# Inference in Bayesian Networks

CS 5804 Introduction to Artificial Intelligence
Virginia Tech

# Plan

- Briefly review probability
- Brute-force inference
- Smarter inference: variable elimination

# Probability Identities

# Probability Identities

- Random variables in caps (**A**)
  - values in lowercase: **A = a** or just **a** for shorthand

# Probability Identities

- Random variables in caps (**A**)
    - values in lowercase: **A = a** or just **a** for shorthand
- $P(a \mid b) = P(a, b) / P(b)$

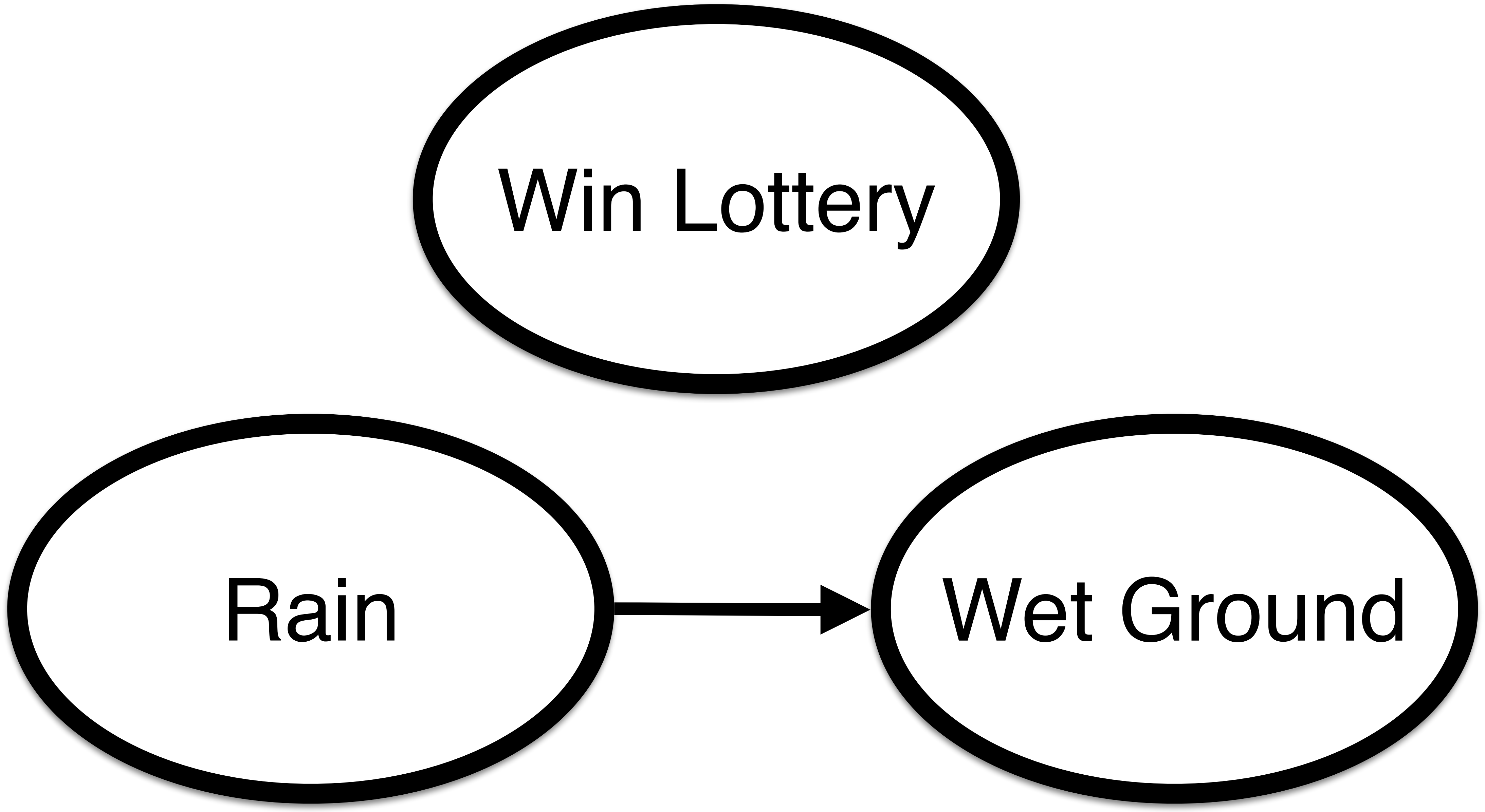
# Probability Identities

- Random variables in caps (**A**)
  - values in lowercase: **A = a** or just **a** for shorthand
- $P(a \mid b) = P(a, b) / P(b)$
- $P(a, b) = P(a \mid b) P(b)$

# Probability Identities

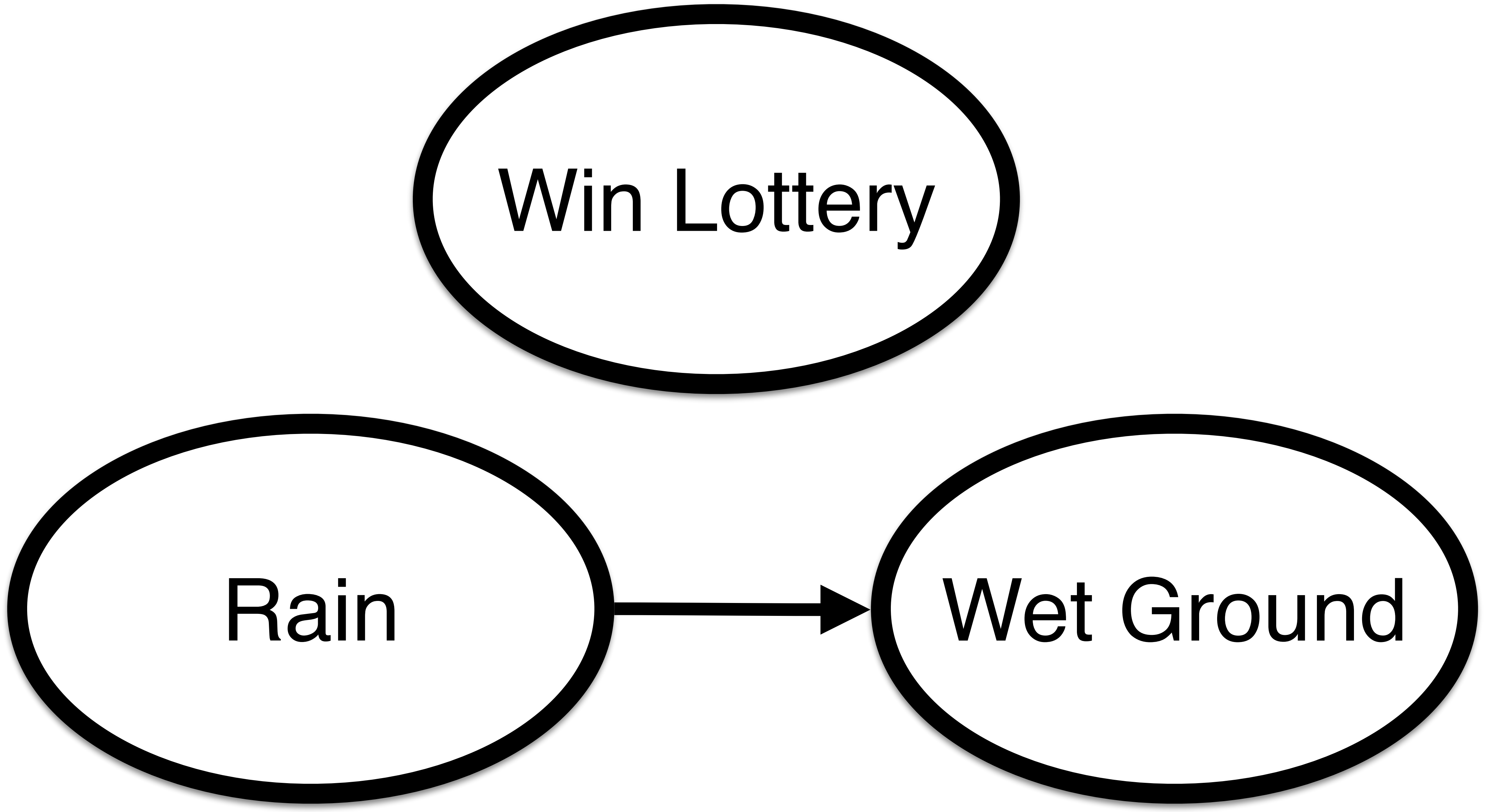
- Random variables in caps (**A**)
  - values in lowercase: **A = a** or just **a** for shorthand
- $P(a \mid b) = P(a, b) / P(b)$
- $P(a, b) = P(a \mid b) P(b)$
- $P(b \mid a) = P(a \mid b) P(b) / P(a)$

# Bayesian Networks



# Bayesian Networks

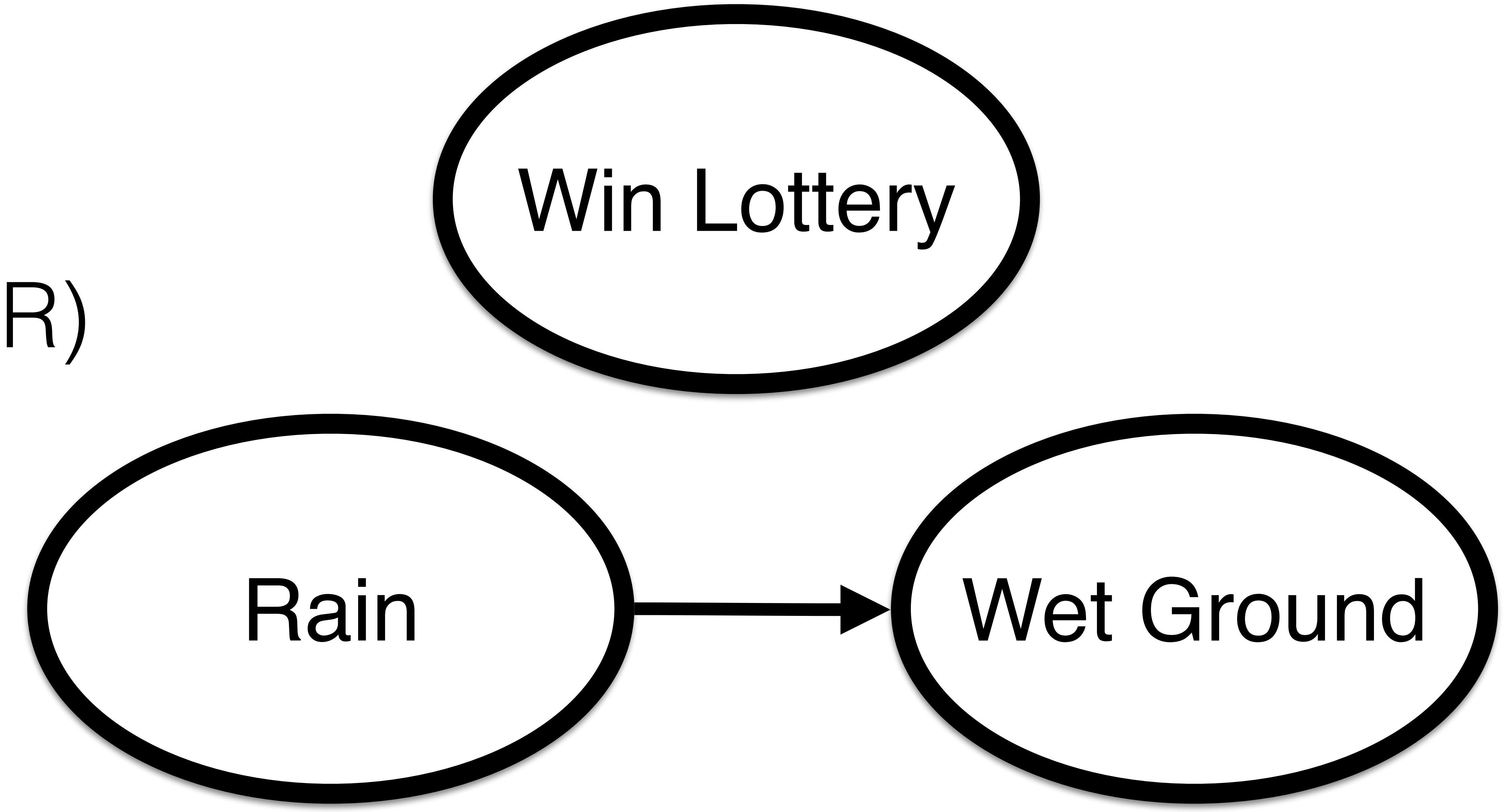
P(L, R, W)



# Bayesian Networks

$P(L, R, W)$

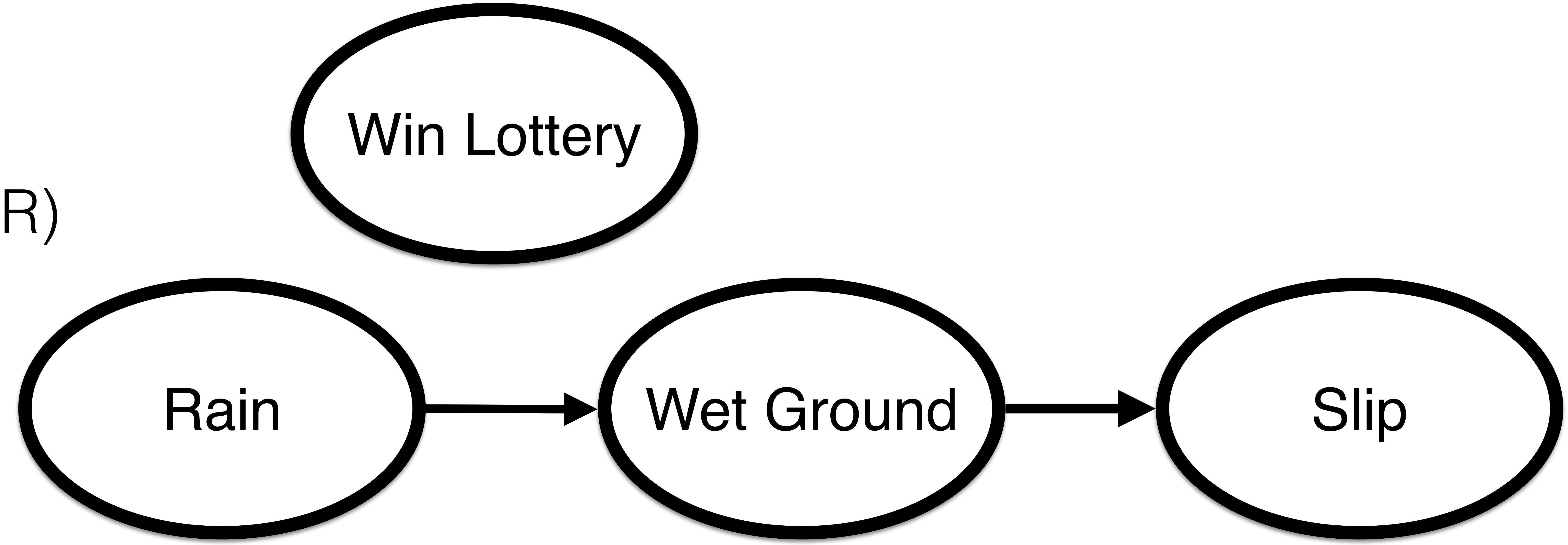
$= P(L)\, P(R)\, P(W \mid R)$



# Bayesian Networks

P(L, R, W)

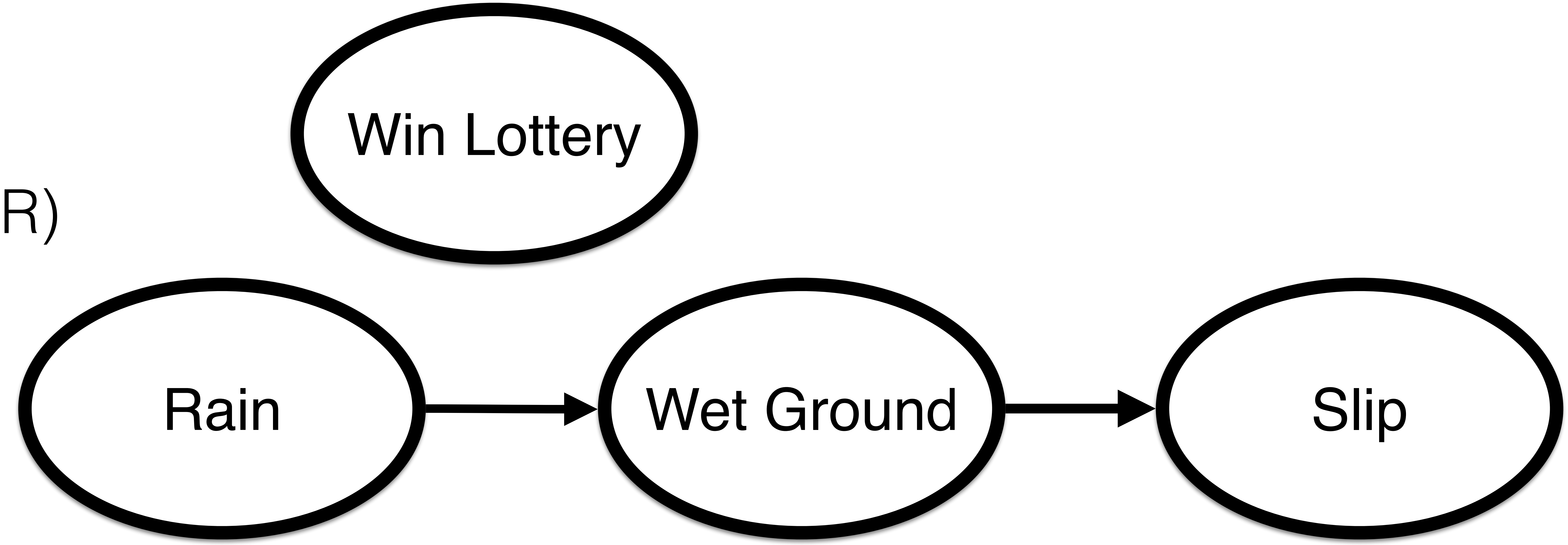
= P(L) P(R) P(W | R)



# Bayesian Networks

P(L, R, W)

= P(L) P(R) P(W | R)

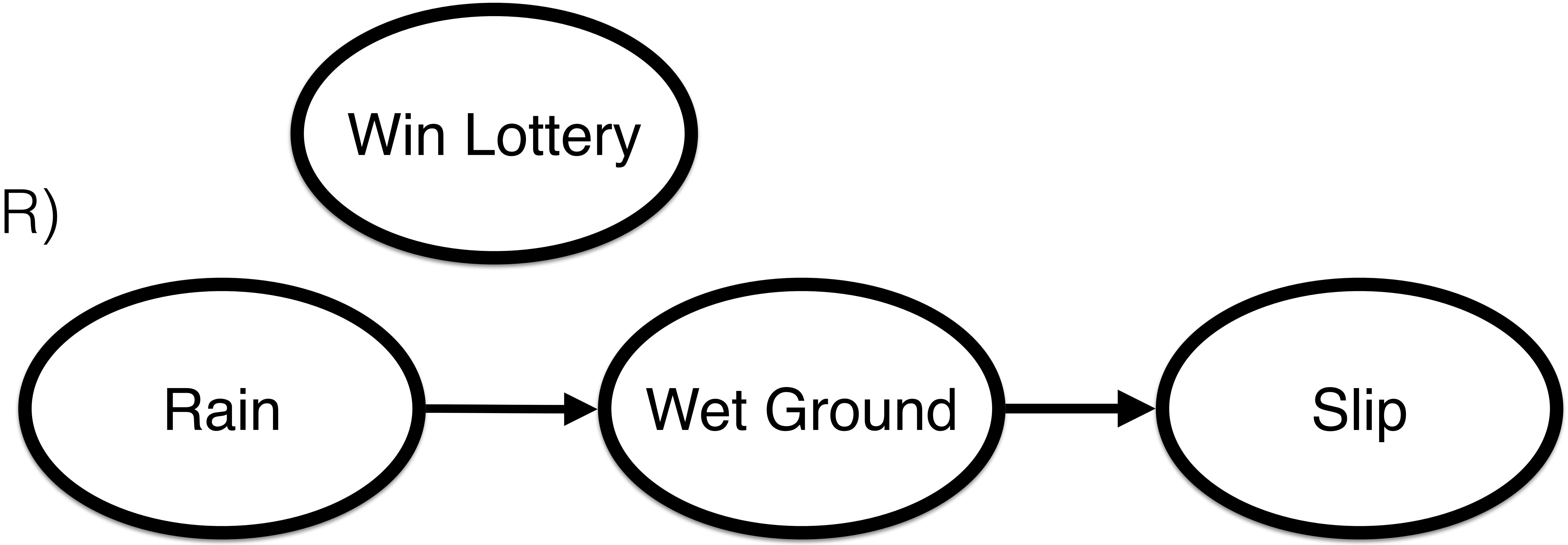



P(L, R, W, S) = P(L) P(R) P(W | R) P(S | W)

# Bayesian Networks

P(L, R, W)

= P(L) P(R) P(W | R)



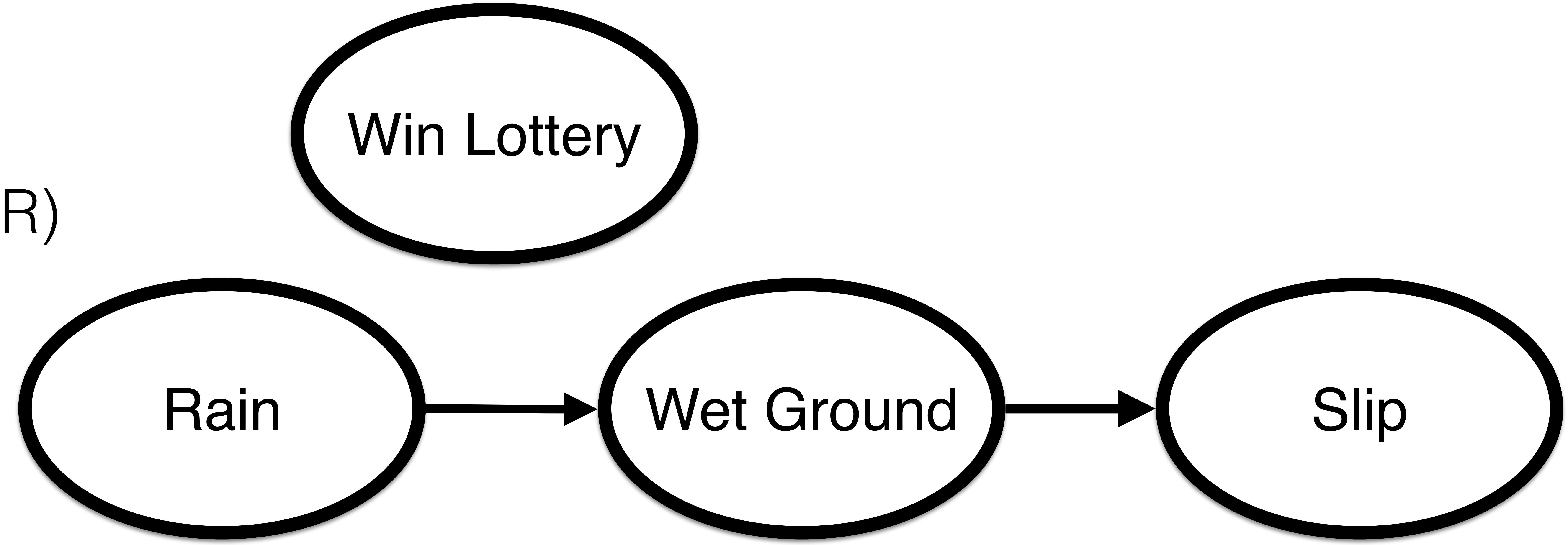

P(L, R, W, S) = P(L) P(R) P(W | R) P(S | W)

P(S | W, R)

# Bayesian Networks

P(L, R, W)

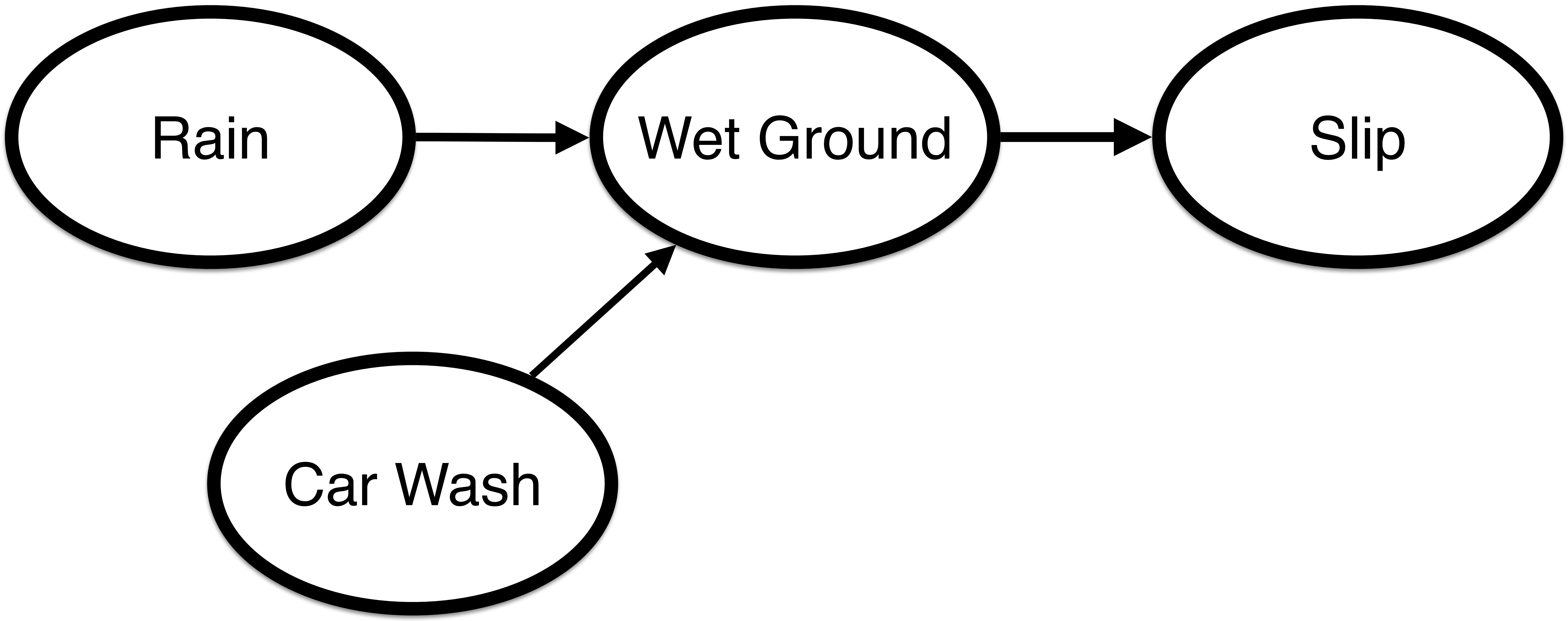
= P(L) P(R) P(W | R)




P(L, R, W, S) = P(L) P(R) P(W | R) P(S | W)

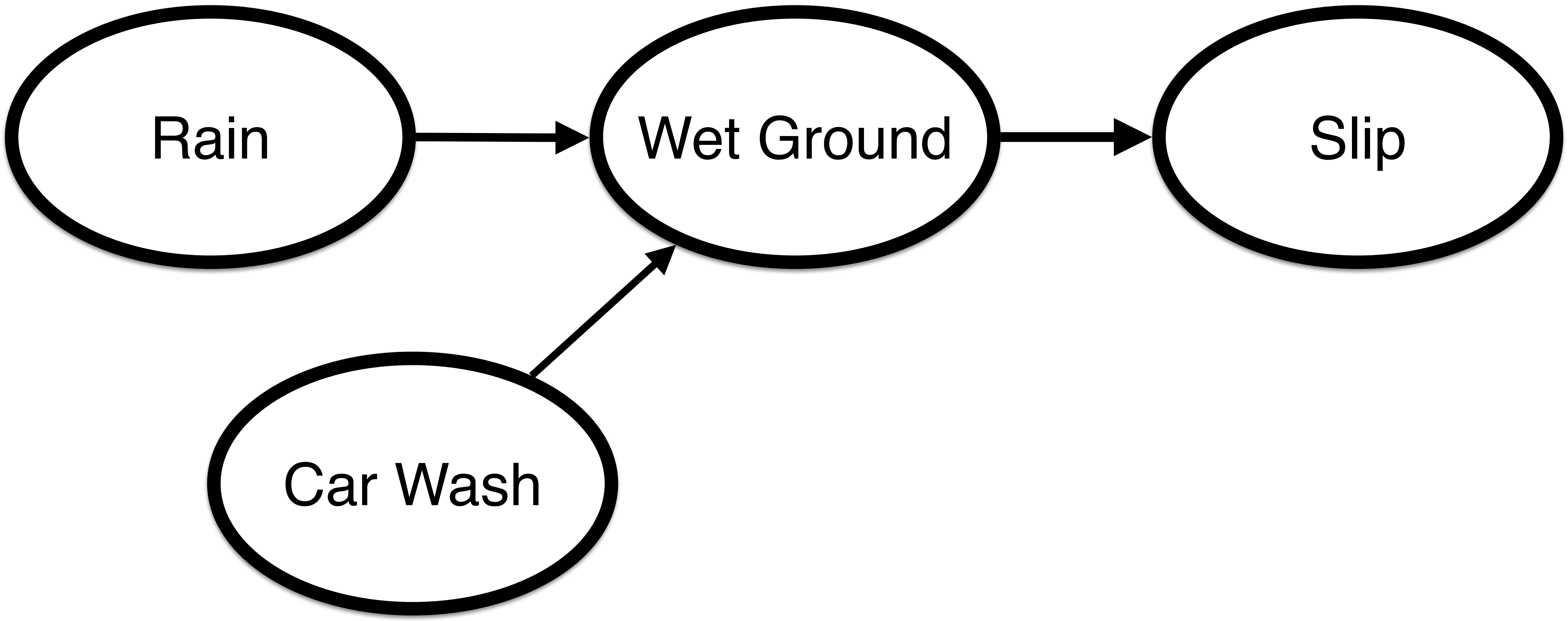
~~P(S | W, R)~~

# Bayesian Networks



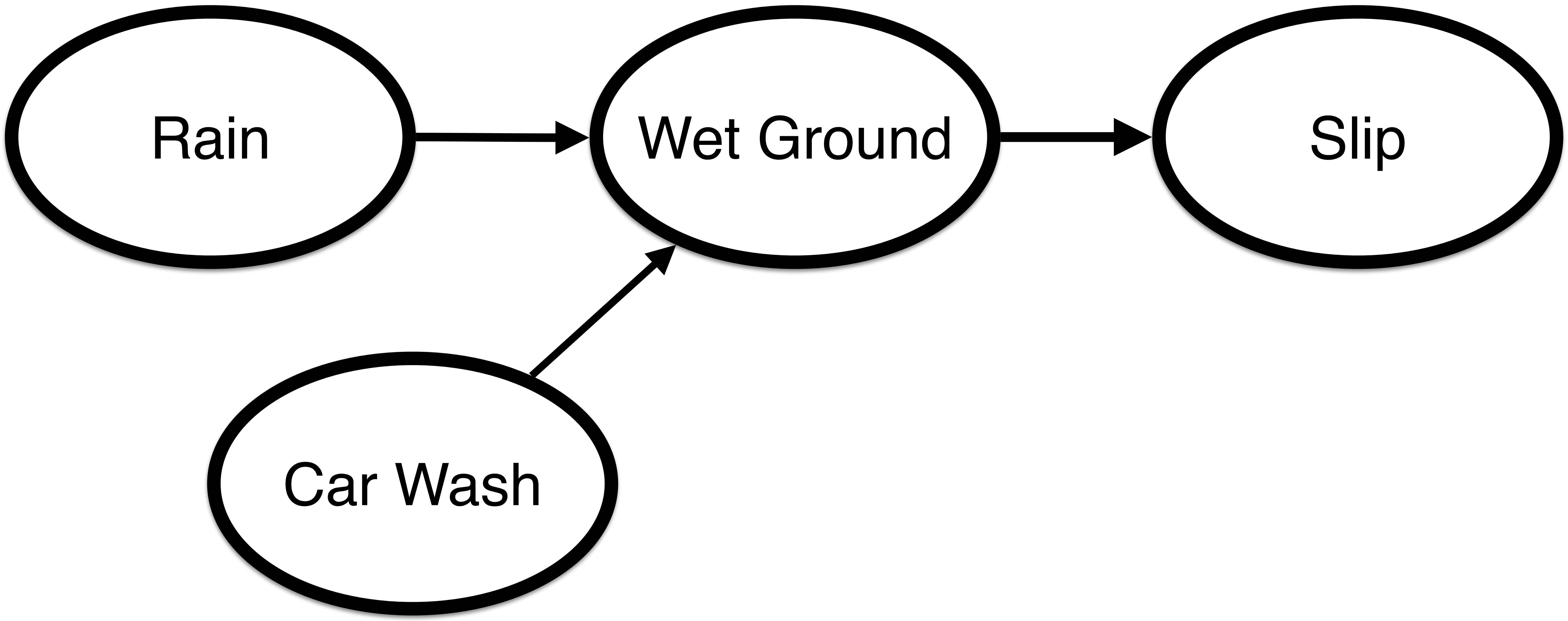
# Bayesian Networks

$P(R, W, S, C) = P(R)\, P(C)\, P(W \mid C, R)\, P(S \mid W)$



# Bayesian Networks

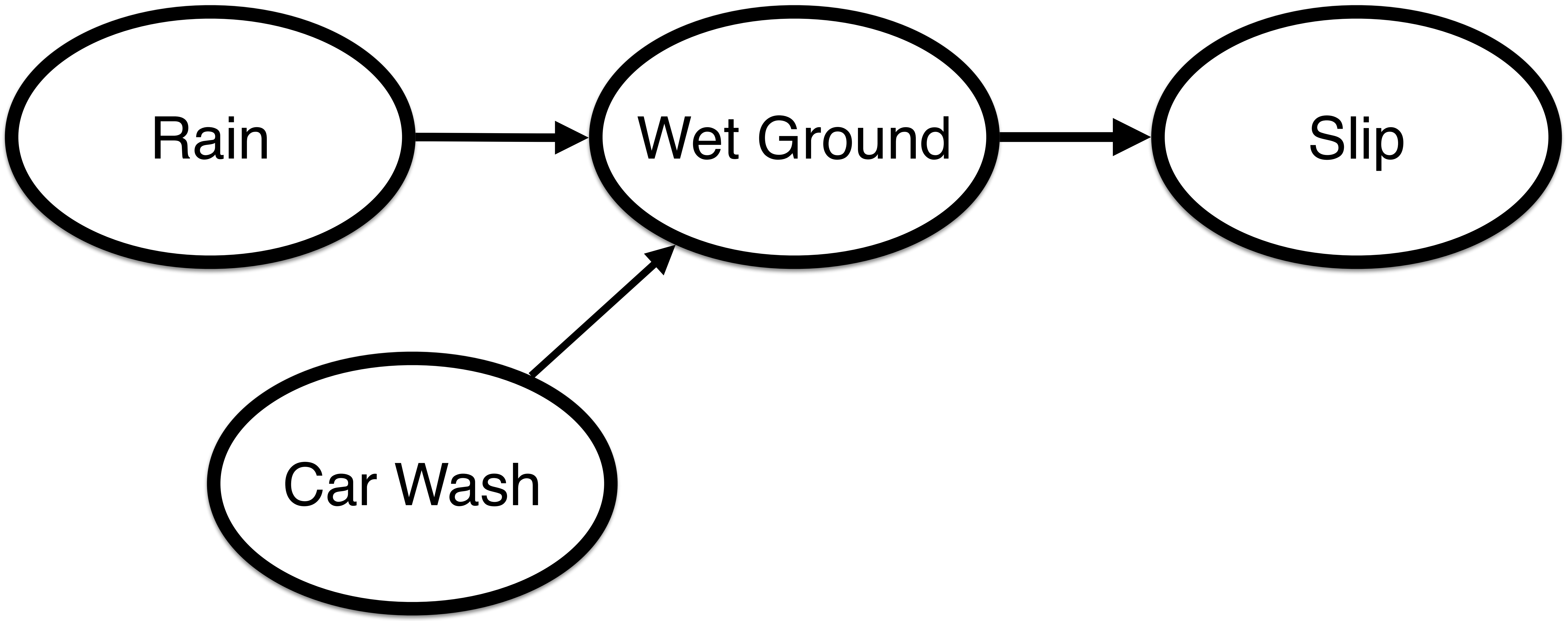
$P(R, W, S, C) = P(R)\,P(C)\,\boxed{P(W \mid C, R)}\,P(S \mid W)$



# Bayesian Networks

$P(R, W, S, C) = P(R)\,P(C)\,\boxed{P(W \mid C, R)}\,P(S \mid W)$

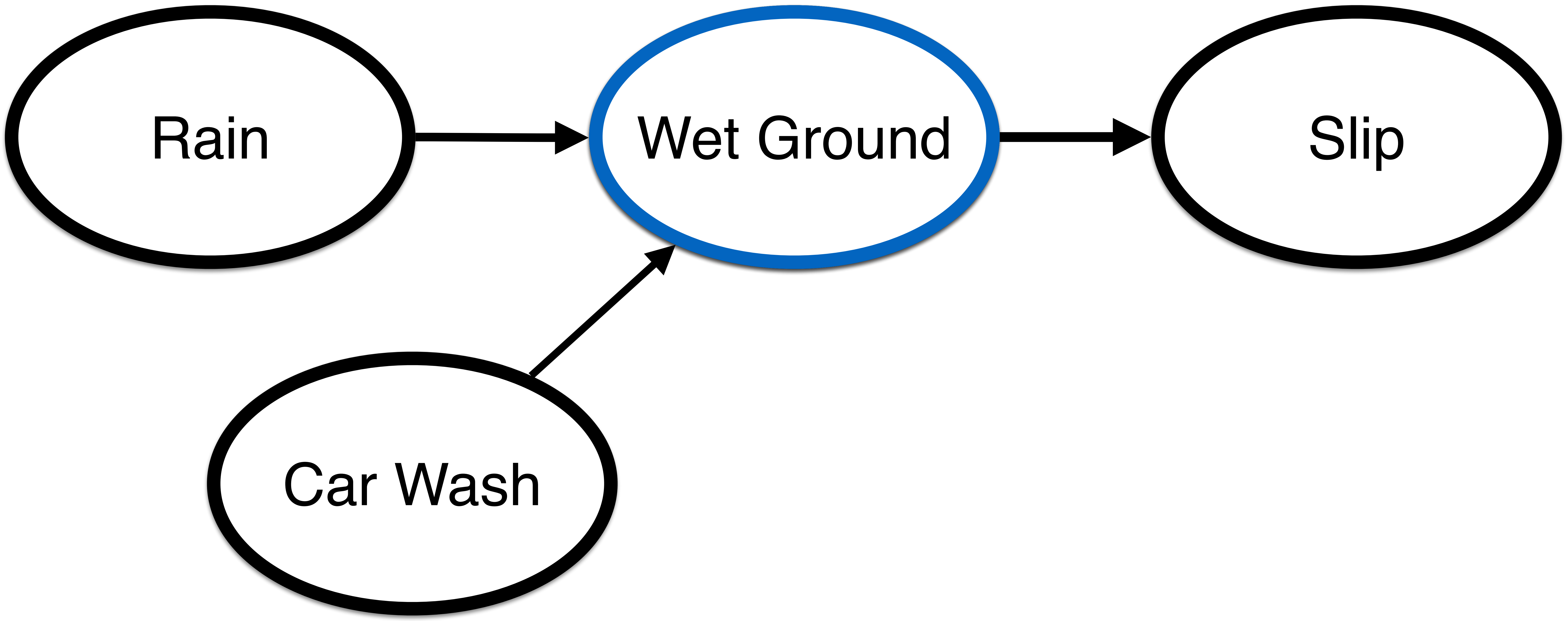
$P(X \mid \text{Parents}(X))$



# Bayesian Networks

$P(R, W, S, C) = P(R)\,P(C)\,\boxed{P(W \mid C, R)}\,P(S \mid W)$

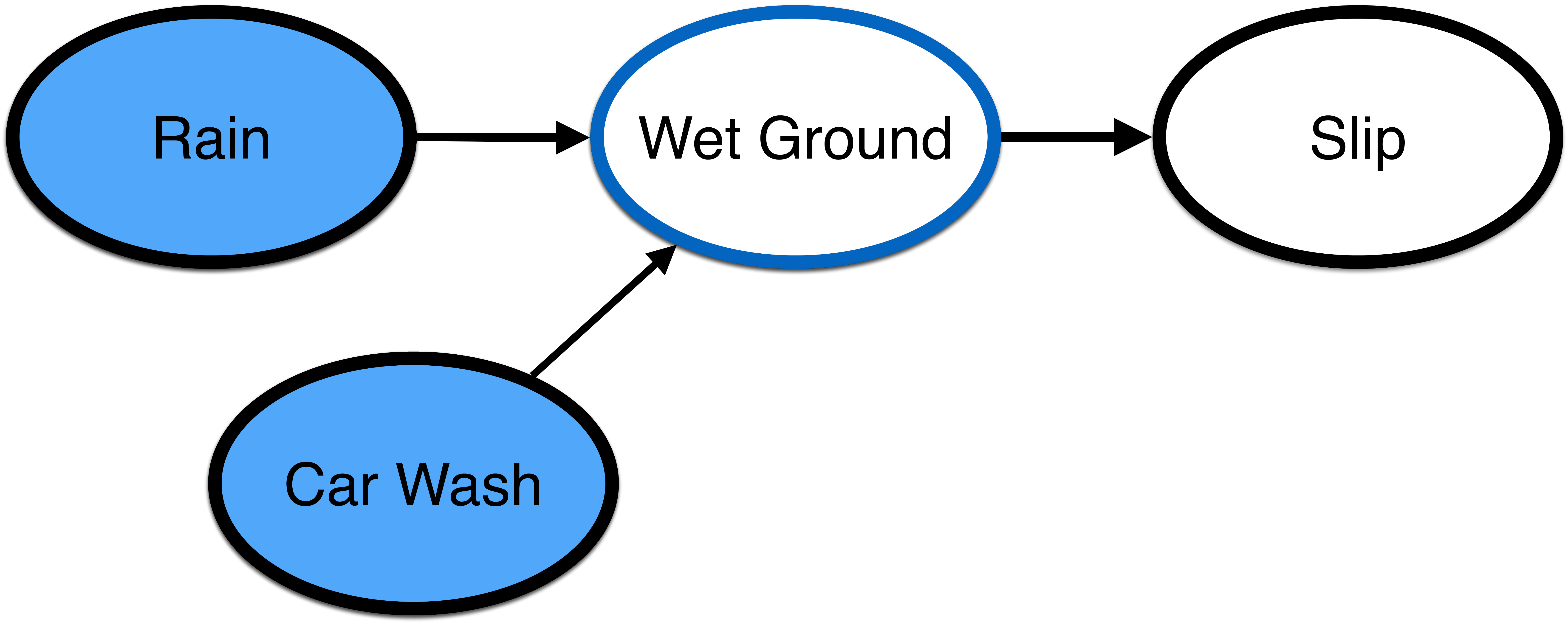
$P(X \mid \text{Parents}(X))$



# Bayesian Networks

$P(R, W, S, C) = P(R)\,P(C)\,\boxed{P(W \mid C, R)}\,P(S \mid W)$

$P(X \mid Parents(X))$



# Inference

- Given a Bayesian Network describing P(X, Y, Z), what is P(Y)
    - First approach: **enumeration**

$$P(R, W, S, C) = P(R)\, P(C)\, P(W \mid C, R)\, P(S \mid W)$$

P(R, W, S, C) = P(R) P(C) P(W | C, R) P(S | W)

$$P(r|s) = \sum_{w} \sum_{c} P(r, w, s, c) / P(s)$$

P(R, W, S, C) = P(R) P(C) P(W | C, R) P(S | W)

$$P(r|s) = \sum_{w} \sum_{c} P(r, w, s, c) / P(s)$$

$$P(r|s) \propto \sum_{w} \sum_{c} P(r) P(c) P(w|c, r) P(s|w)$$

P(R, W, S, C) = P(R) P(C) P(W | C, R) P(S | W)

$$P(r|s) = \sum_{w} \sum_{c} P(r, w, s, c)/P(s)$$

$$P(r|s) \propto \sum_{w} \sum_{c} P(r)P(c)P(w|c, r)P(s|w)$$

$$P(r|s) \propto P(r) \sum_{w} P(s|w) \sum_{c} P(c)P(w|c, r)$$

P(R, W, S, C) = P(R) P(C) P(W | C, R) P(S | W)

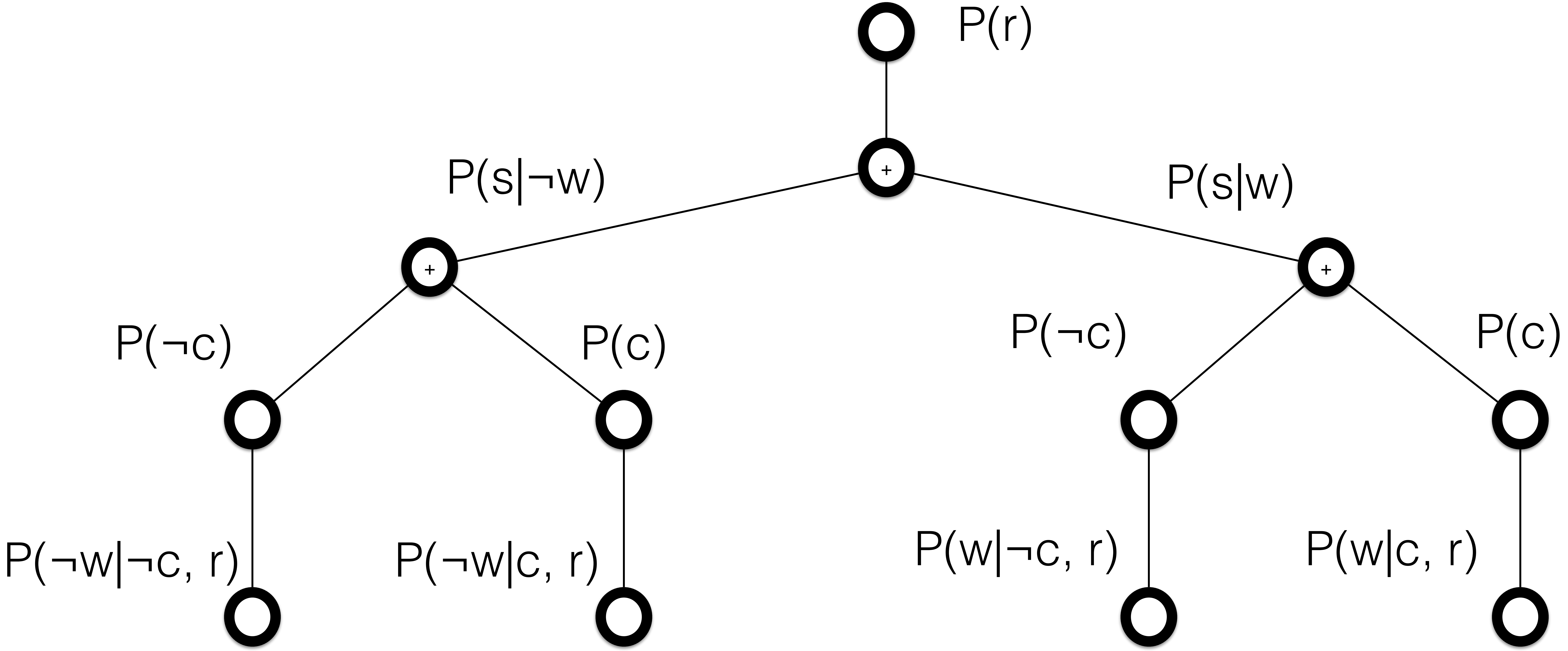
$$P(r|s) = \sum_{w} \sum_{c} P(r, w, s, c)/P(s)$$

$$P(r|s) \propto \sum_{w} \sum_{c} P(r)P(c)P(w|c, r)P(s|w)$$

$$P(r|s) \propto P(r) \sum_{w} P(s|w) \sum_{c} P(c)P(w|c, r)$$

$O(2^n)$

$$P(r|s) \propto P(r) \sum_{w} P(s|w) \sum_{c} P(c) P(w|c, r)$$



# Second Approach: Variable Elimination

# Second Approach: Variable Elimination

$$P(r|s) \propto \sum_{w} \sum_{c} P(r)P(c)P(w|c,r)P(s|w)$$

# Second Approach: Variable Elimination

$$P(r|s) \propto \sum_{w} \sum_{c} P(r) \boxed{P(c) P(w|c,r)} P(s|w)$$

# Second Approach: Variable Elimination

$$P(r|s) \propto \sum_{w} \sum_{c} P(r) P(c) P(w|c, r) P(s|w)$$

$$f_C(w) = \sum_{c} P(c) P(w|c, r)$$

# Second Approach: Variable Elimination

$$P(r|s) \propto \sum_w \sum_c P(r) \boxed{P(c)P(w|c,r)} P(s|w)$$

$$f_C(w) = \sum_c P(c)P(w|c,r)$$

$$P(r|s) \propto \sum_w P(r)P(s|w)f_c(w)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

$$P(Y) = \sum_{w} \sum_{x} \sum_{z} P(w)P(x|w)P(Y|x)P(z|Y)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

$$P(Y) = \sum_w \sum_x \sum_z P(w)P(x|w)P(Y|x)P(z|Y)$$

$$f_w(x) = \sum_w P(w)P(x|w)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

$$P(Y) = \sum_w \sum_x \sum_z P(w)P(x|w)P(Y|x)P(z|Y)$$

$$f_w(x) = \sum_w P(w)P(x|w)$$

$$P(Y) = \sum_x \sum_z f_w(x)P(Y|x)P(z|Y)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

$$P(Y) = \sum_w \sum_x \sum_z P(w)P(x|w)P(Y|x)P(z|Y)$$

$$f_w(x) = \sum_w P(w)P(x|w)$$

$$P(Y) = \sum_x \sum_z f_w(x)P(Y|x)P(z|Y)$$

$$f_x(Y) = \sum_x f_w(x)P(Y|x)$$

$$P(W, X, Y, Z) = P(W)P(X|W)P(Y|X)P(Z|Y)$$

$P(Y)$?

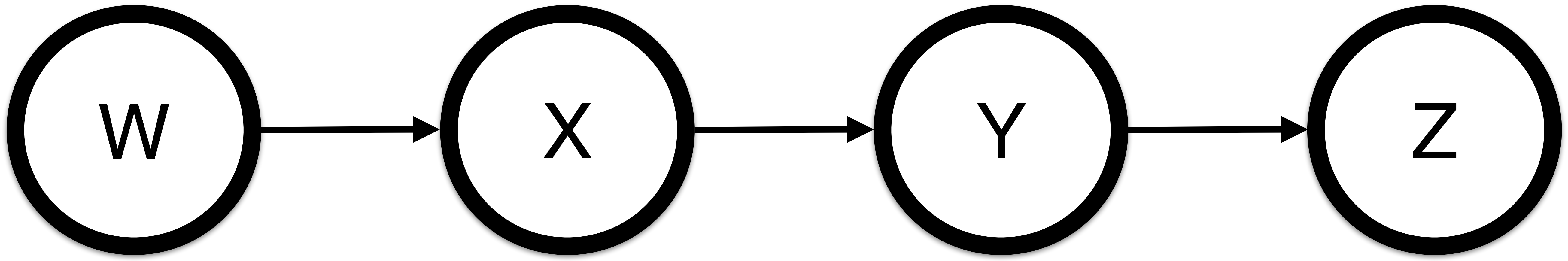
$$P(Y) = \sum_w \sum_x \sum_z P(w)P(x|w)P(Y|x)P(z|Y)$$

$$f_w(x) = \sum_w P(w)P(x|w)$$

$$P(Y) = \sum_x \sum_z f_w(x)P(Y|x)P(z|Y)$$

$$f_x(Y) = \sum_x f_w(x)P(Y|x)$$

$$P(Y) = \sum_z f_x(Y)P(z|Y)$$




$$P(Y) = \sum_w \sum_x \sum_z P(w)P(x|w)P(Y|x)P(z|Y)$$

$$f_w(x) = \sum_w P(w)P(x|w)$$

$$P(Y) = \sum_x \sum_z f_w(x)P(Y|x)P(z|Y)$$

$$f_x(Y) = \sum_x f_w(x)P(Y|x)$$

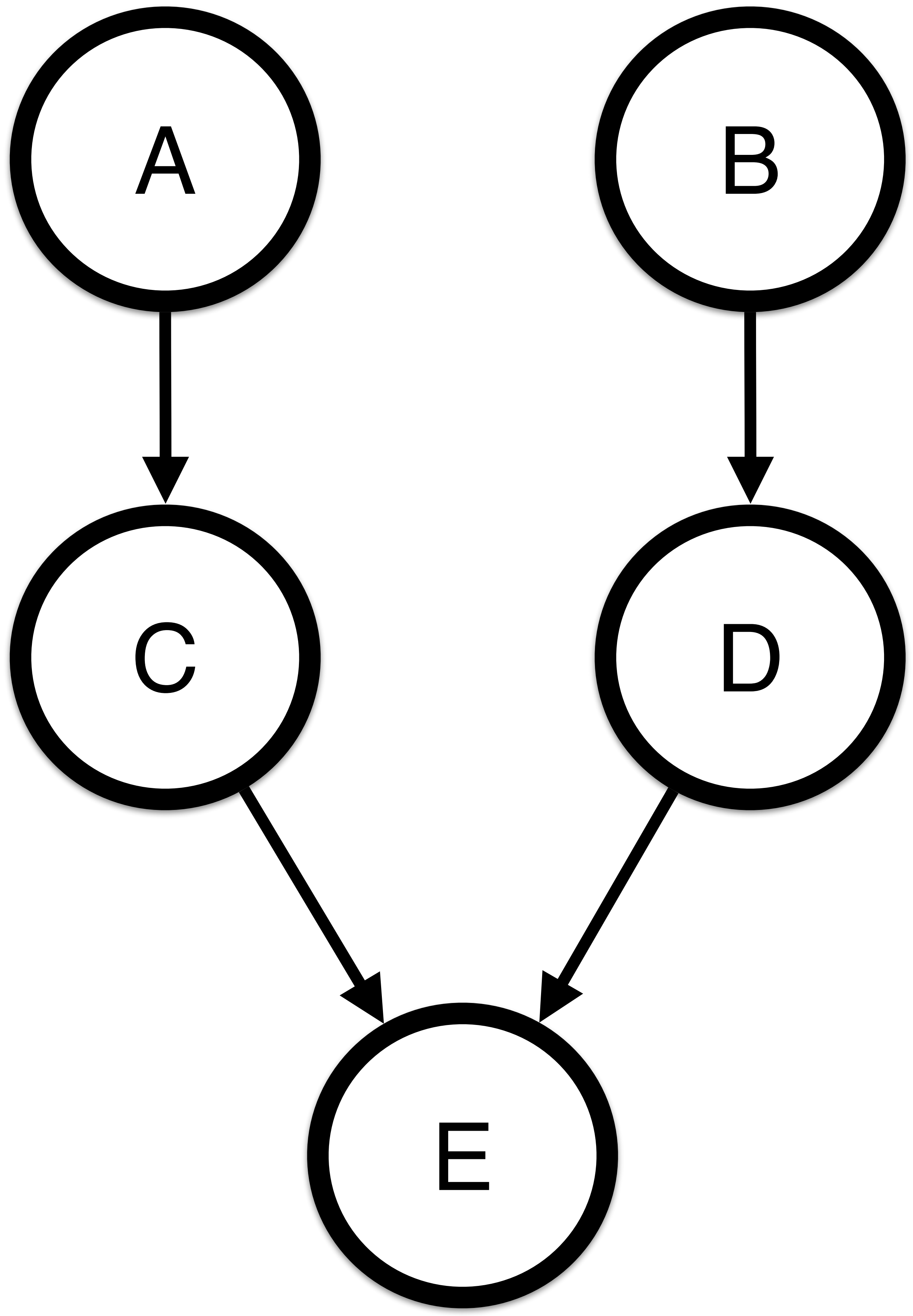
$$P(Y) = \sum_z f_x(Y)P(z|Y)$$

# Variable Elimination

- Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query
- Iterate:
  - choose variable to eliminate
  - sum terms relevant to variable, generate new factor
  - until no more variables to eliminate
- Exact inference is #P-Hard
  - in tree-structured BNs, linear time (in number of table entries)

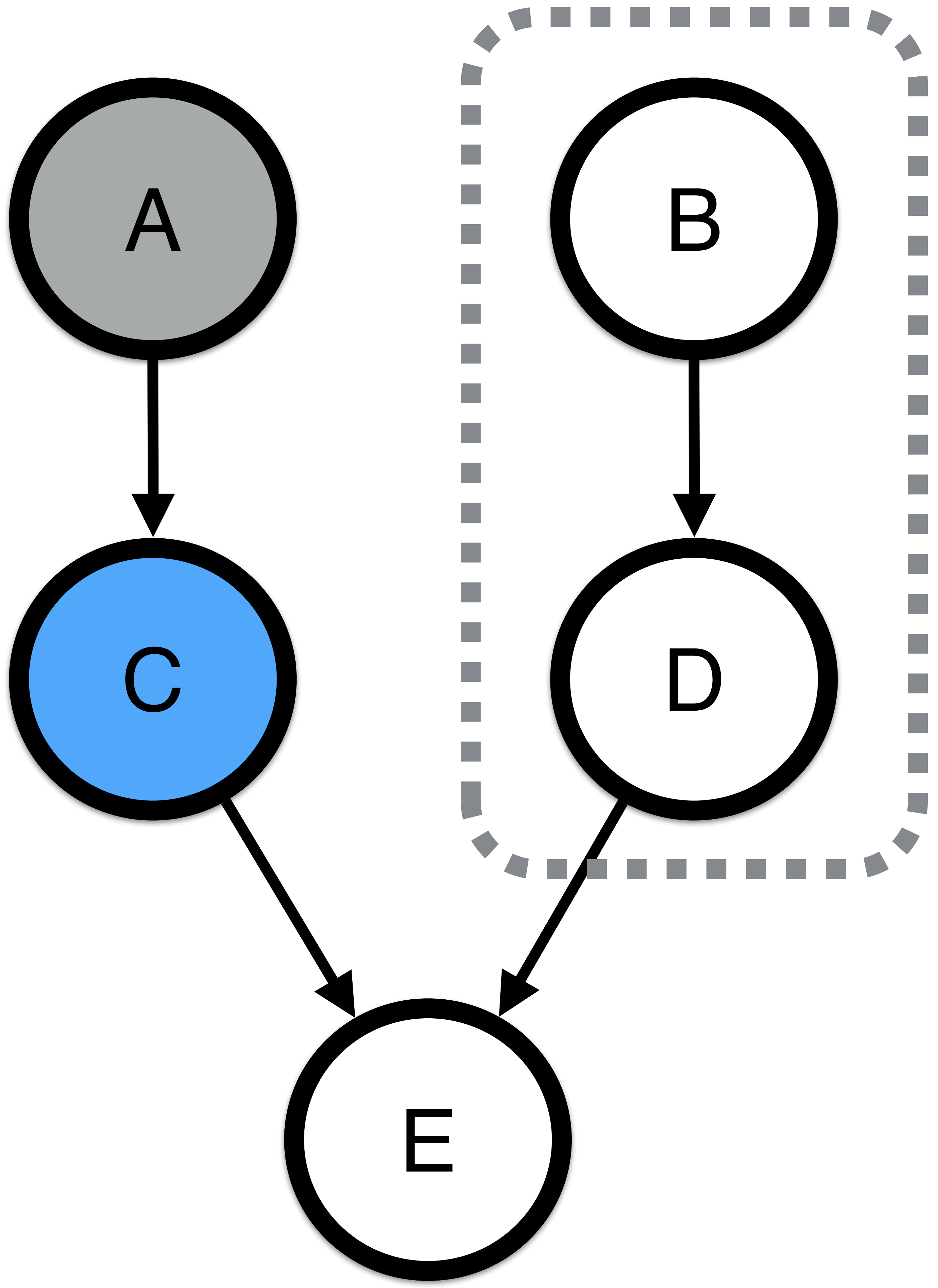
# Independence in Bayes Nets

- Each variable is conditionally independent of its **non-descendents** given its **parents**



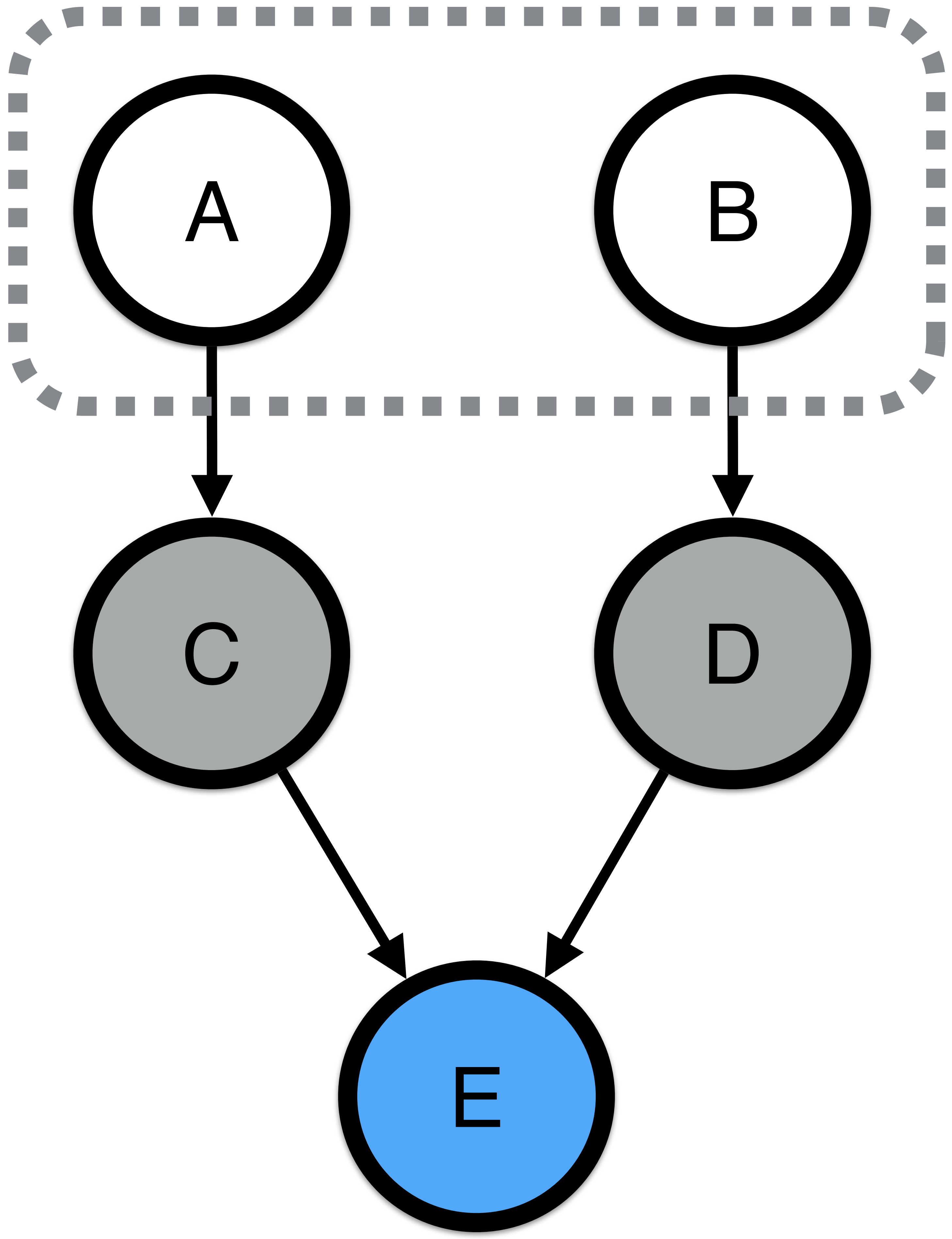
# Independence in Bayes Nets

- Each variable is conditionally independent of its **non-descendents** given its **parents**



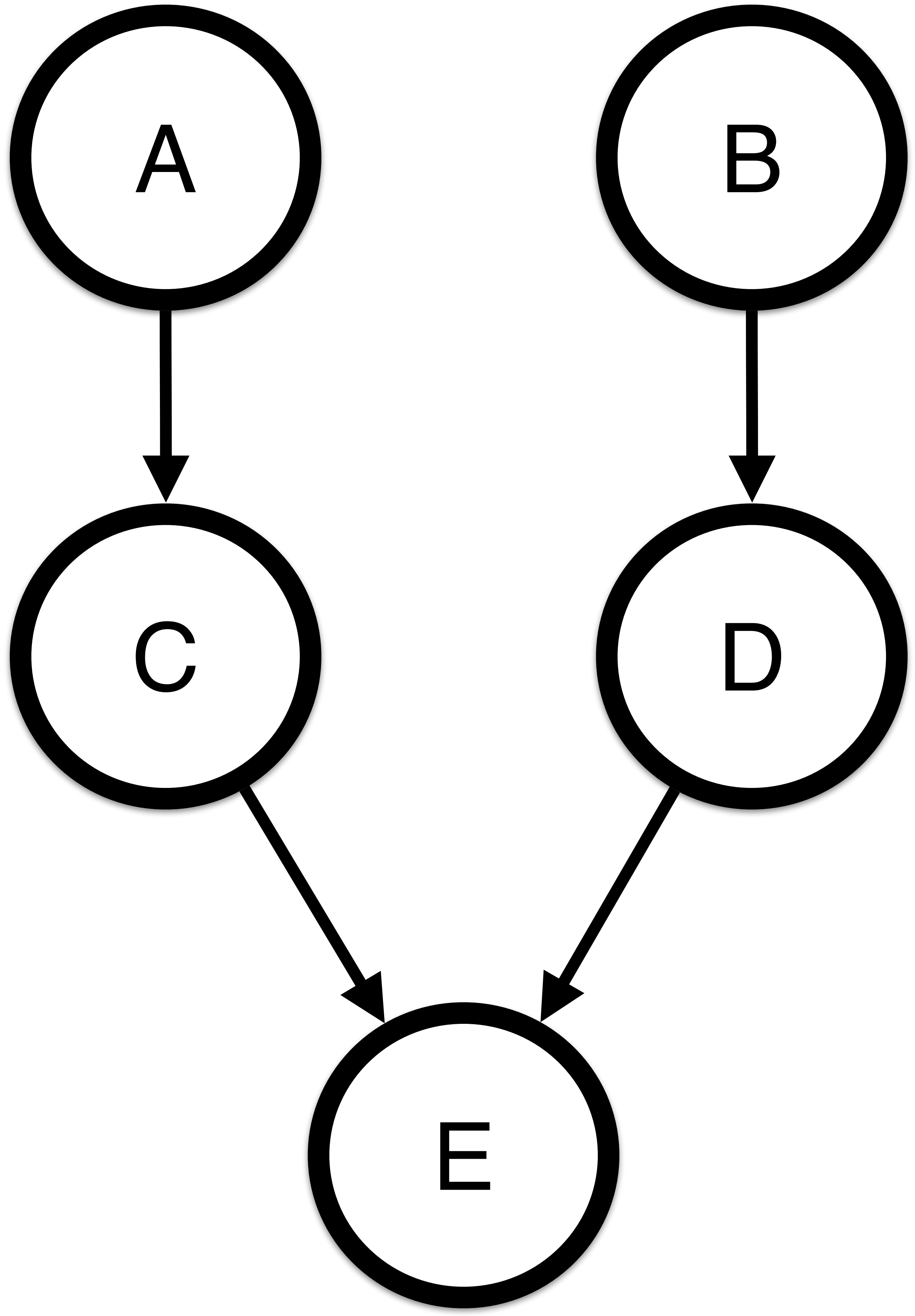
# Independence in Bayes Nets

- Each variable is conditionally independent of its **non-descendents** given its **parents**



# Independence in Bayes Nets

- Each variable is conditionally independent of its **non-descendents** given its **parents**
- Each variable is conditionally independent of any other variable given its **Markov blanket**
  - Parents, children, and children's parents

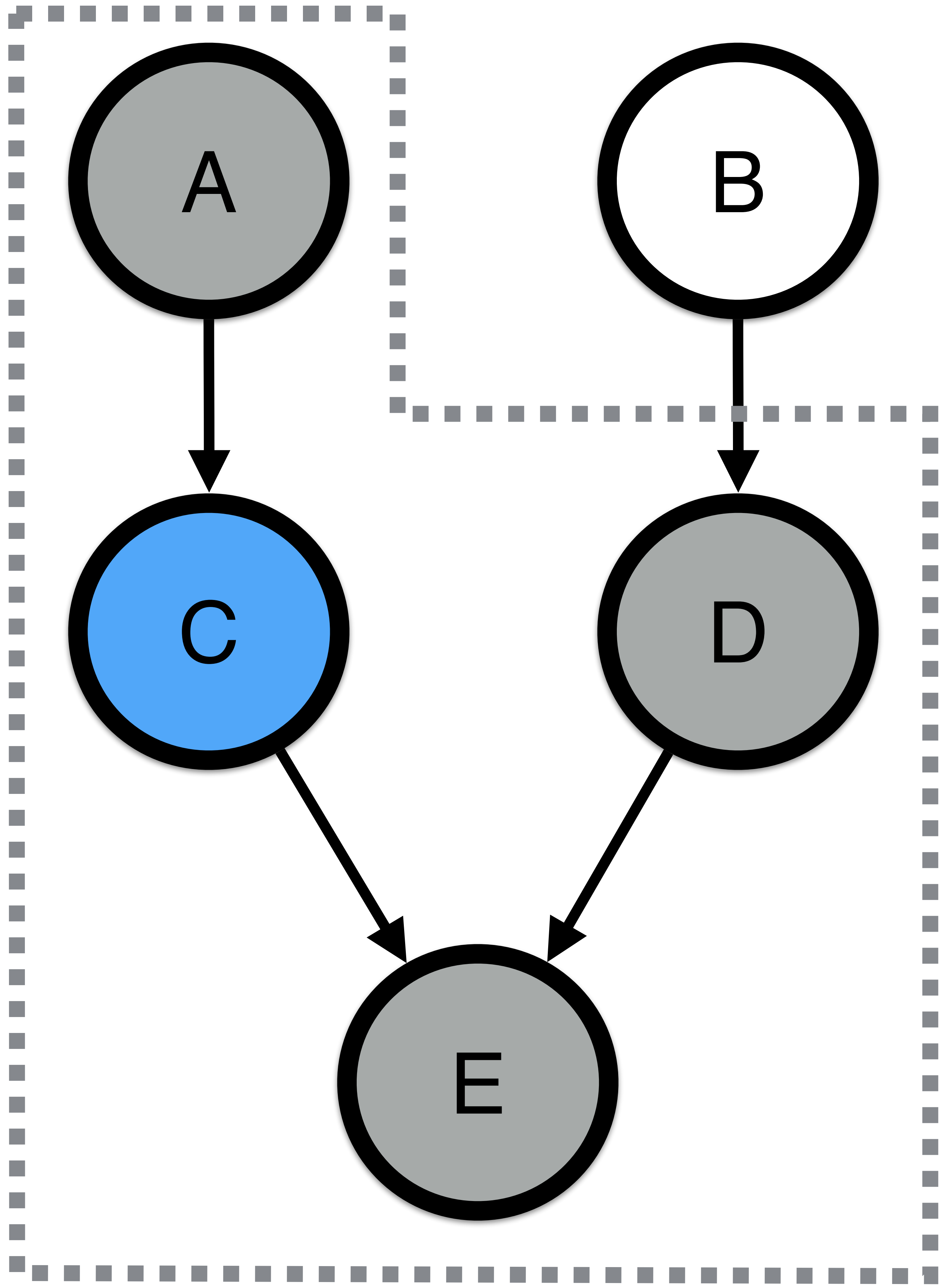


# Independence in Bayes Nets

- Each variable is conditionally independent of its **non-descendents** given its **parents**

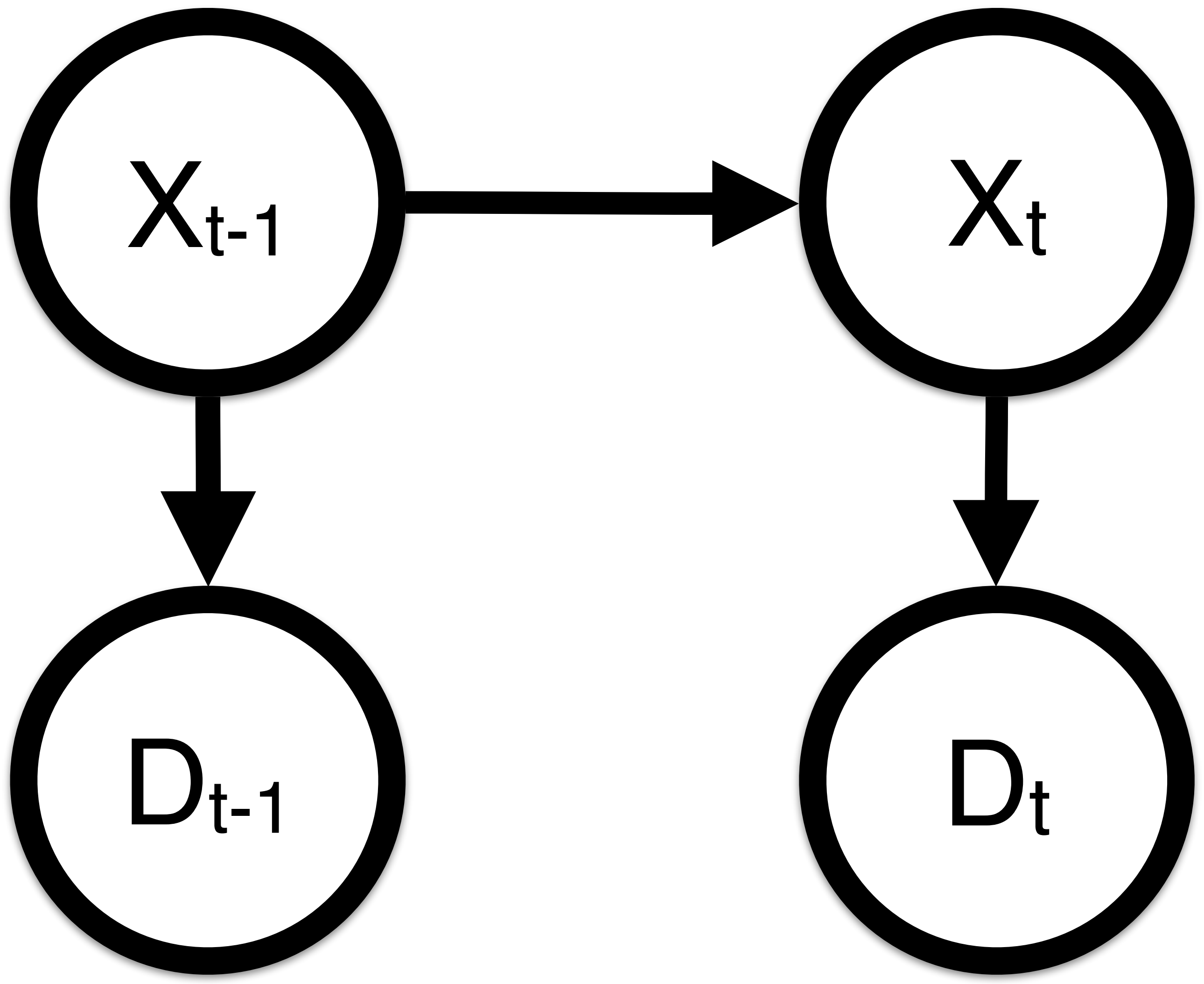
- Each variable is conditionally independent of any other variable given its **Markov blanket**

  - Parents, children, and children's parents



# Pacman Homework Warmup

- Given: P(ghostLocation | prevGhostLocation)
- Given: P(noisyDistance | ghostLocation)
- Goal: P(ghostLocation | noisyDistance)
  - Need: P(ghostLocation | all previous evidence)



# Reading

- Chapter 13
- Chapter 14 - 14.2, 14.4-14.4.3