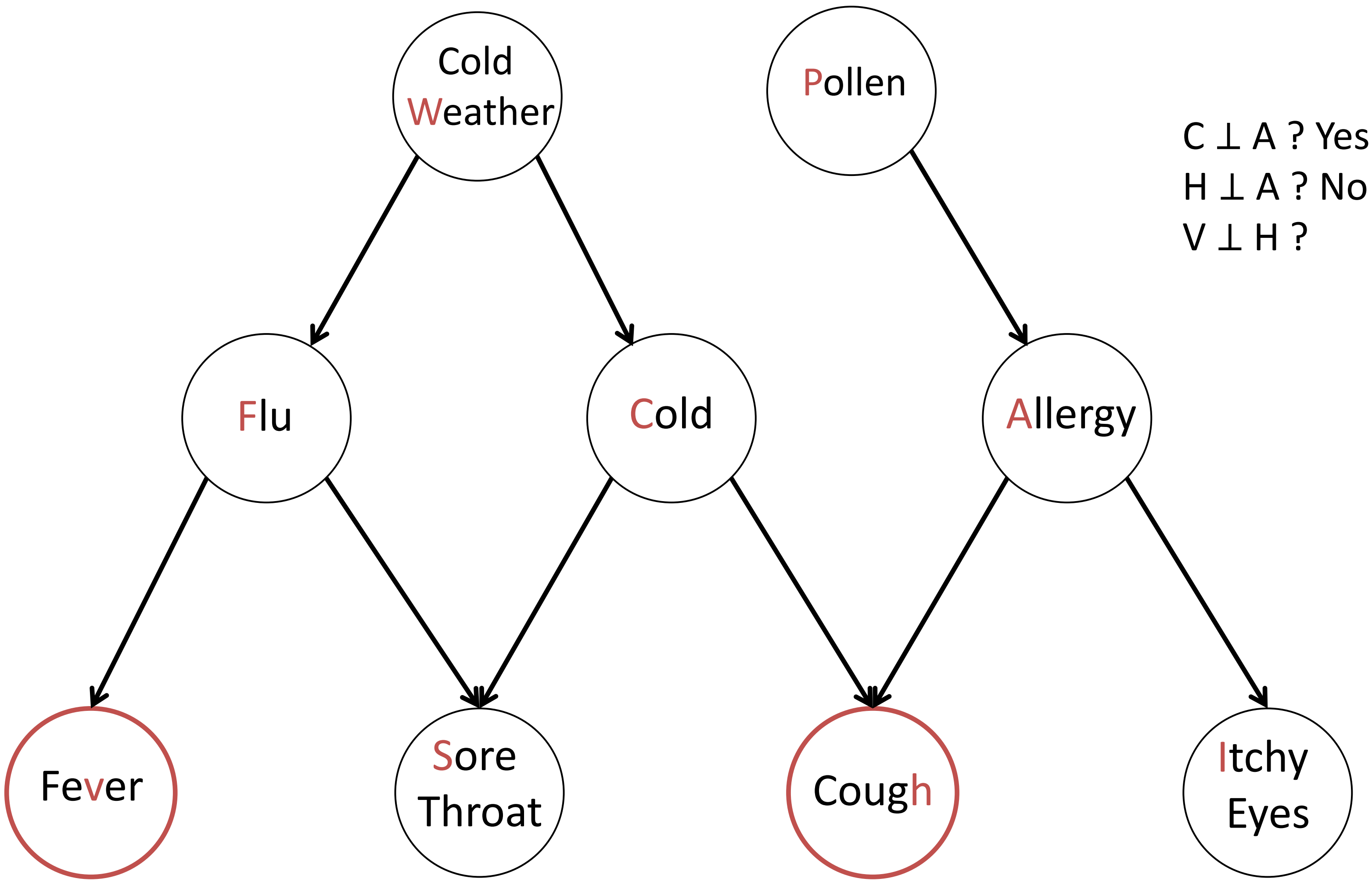
Conditional Independence



$C \perp A$? Yes

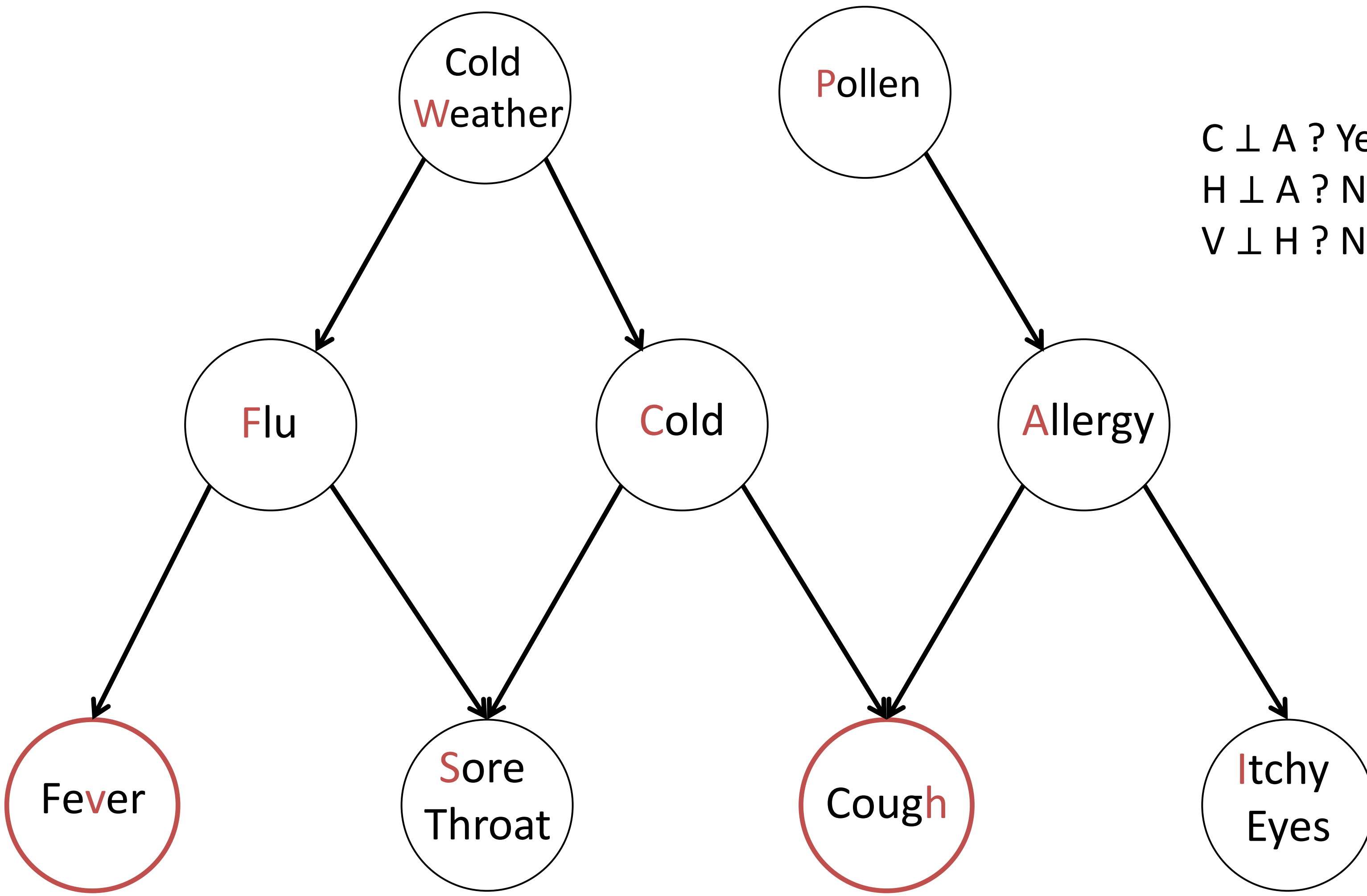
$H \perp A$? No

Conditional Independence



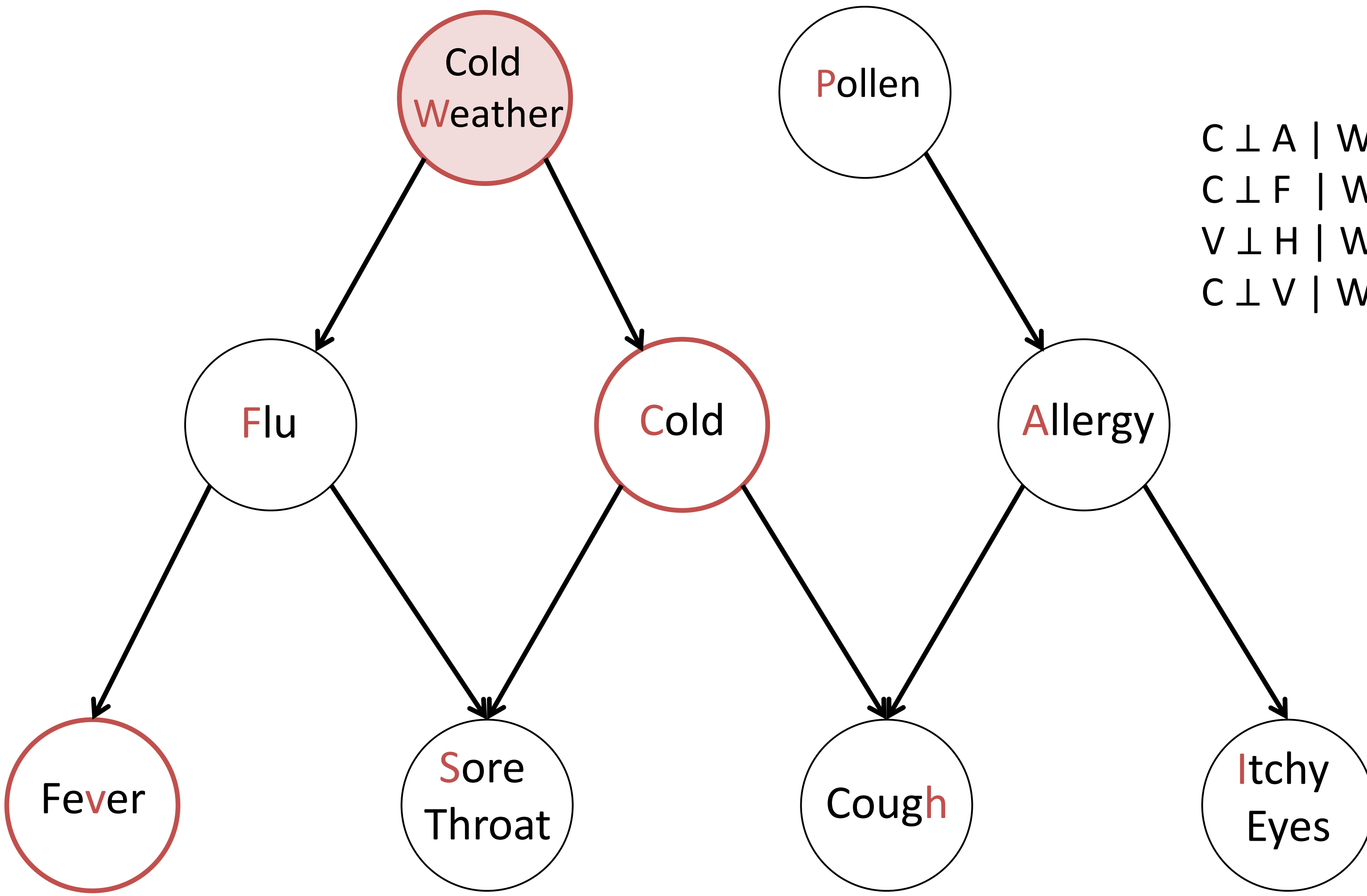
C ⊥ A ? Yes
H ⊥ A ? No
V ⊥ H ?

Conditional Independence



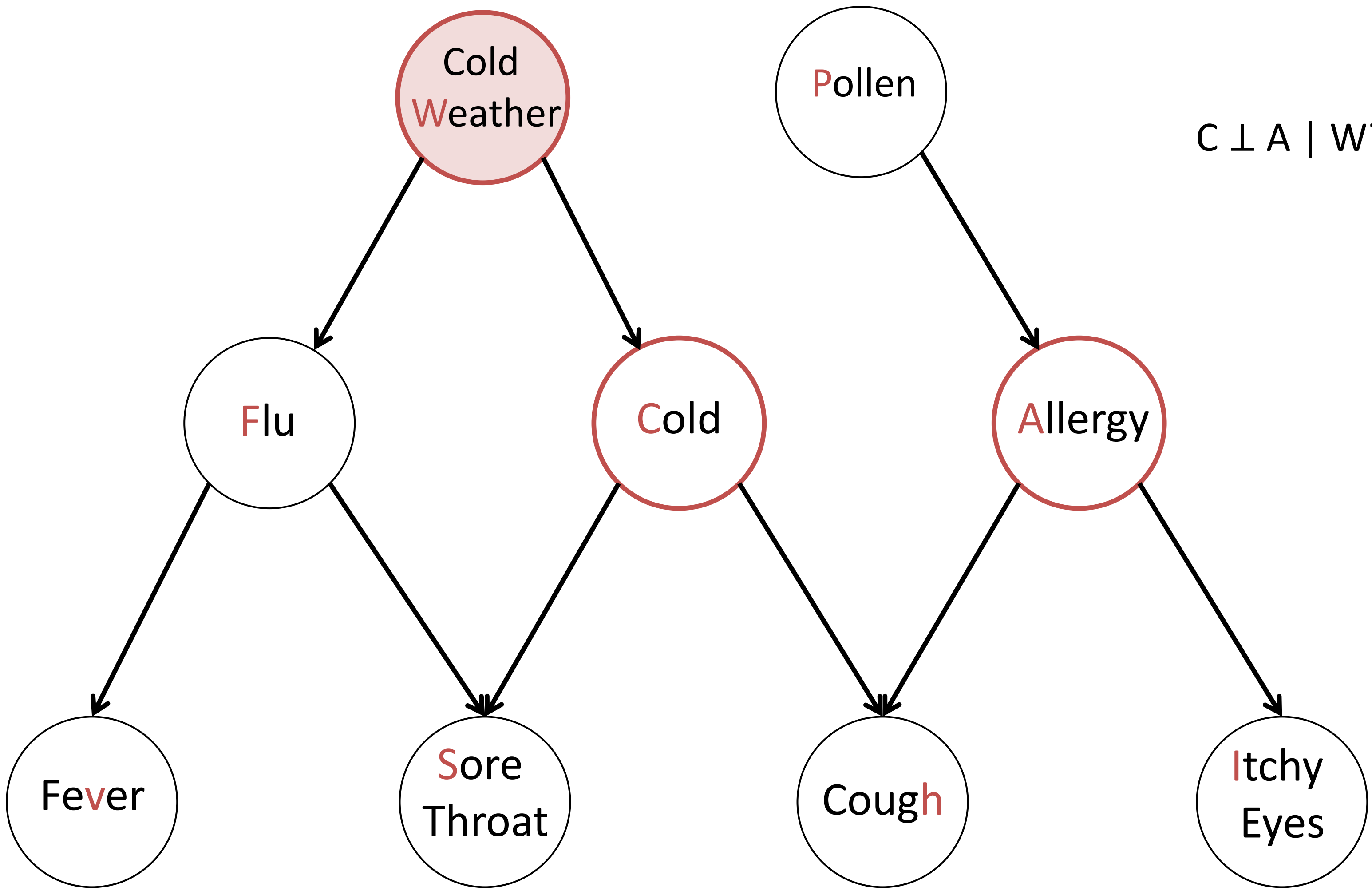
C ⊥ A ? Yes
H ⊥ A ? No
V ⊥ H ? No

Conditional Independence



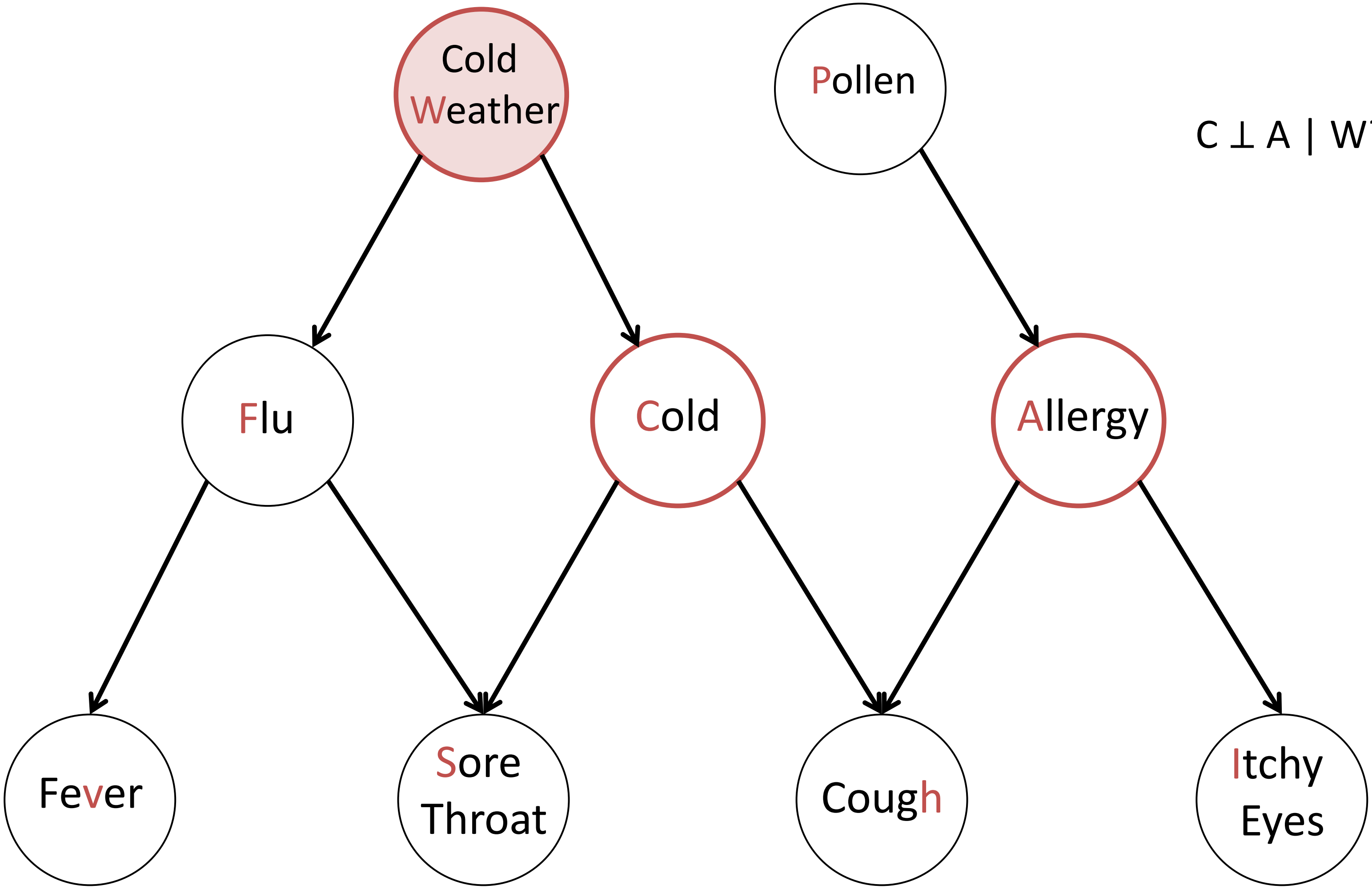
$C \perp A \mid W$?
$C \perp F \mid W$?
$V \perp H \mid W$?
$C \perp V \mid W$?

Conditional Independence



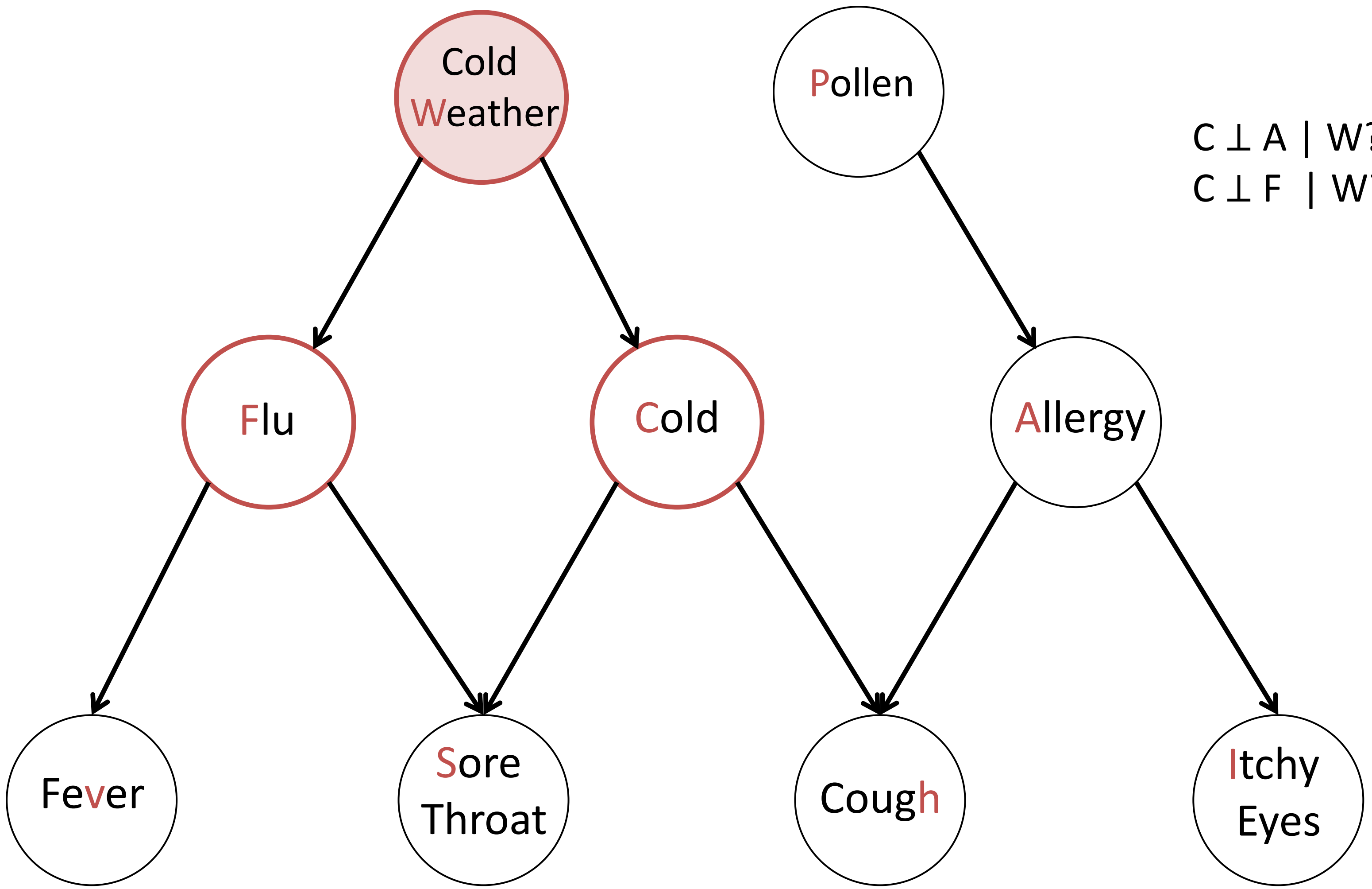
C ⊥ A | W?

Conditional Independence



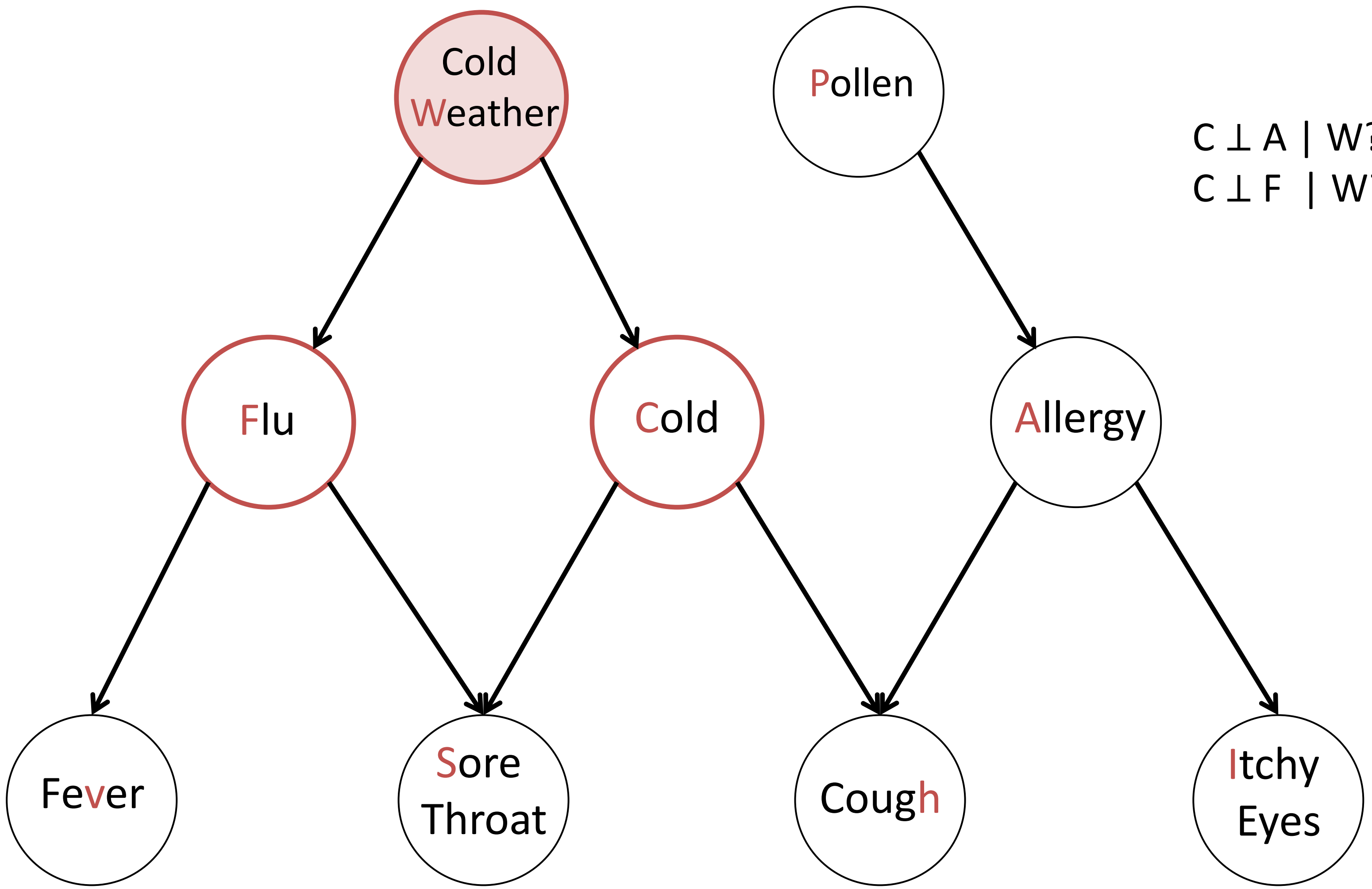
C ⊥ A | W? Yes

Conditional Independence



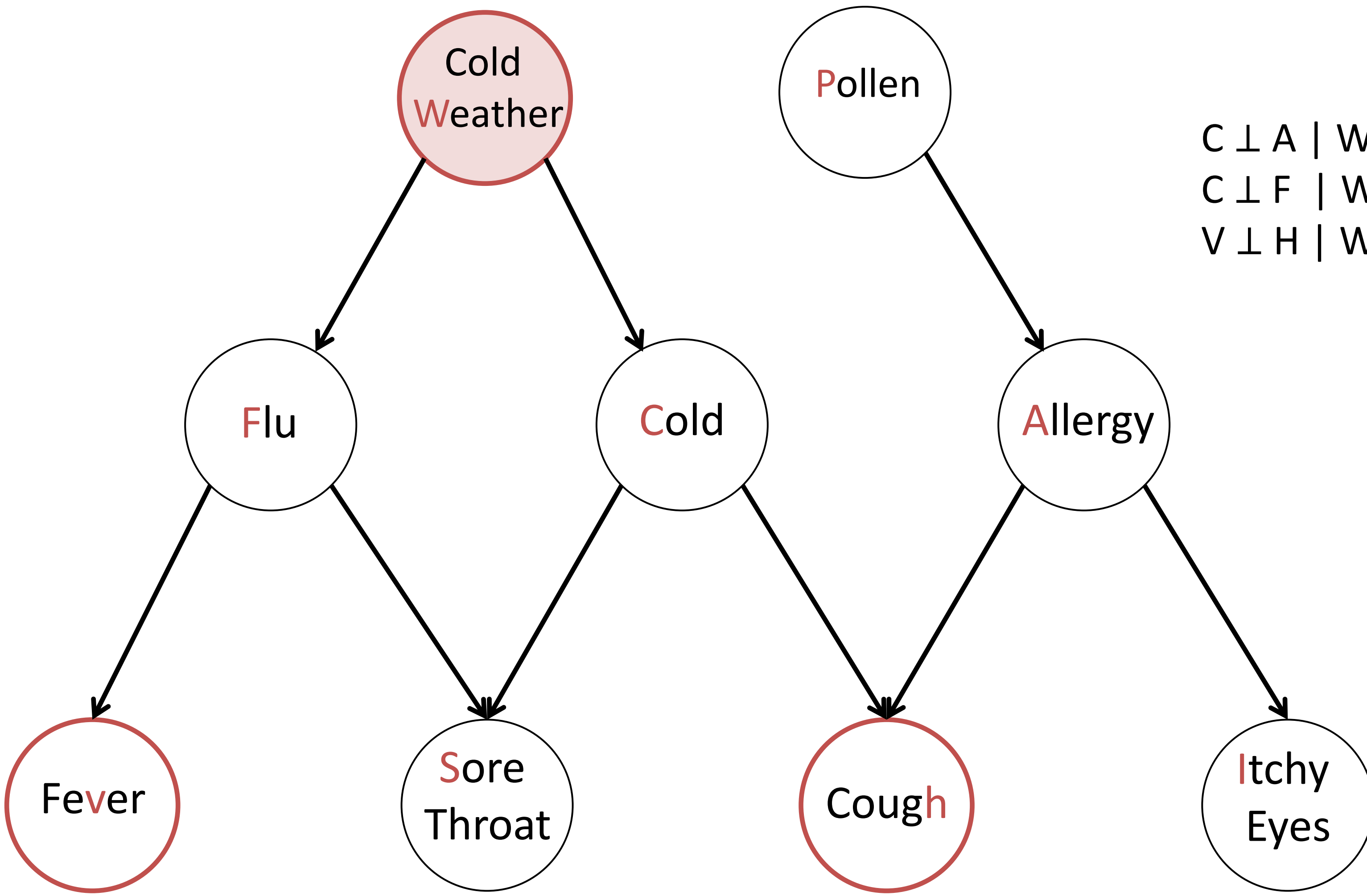
C ⊥ A | W? Yes
C ⊥ F | W?

Conditional Independence



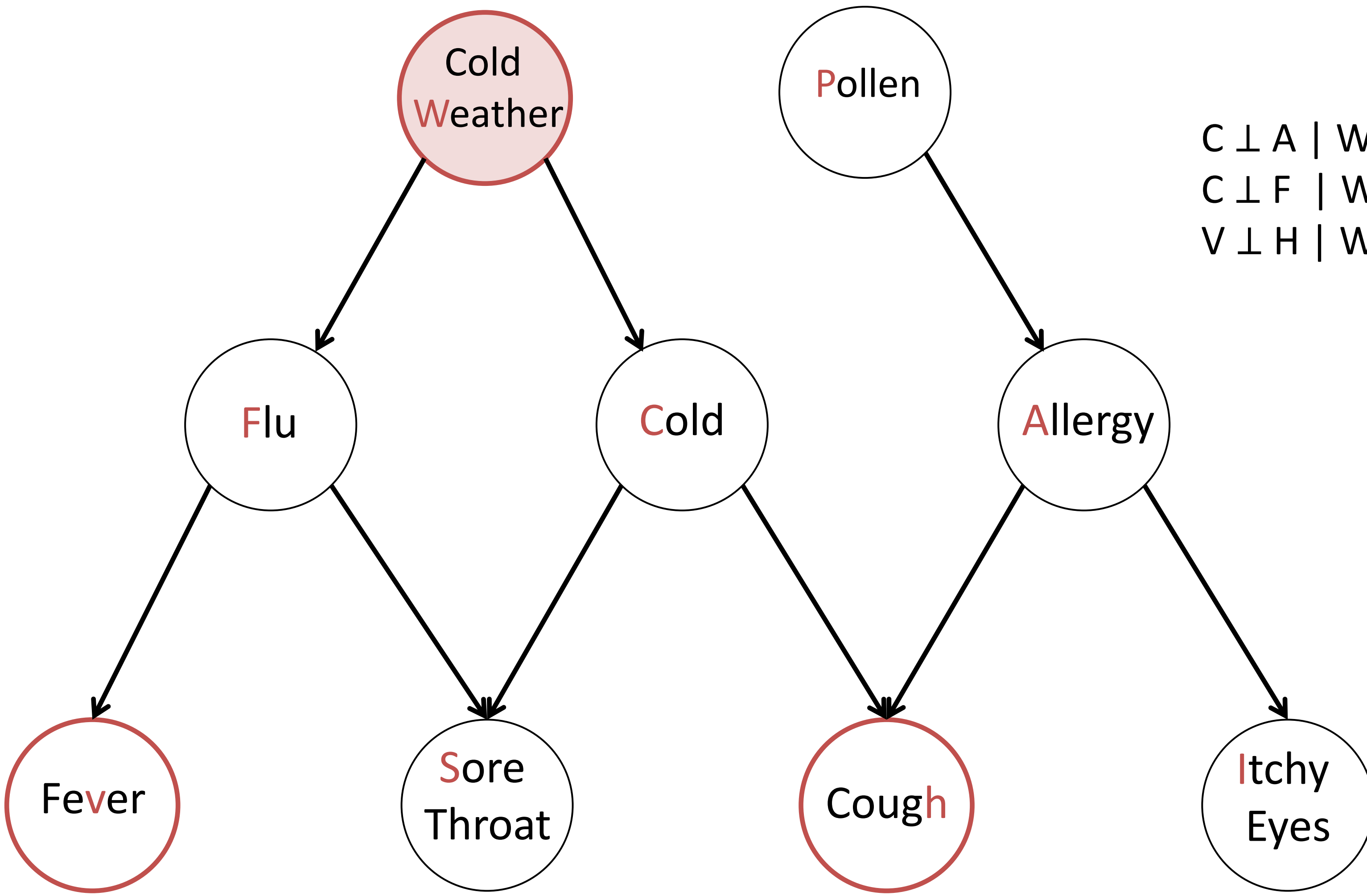
C ⊥ A | W? Yes
C ⊥ F | W? Yes

Conditional Independence



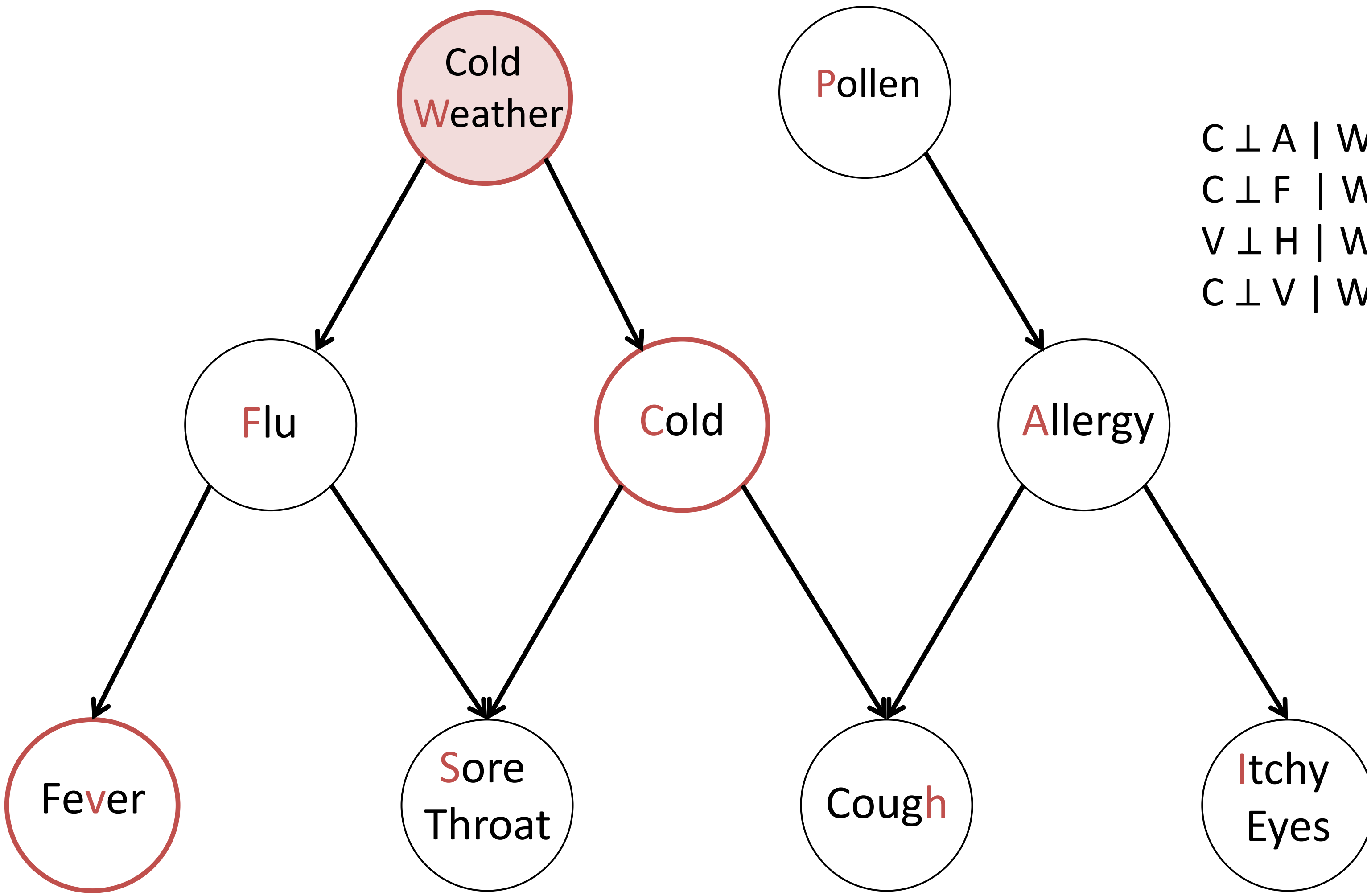
C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W?

Conditional Independence



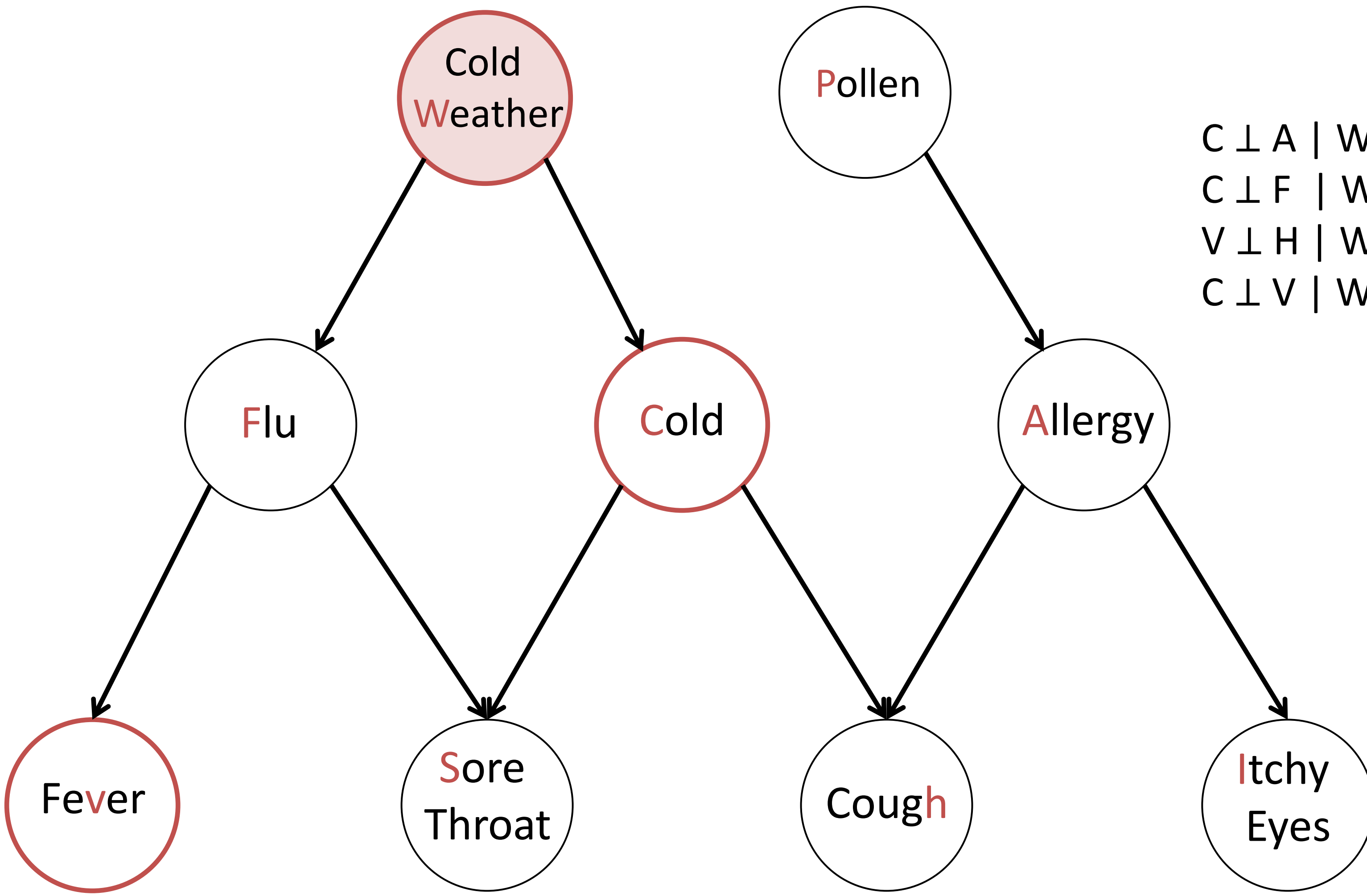
C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W? Yes

Conditional Independence



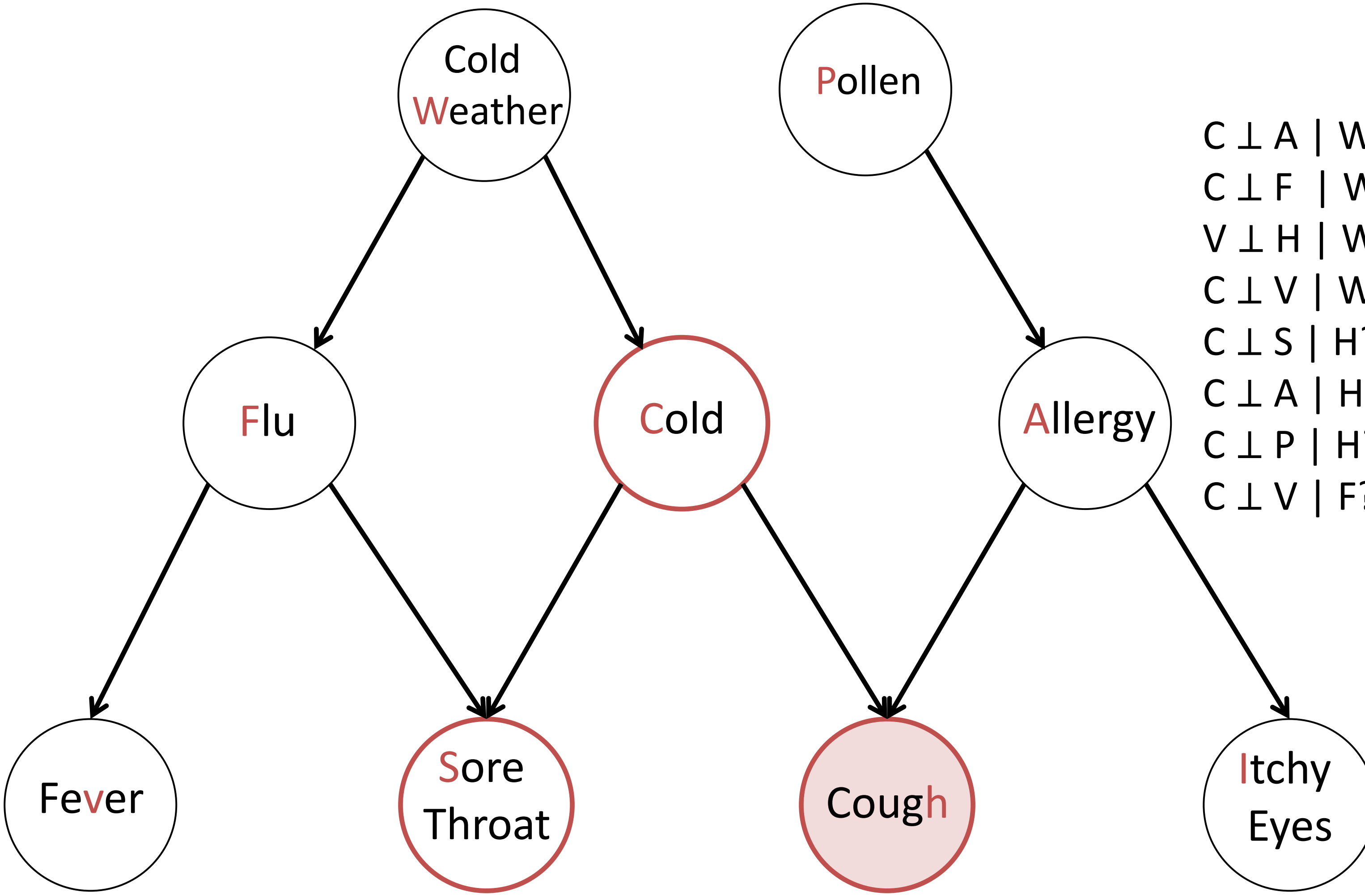
C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W? Yes
C ⊥ V | W?

Conditional Independence



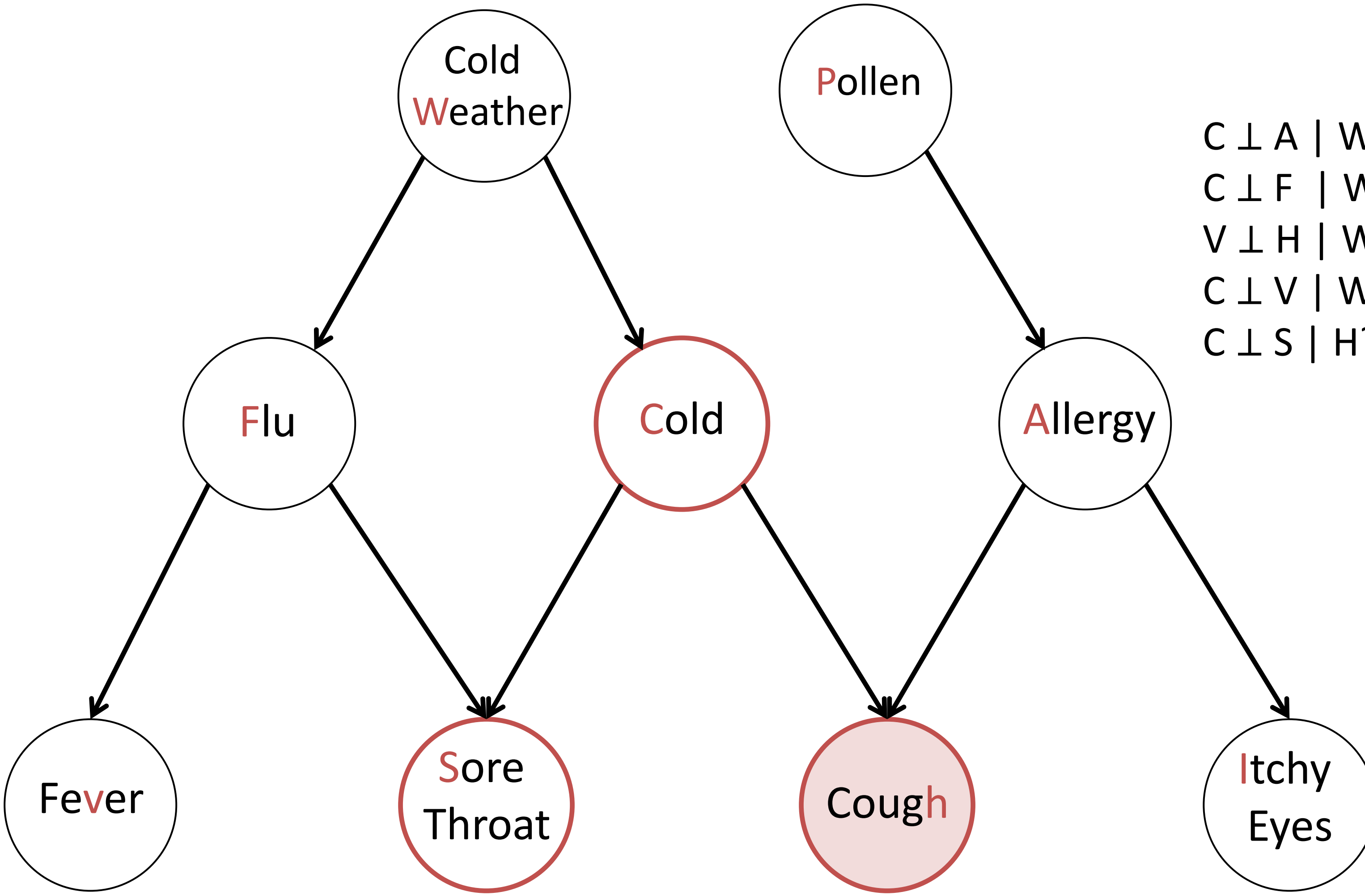
C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W? Yes
C ⊥ V | W? Yes

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
$C \perp S \mid H$?
$C \perp A \mid H$?
$C \perp P \mid H$?
$C \perp V \mid F$?

Conditional Independence



$C \perp A \mid W$? Yes

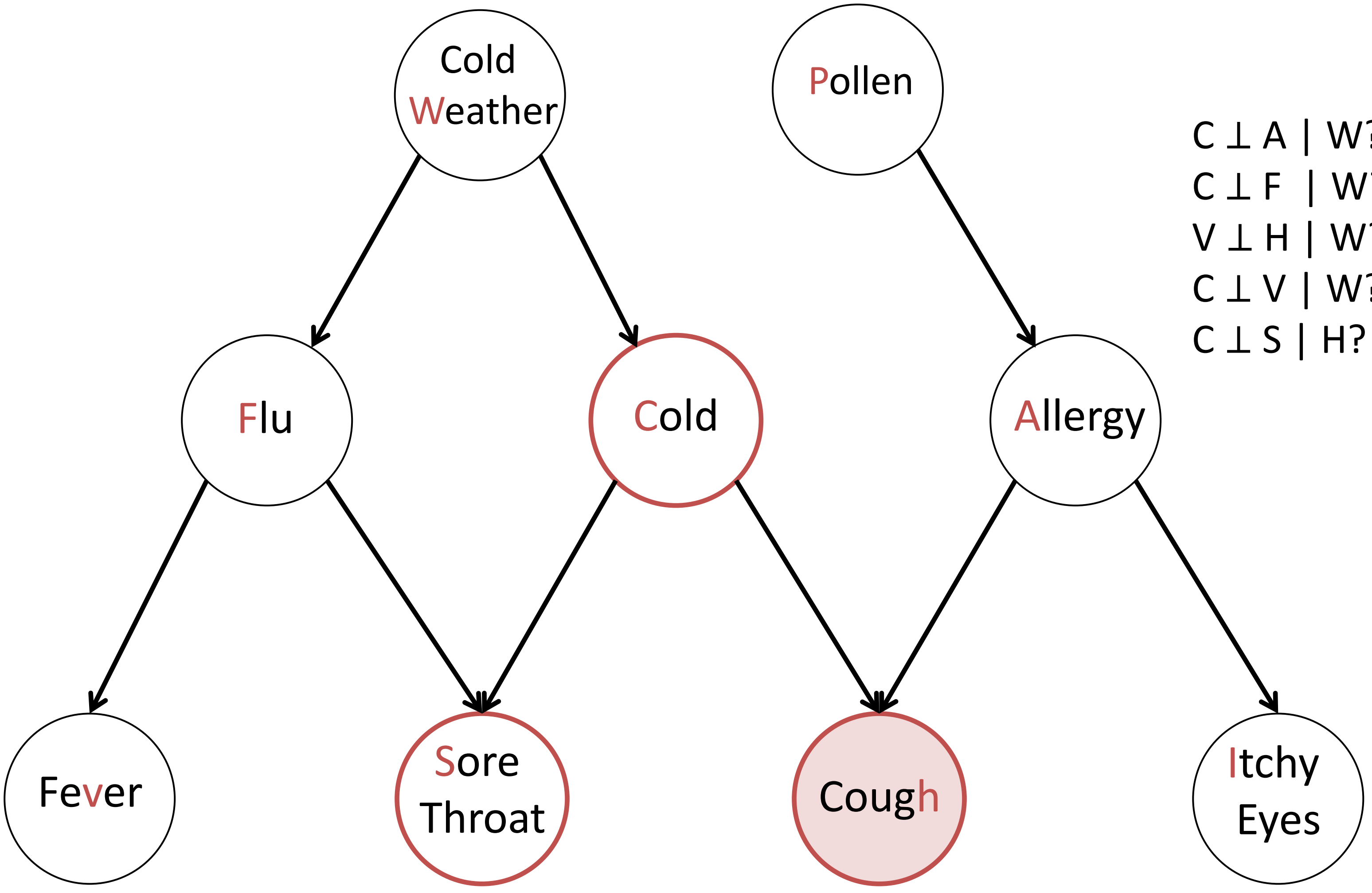
$C \perp F \mid W$? Yes

$V \perp H \mid W$? Yes

$C \perp V \mid W$? Yes

$C \perp S \mid H$?

Conditional Independence



C ⊥ A | W? Yes

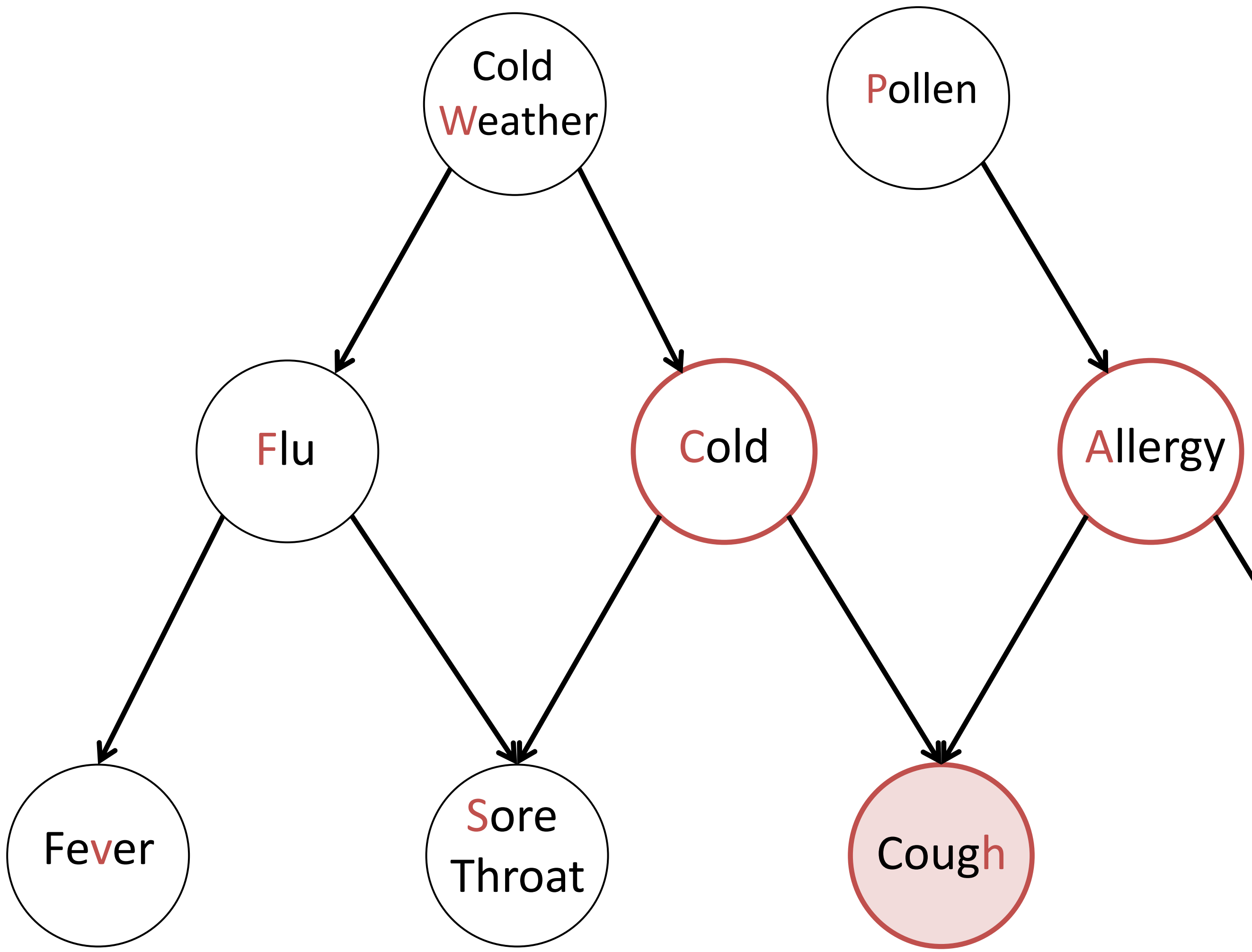
C ⊥ F | W? Yes

V ⊥ H | W? Yes

C ⊥ V | W? Yes

C ⊥ S | H? No

Conditional Independence



$C \perp A \mid W$? Yes

$C \perp F \mid W$? Yes

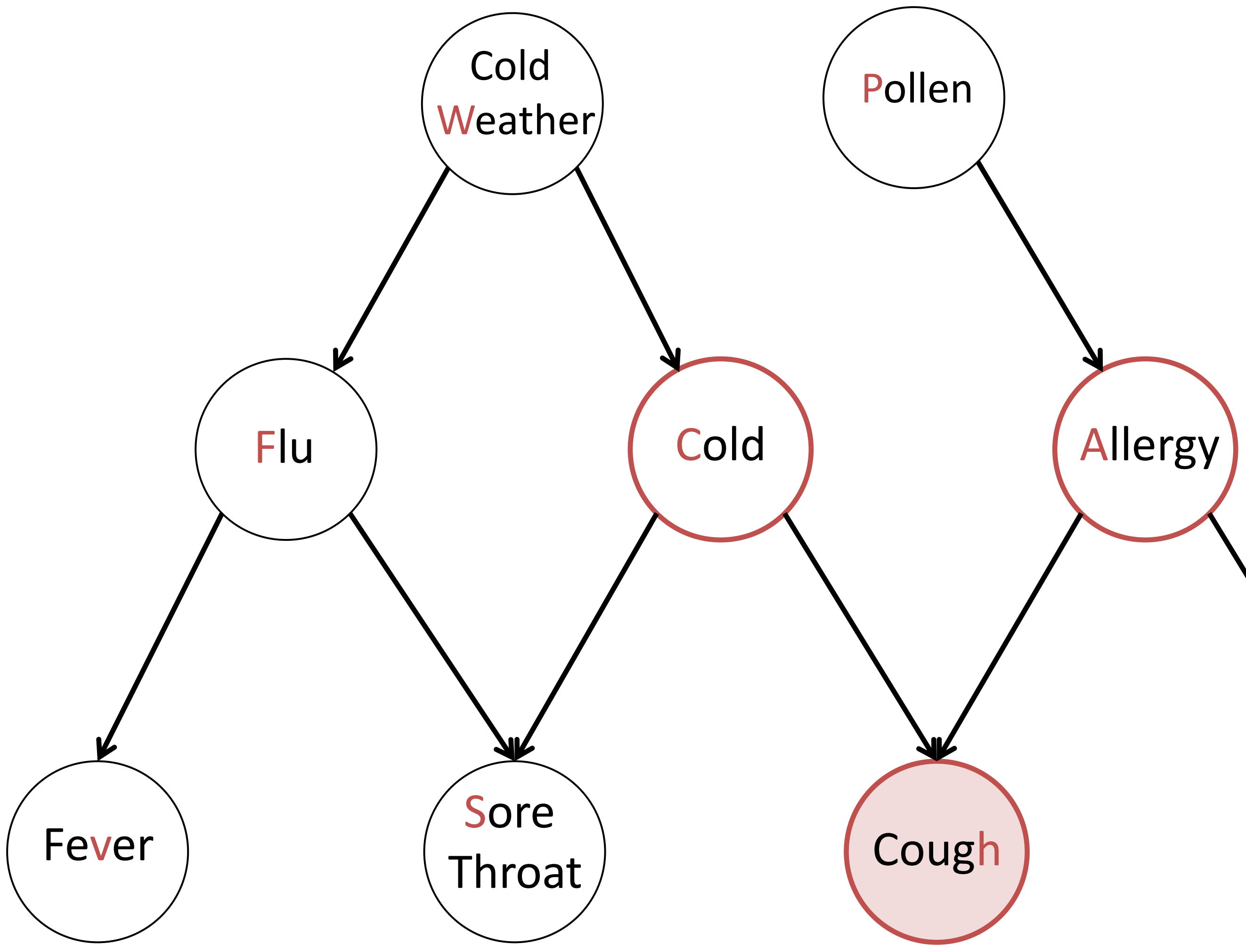
$V \perp H \mid W$? Yes

$C \perp V \mid W$? Yes

$C \perp S \mid H$? No

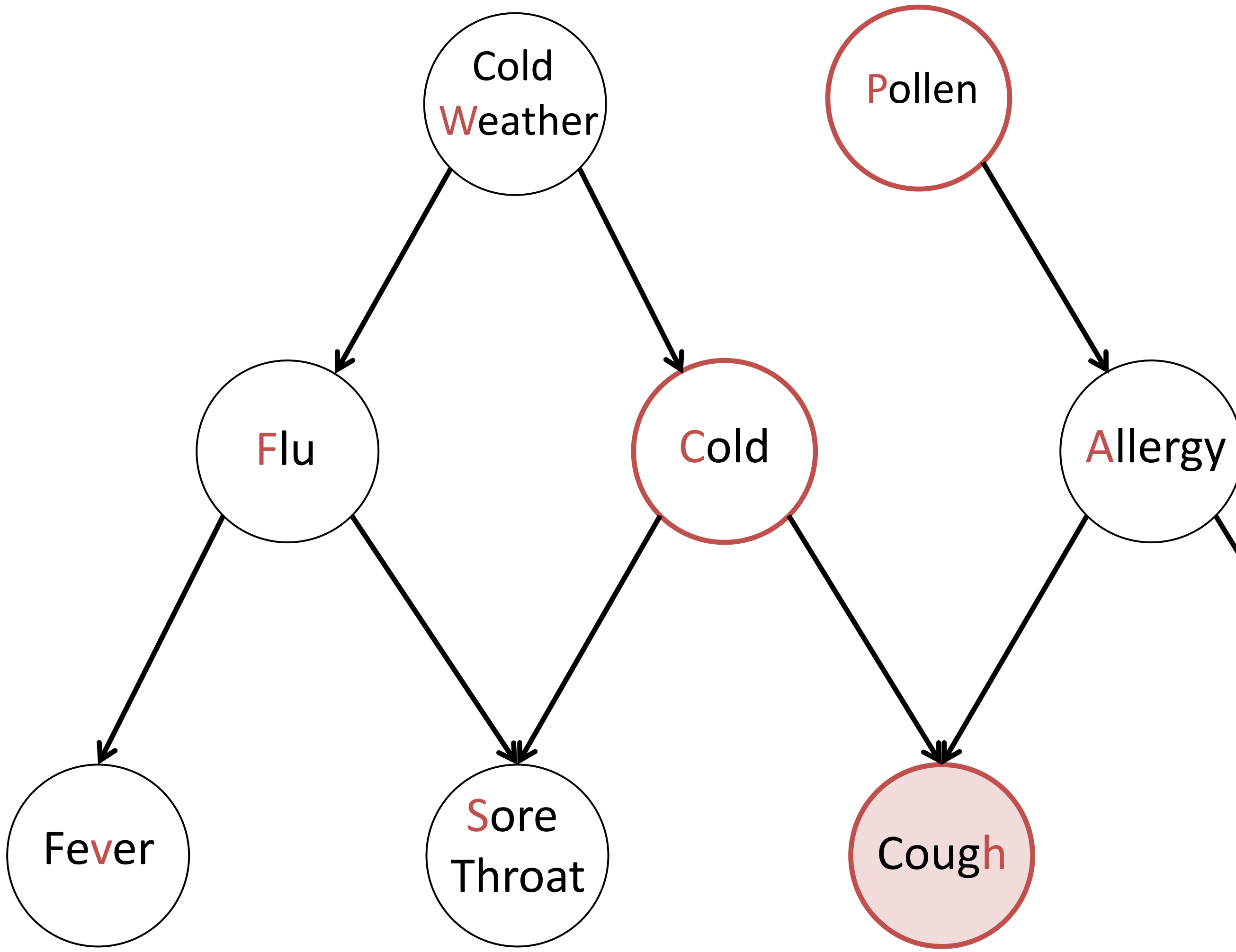
$C \perp A \mid H$?

Conditional Independence



C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W? Yes
C ⊥ V | W? Yes
C ⊥ S | H? No
C ⊥ A | H? No

Conditional Independence



$C \perp A \mid W$? Yes

$C \perp F \mid W$? Yes

$V \perp H \mid W$? Yes

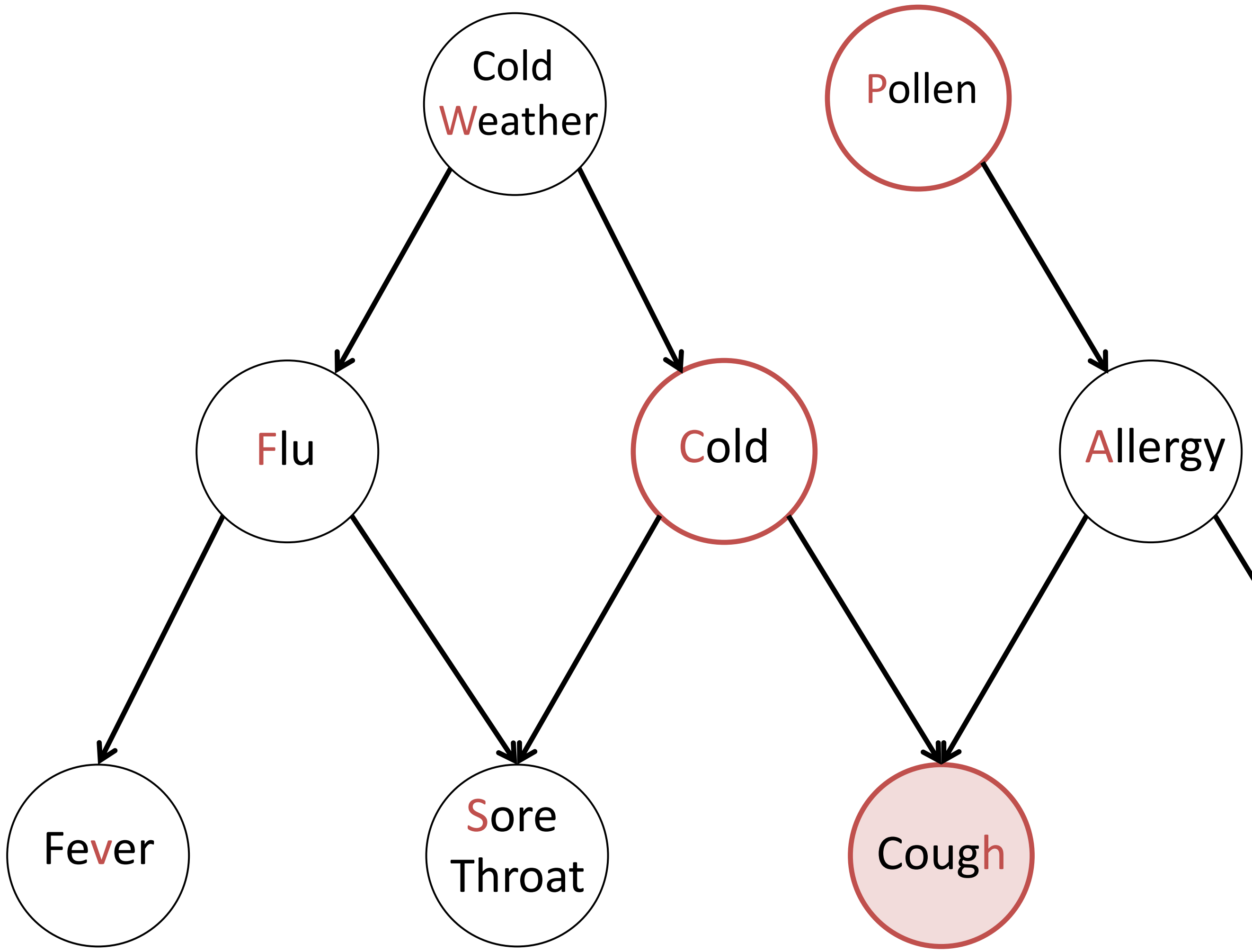
$C \perp V \mid W$? Yes

$C \perp S \mid H$? No

$C \perp A \mid H$? No

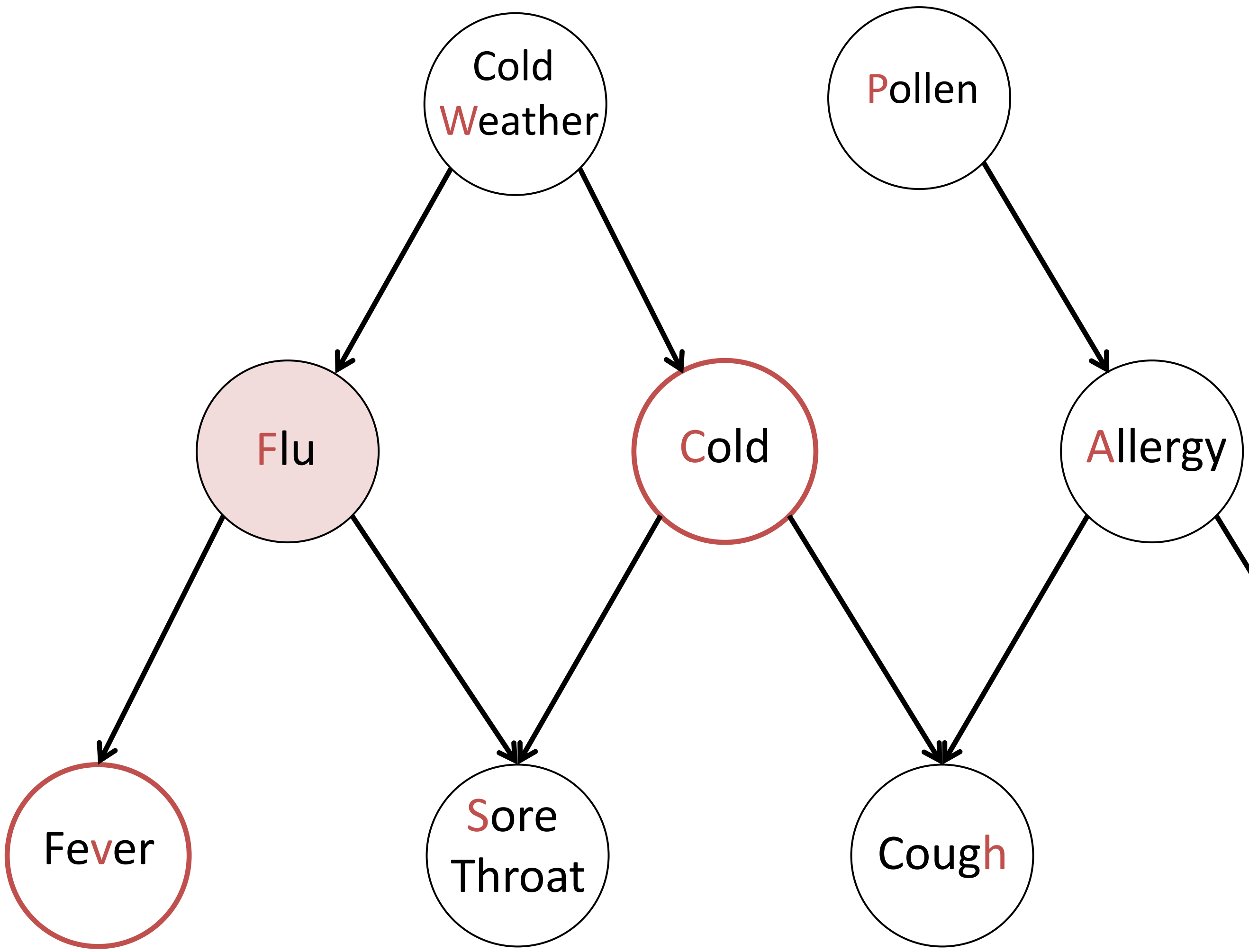
$C \perp P \mid H$?

Conditional Independence



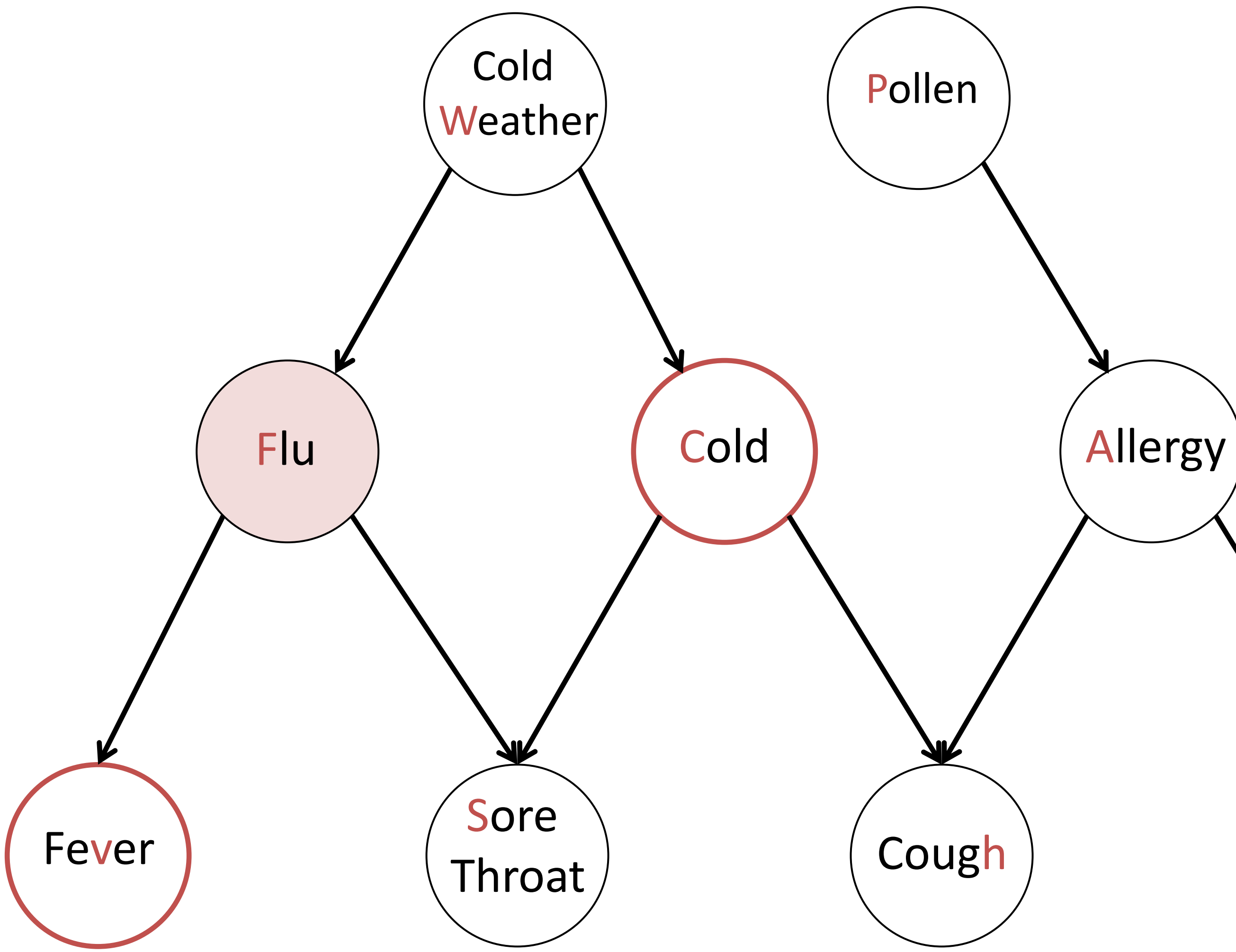
C ⊥ A | W? Yes
C ⊥ F | W? Yes
V ⊥ H | W? Yes
C ⊥ V | W? Yes
C ⊥ S | H? No
C ⊥ A | H? No
C ⊥ P | H? No

Conditional Independence



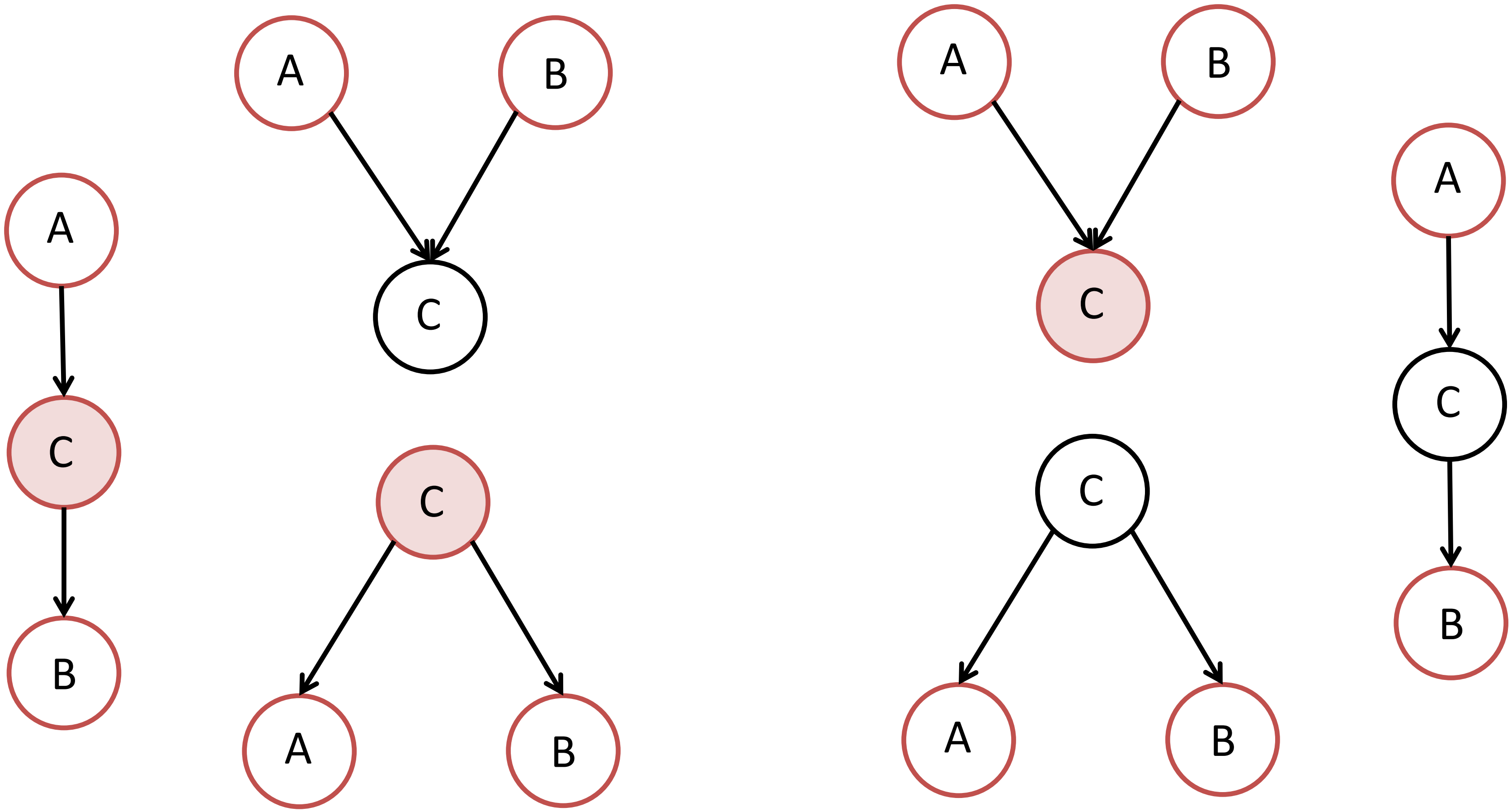
$C \perp A \mid W$? Yes
$C \perp F \mid W$? Yes
$V \perp H \mid W$? Yes
$C \perp V \mid W$? Yes
$C \perp S \mid H$? No
$C \perp A \mid H$? No
$C \perp P \mid H$? No
$C \perp V \mid F$?

Conditional Independence



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

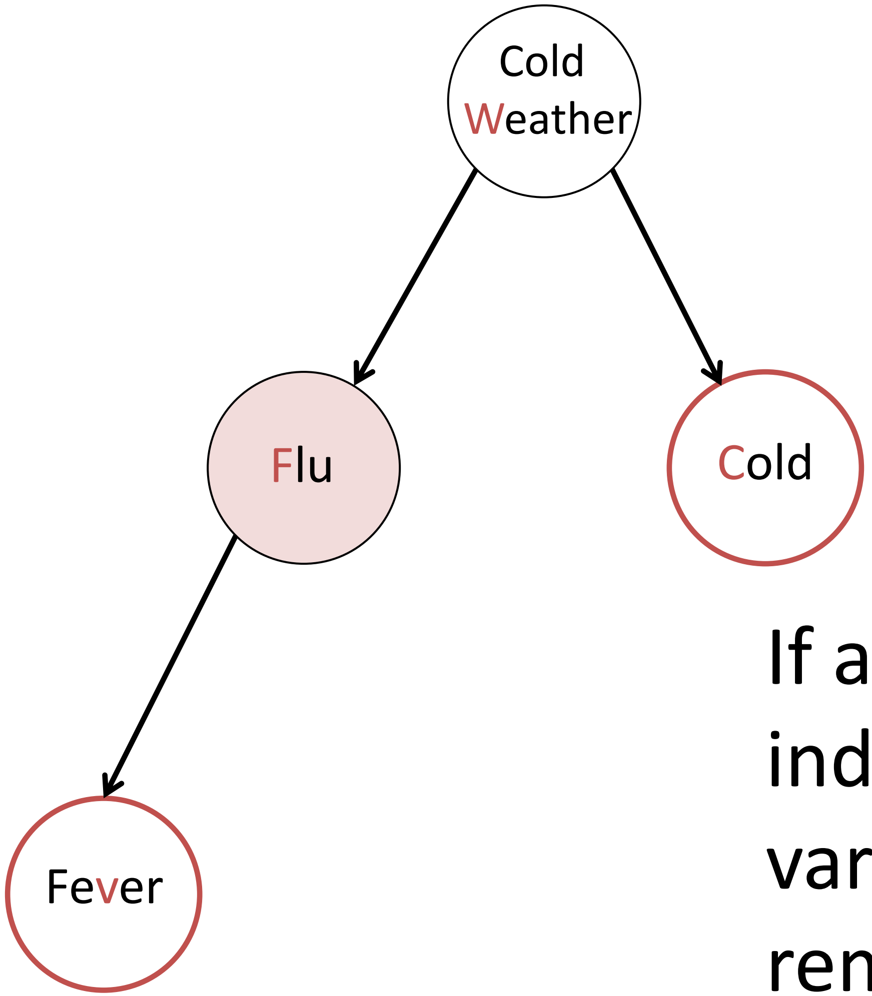
Patterns



Independent

Dependent

Conditional Independence



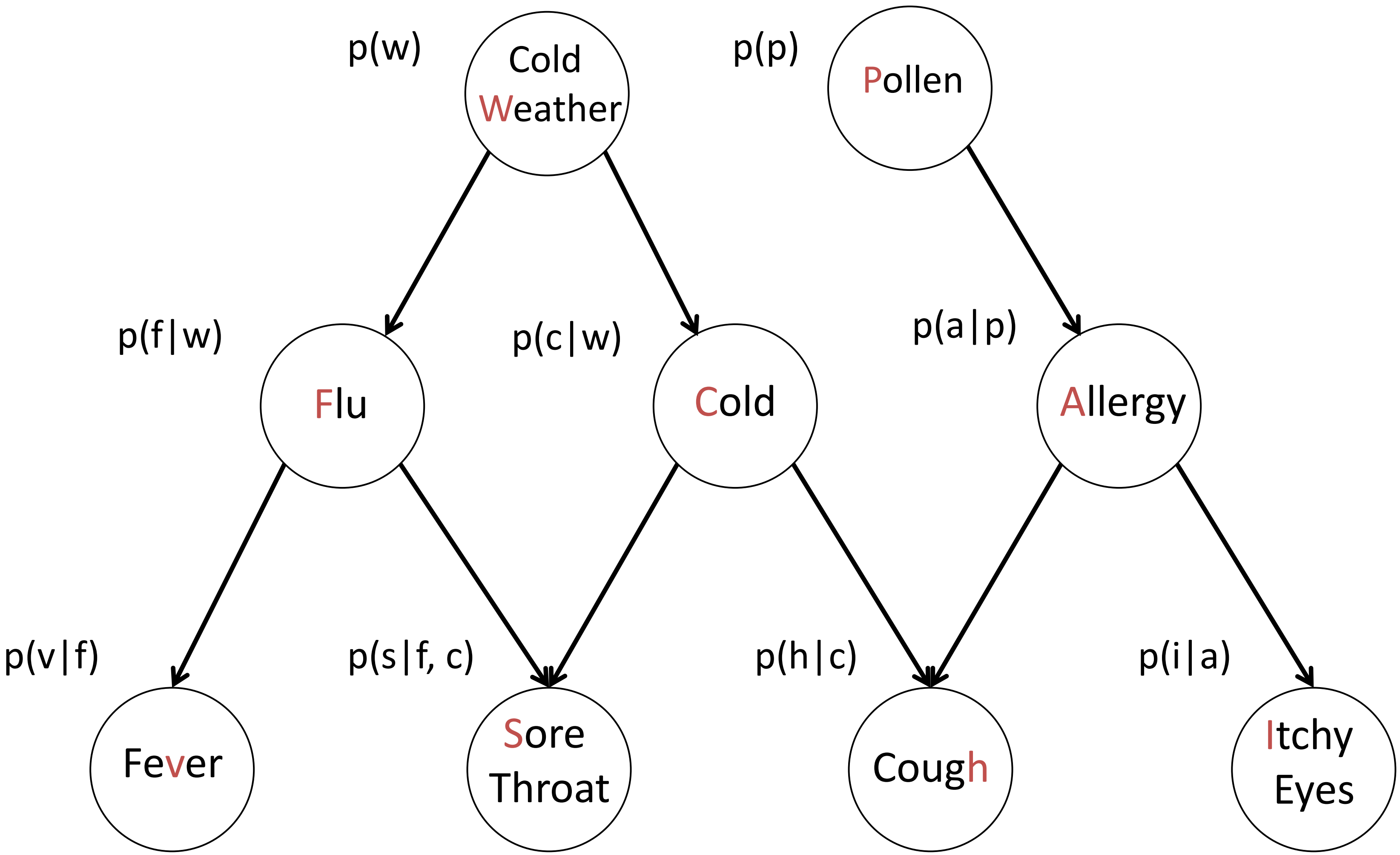
P(C=c|F=f) = ?

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

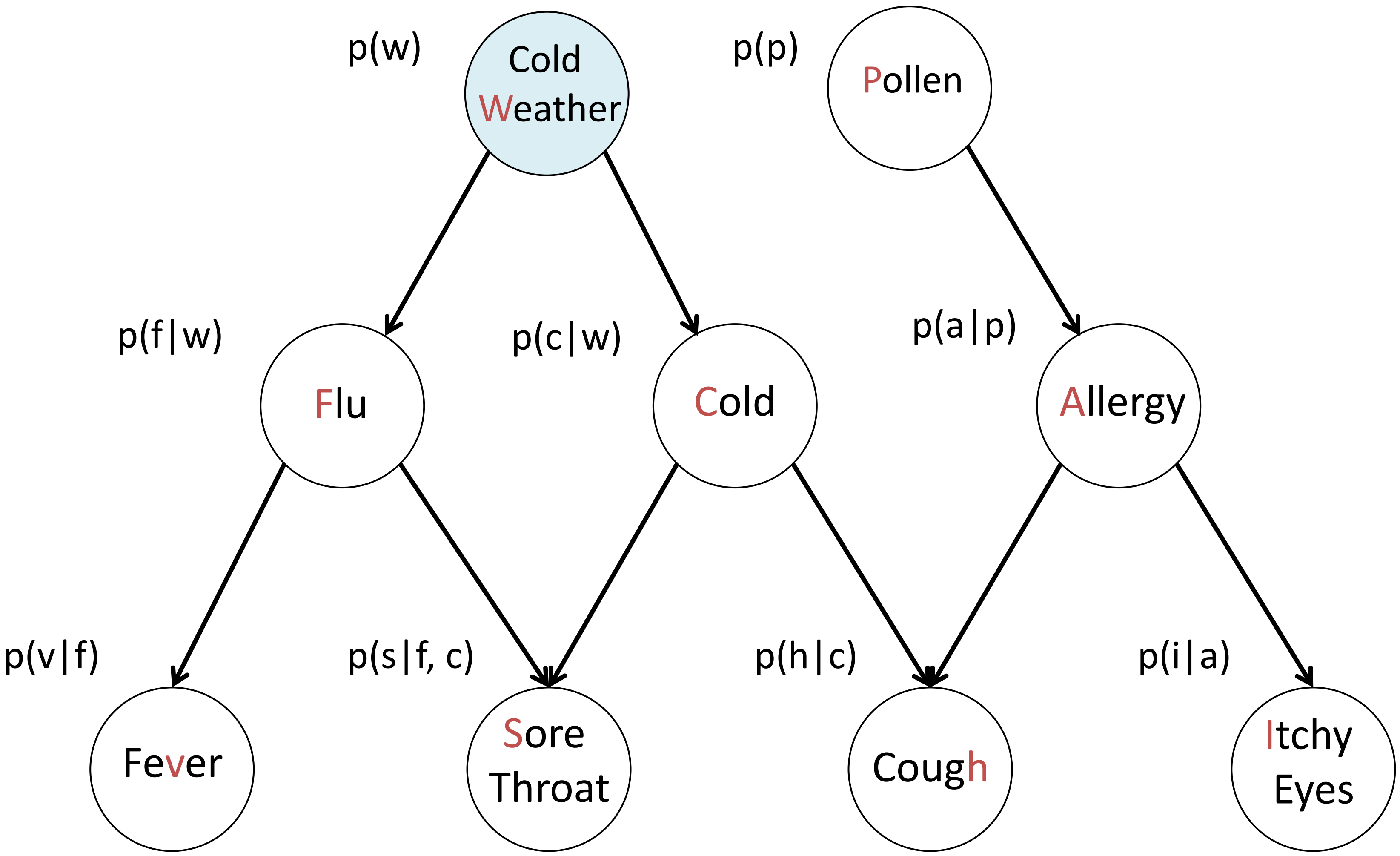
Roadmap

- Bayesian Networks Introduction
- Probabilistic Queries
- Conditional Independence
- Gibbs Sampling

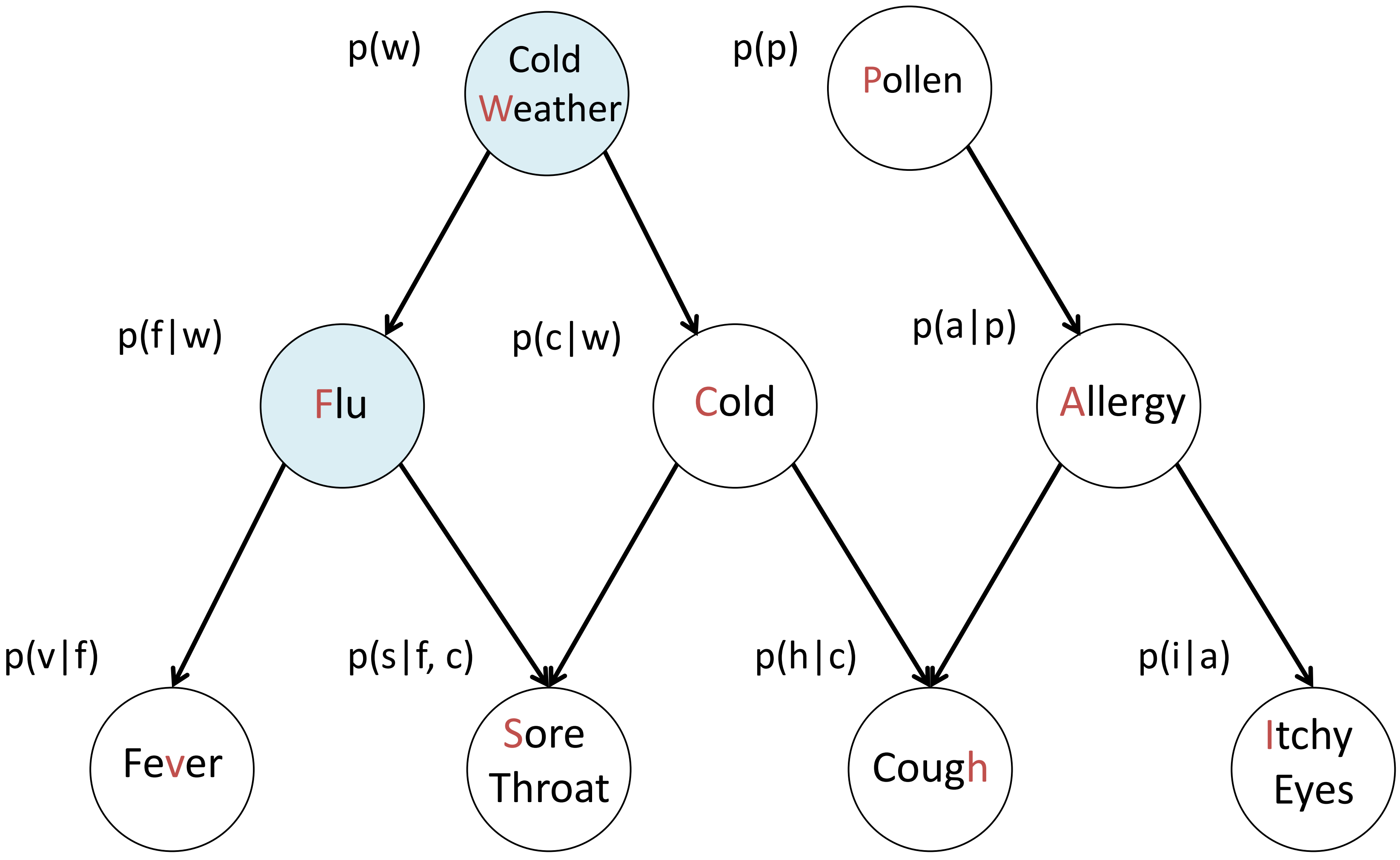
Sample 1M samples from joint distribution



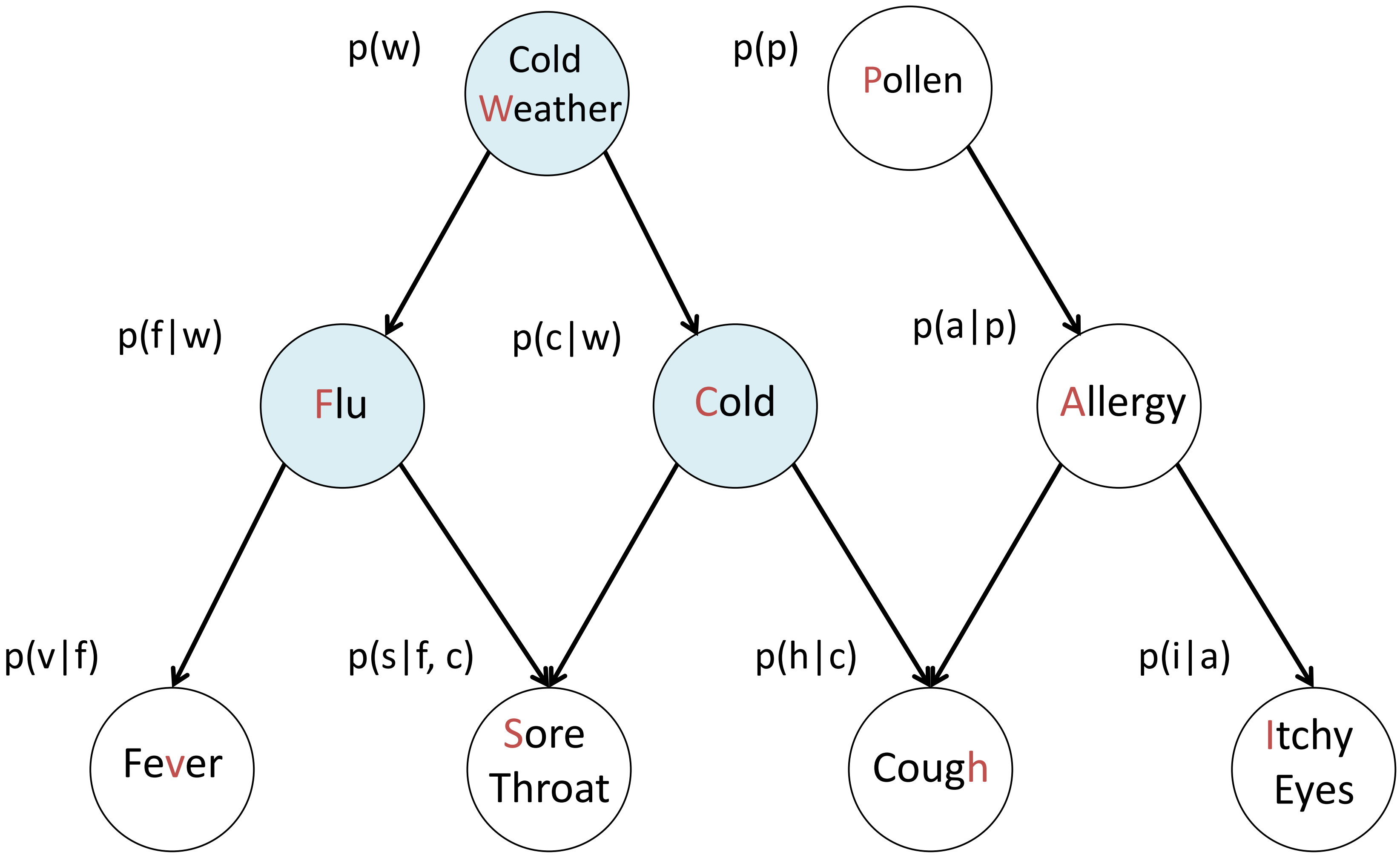
Forward Sampling



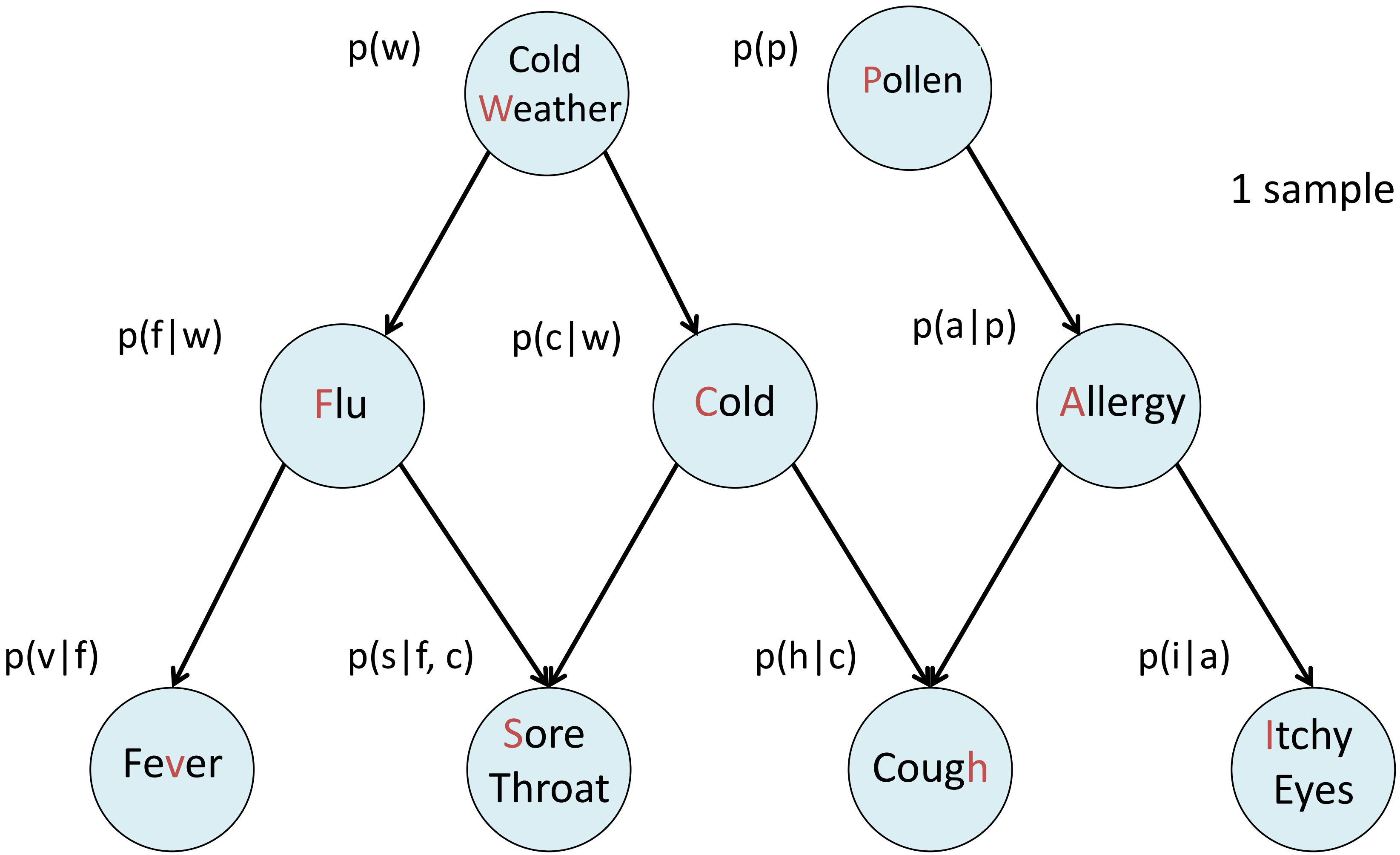
Forward Sampling



Forward Sampling



Forward Sampling



Gibbs Sampling

Algorithm: Gibbs sampling

Initialize x to a random complete assignment

Loop through $i = 1, \ldots, n$ until convergence:

  for each v, compute weight of $\{X_i : v\} \cup x \setminus \{ x_i\}$

  Choose $\{X_i : v\} \cup x \setminus \{ x_i\}$ with prob prop. to weight

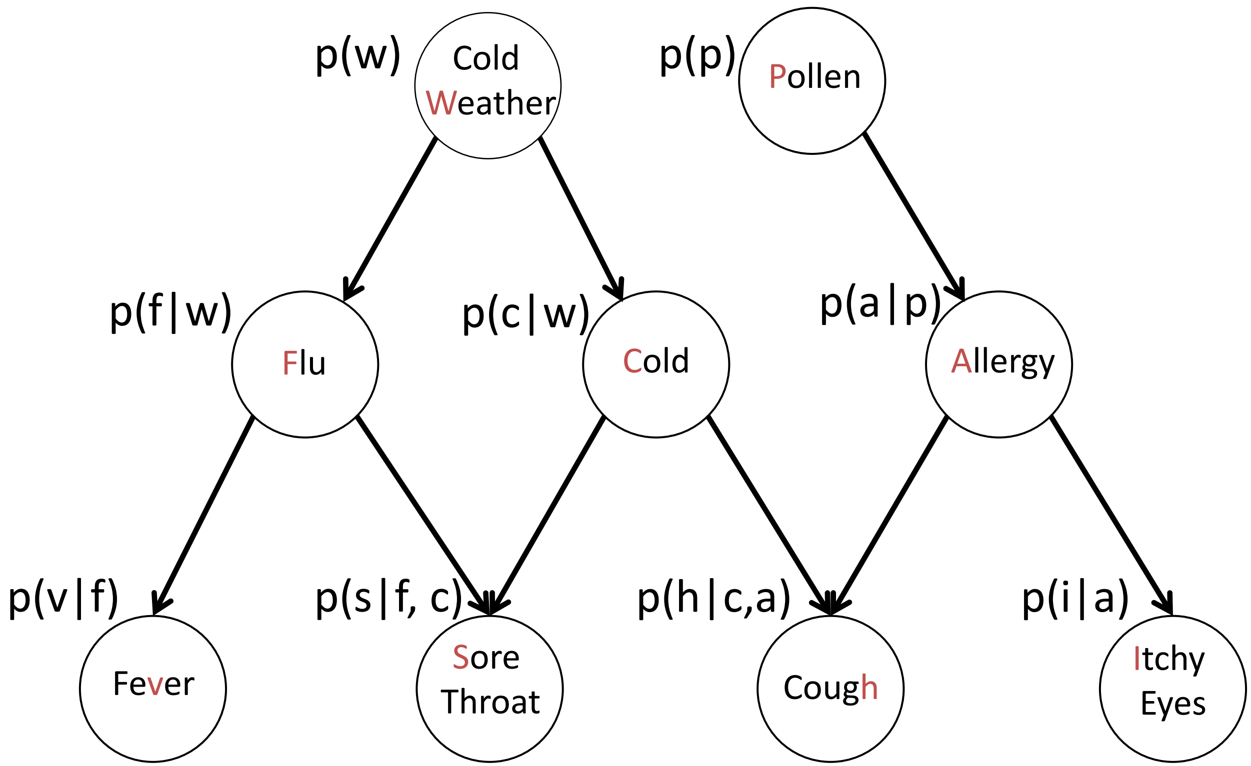
Gibbs sampling (probabilistic interpretation)

Loop through $i = 1, \ldots, n$ until convergence:

  Set $X_i = v$ with prob. $\mathbb{P}(X_i = v \mid X_{-i} = x_{-i})$

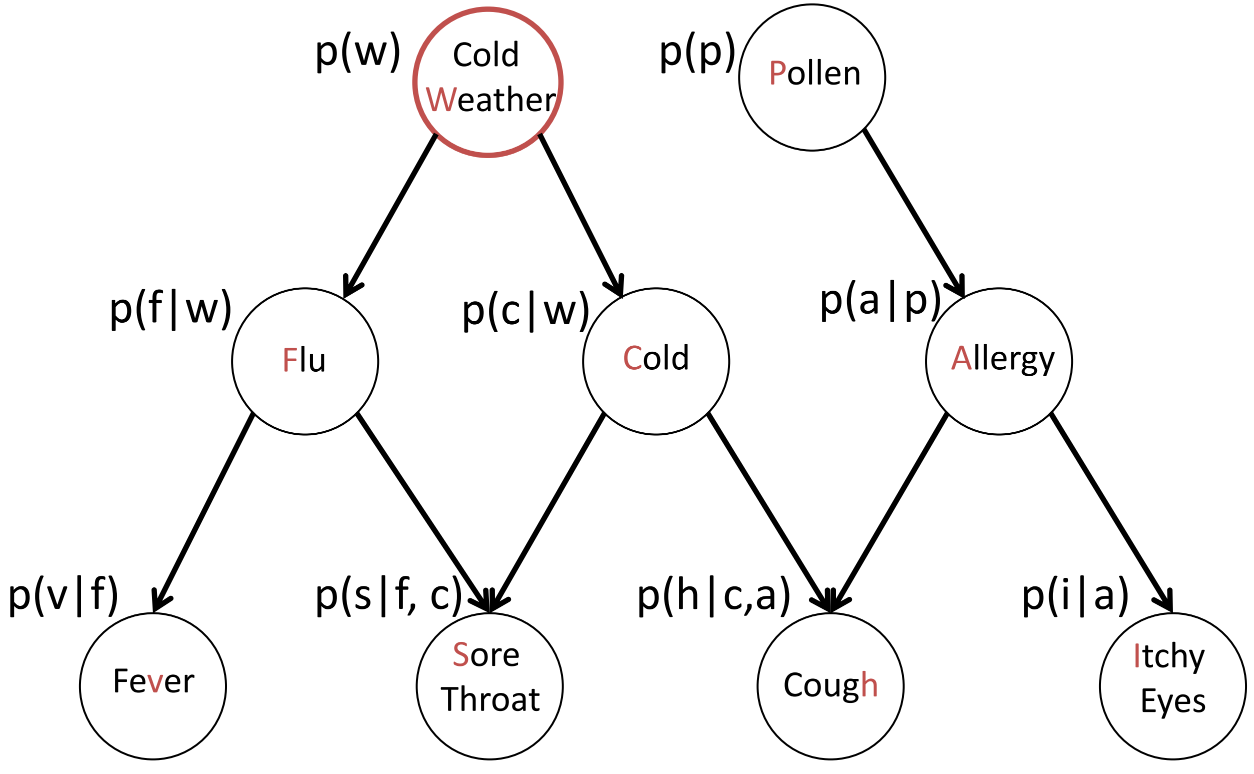
  (notation: $X_{-i} = X \setminus \{X_i\}$)

Gibbs Sampling



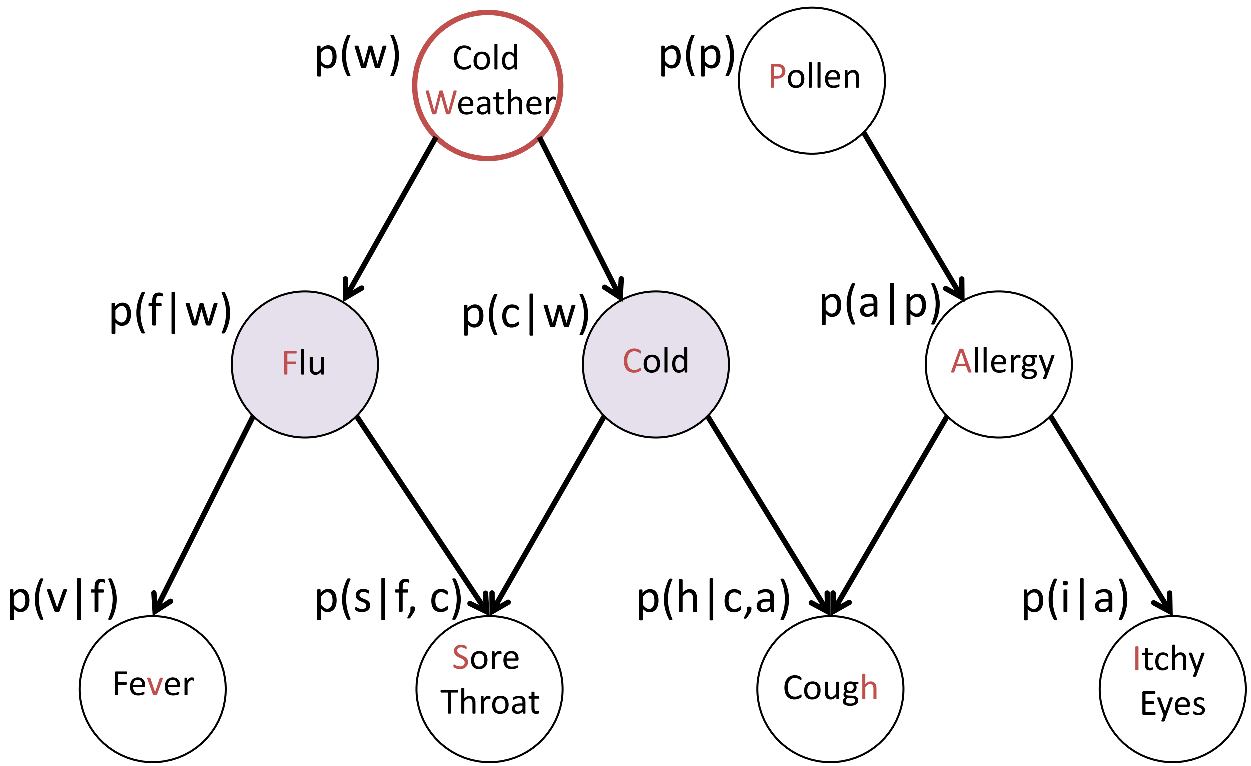
Gibbs Sampling

How do we sample a new value for W?



Gibbs Sampling

How do we sample a new value for W?

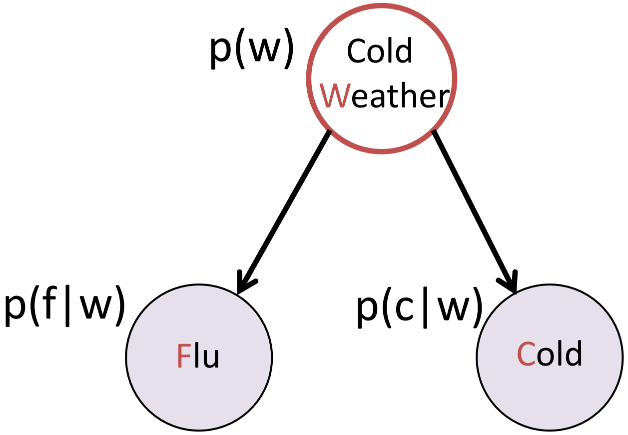


$$P(W{=}w|F{=}1, P{=}1, C{=}0, \ldots, I{=}0)$$

$$= P(W{=}w|F{=}1, C{=}0)$$

Markov Blanket!

Gibbs Sampling

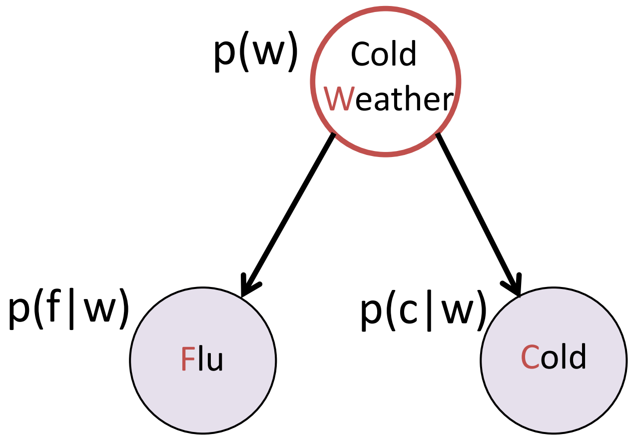


w	p(w)
0	0.4
1	0.6

w	f	p(f|w)
0	0	0.95
0	1	0.05
1	0	0.80
1	1	0.20

w	c	p(f|w)
0	0	0.88
0	1	0.12
1	0	0.70
1	1	0.30

Gibbs Sampling



w	p(w)
0	0.4
1	0.6

w	f	p(f|w)
0	0	0.95
0	1	0.05
1	0	0.80
1	1	0.20

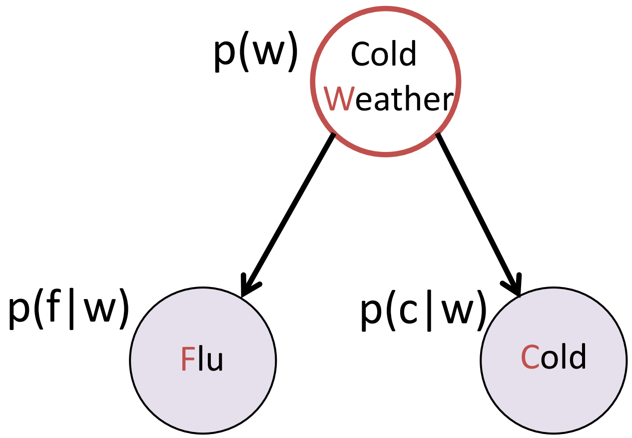
w	c	p(f|w)
0	0	0.88
0	1	0.12
1	0	0.70
1	1	0.30

$P(W=w|F=1, P=1, C=0, \ldots, I=0)$

$= P(W=w|F=1, C=0)$

µ $P(F=1|W=w)*P(C=0|W=w)*P(W=w)$

Gibbs Sampling



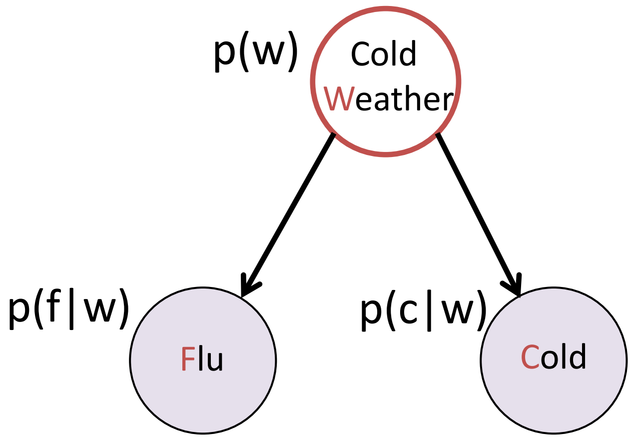
w	p(w)
0	0.40
1	0.60

w	f	p(f|w)
0	0	0.95
0	1	0.05
1	0	0.80
1	1	0.20

w	c	p(f|w)
0	0	0.88
0	1	0.12
1	0	0.70
1	1	0.30

$$P(W=w|F=1, P=1, C=0, \ldots, I=0)$$
$$= P(W=w|F=1, C=0)$$
$$\mu \; P(F=1|W=w)*P(C=0|W=w)*P(W=w)$$
$$= \quad 0.05*0.88*0.40, \quad W=0$$

Gibbs Sampling



w	p(w)
0	0.40
1	0.60

w	f	p(f|w)
0	0	0.95
0	1	0.05
1	0	0.80
1	1	0.20

w	c	p(f|w)
0	0	0.88
0	1	0.12
1	0	0.70
1	1	0.30

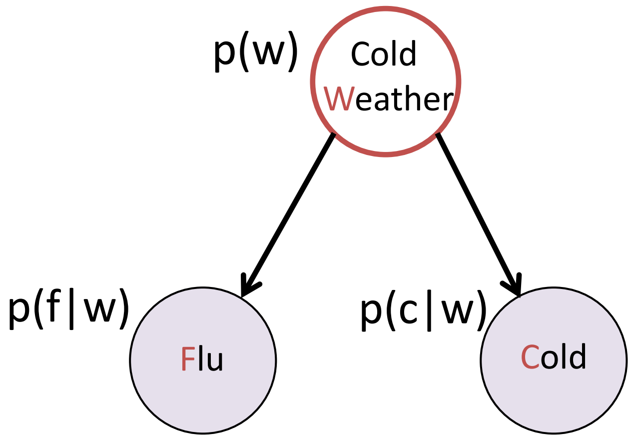
$$P(W=w|F=1, P=1, C=0, \ldots, I=0)$$

$$= P(W=w|F=1, C=0)$$

$$\mu \quad P(F=1|W=w)*P(C=0|W=w)*P(W=w)$$

$$= \begin{array}{ll} 0.05*0.88*0.40, & W=0 \\ 0.20*0.70*0.60, & W=1 \end{array}$$

Gibbs Sampling



w	p(w)
0	0.40
1	0.60

w	f	p(f|w)
0	0	0.95
0	1	0.05
1	0	0.80
1	1	0.20

w	c	p(f|w)
0	0	0.88
0	1	0.12
1	0	0.70
1	1	0.30

$$P(W=w|F=1, P=1, C=0, \ldots, I=0)$$

$$= P(W=w|F=1, C=0)$$

$$\mu \;\; P(F=1|W=w)*P(C=0|W=w)*P(W=w)$$

$$= \begin{matrix} 0.05*0.88*0.40, & W=0 \\ 0.20*0.70*0.60, & W=1 \end{matrix}$$

$$P(W=w|F=1, C=0)$$

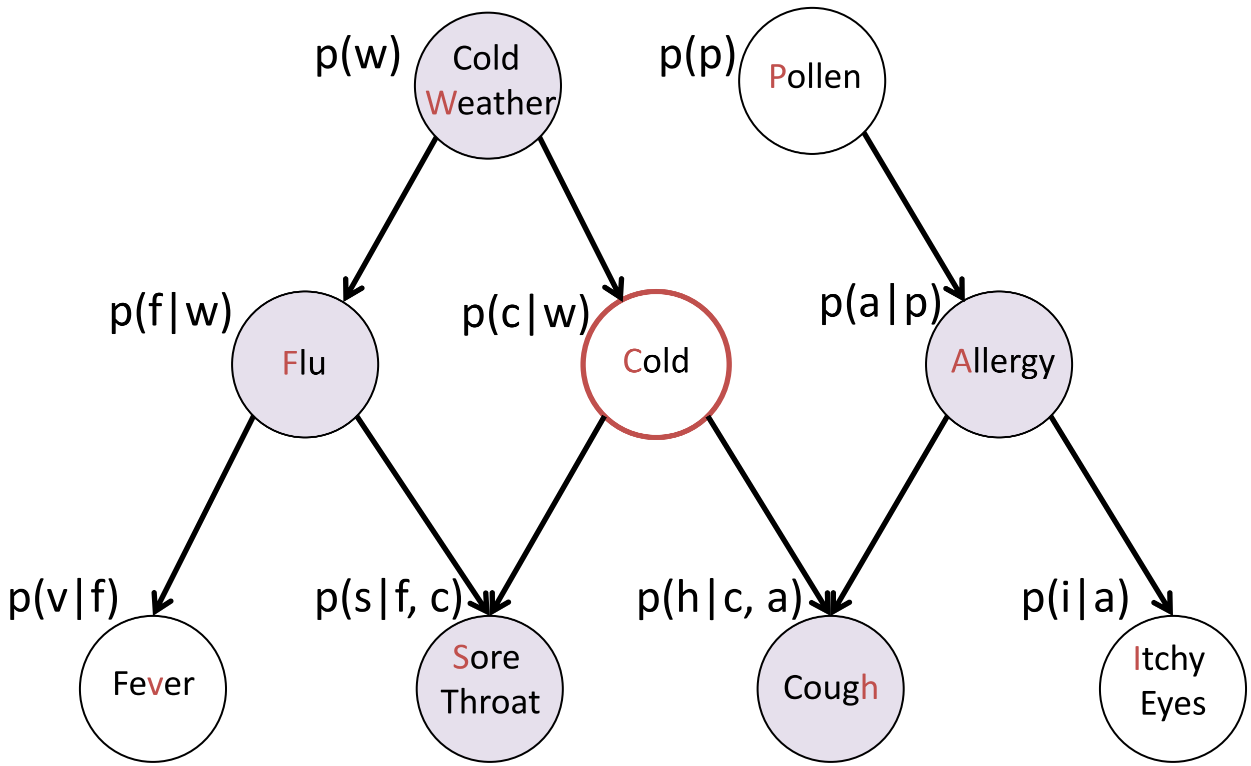
$$= \begin{matrix} 0.0176/(0.0176+0.084), & w=0 \\ 0.084/(0.0176+0.084), & w=1 \end{matrix}$$

$$= \begin{matrix} 0.173, & w=0 \\ 0.827, & w=1 \end{matrix}$$

Sample a new w!

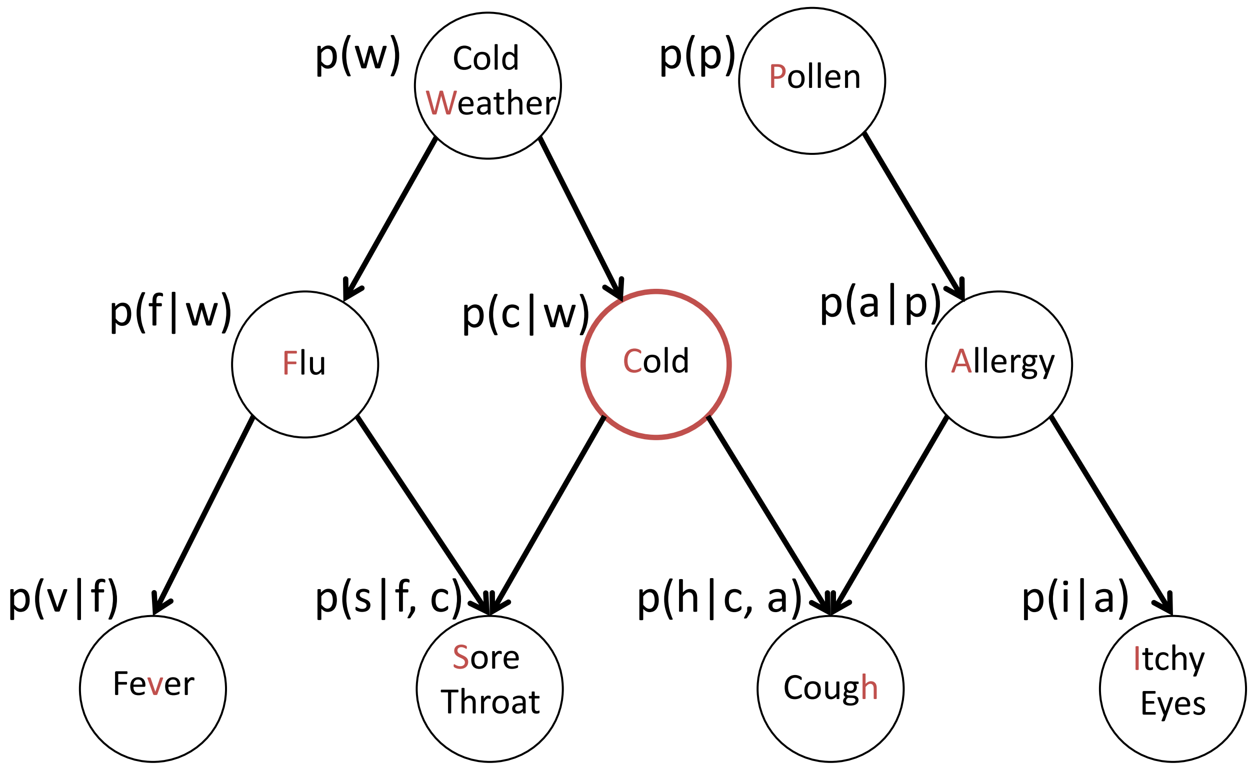
Gibbs Sampling

How do we sample a new value for C?



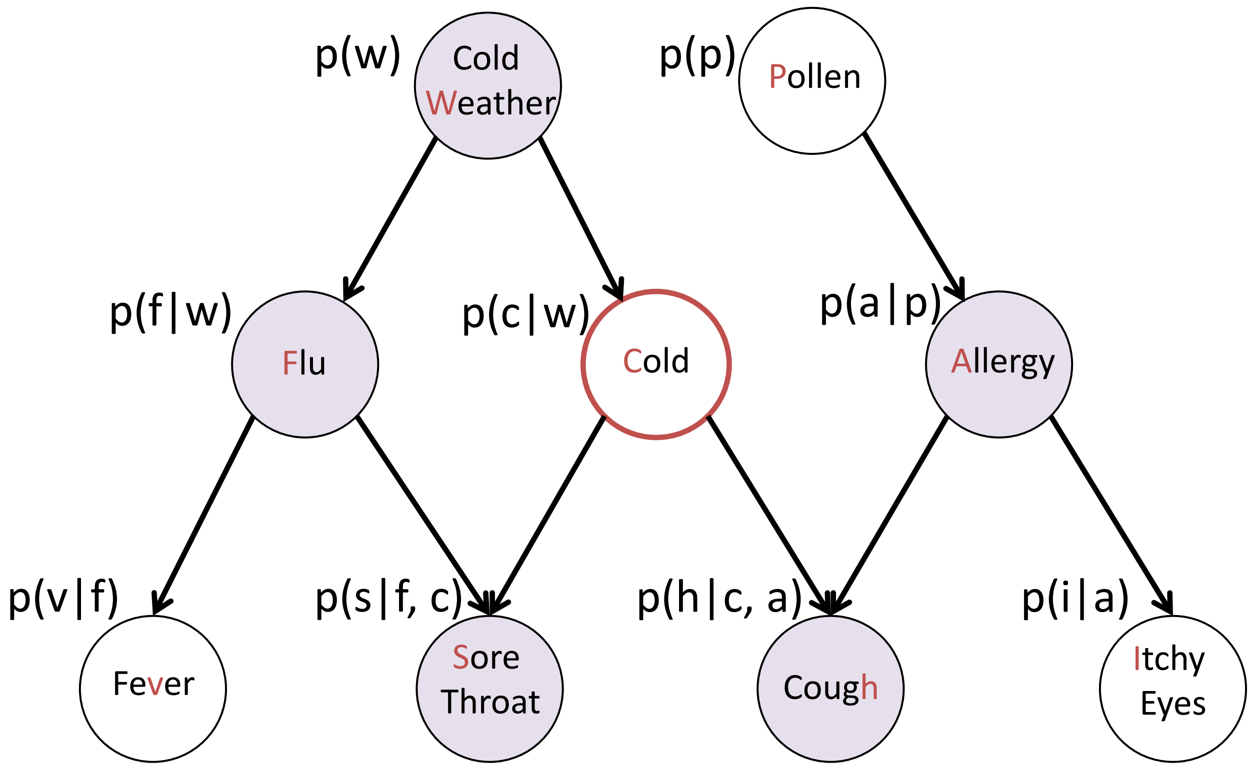
Gibbs Sampling

How do we sample a new value for C?



Gibbs Sampling

How do we sample a new value for C?

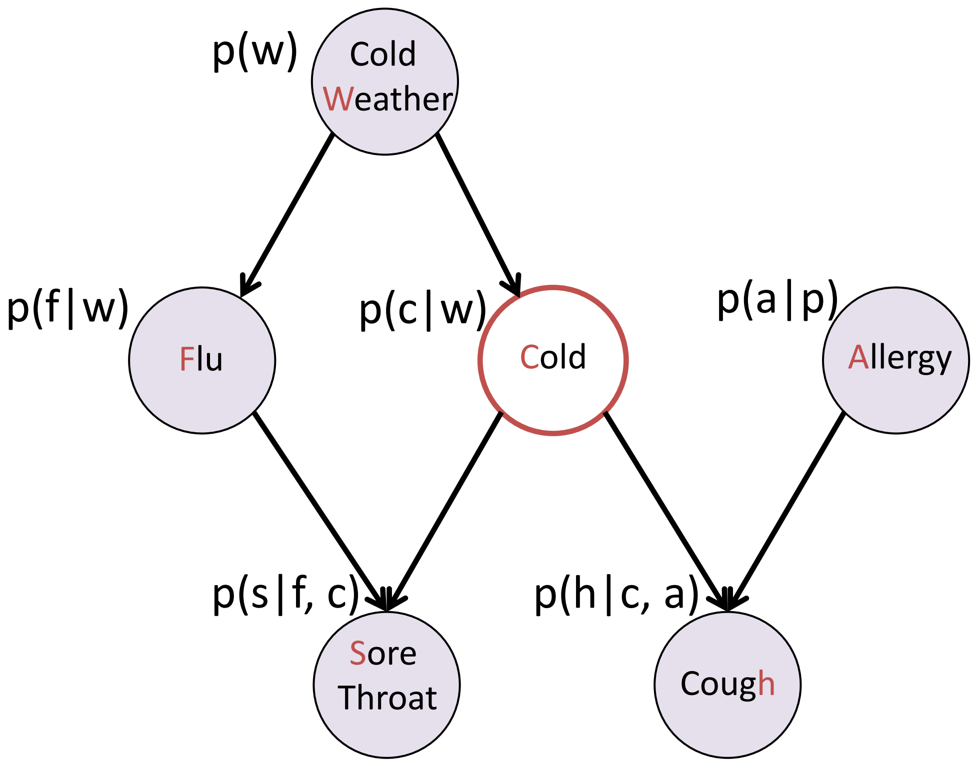


$$P(C{=}c \mid W{=}1, F{=}1, P{=}1, \ldots, I{=}0)$$

$$= P(C{=}c \mid W{=}1, F{=}1, S{=}0, H{=}1, A{=}1)$$

Markov Blanket!

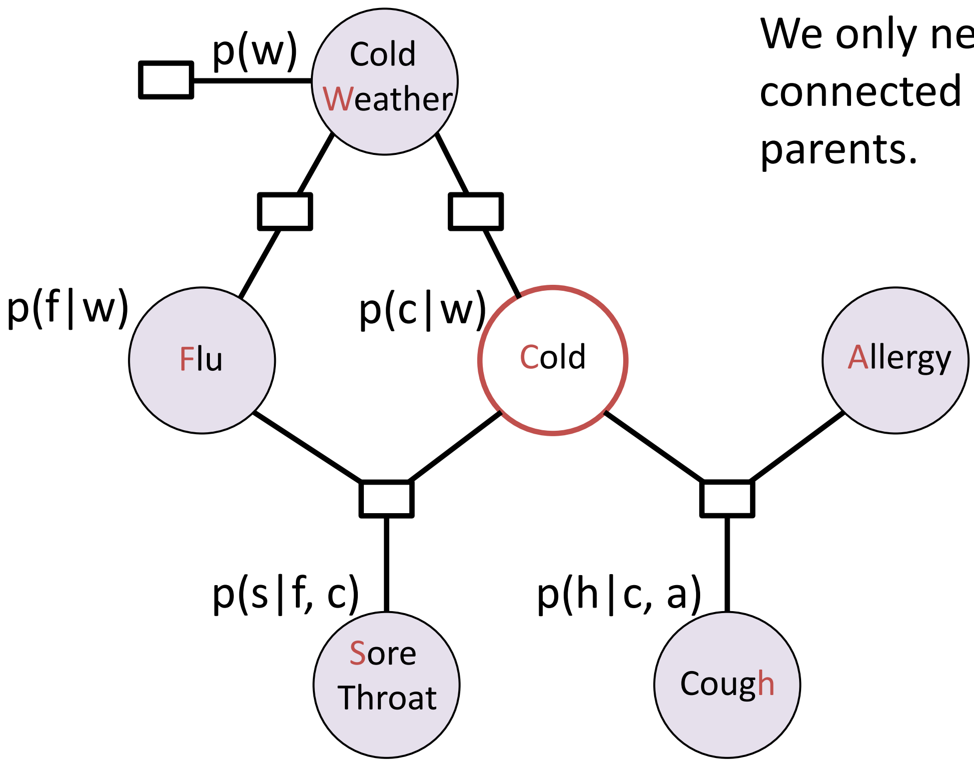
Gibbs Sampling



$$P(C{=}c \mid W{=}1, F{=}1, P{=}1, \ldots, I{=}0$$

$$= P(C{=}c \mid W{=}1, F{=}1, S{=}0, H{=}1, A{=}1)$$

Gibbs Sampling
From a Factor Graph Perspective



We only need look at the factors connected to the children and the parents.

$$P(C{=}c \mid W{=}1, F{=}1, P{=}1, \ldots, I{=}0$$

$$= P(C{=}c \mid W{=}1, F{=}1, S{=}0, H{=}1, A{=}1)$$

$$= p(w)\, p(f \mid w)\, p(c \mid w)\, p(s \mid f, c)\, p(h \mid c, a)$$

Questions?