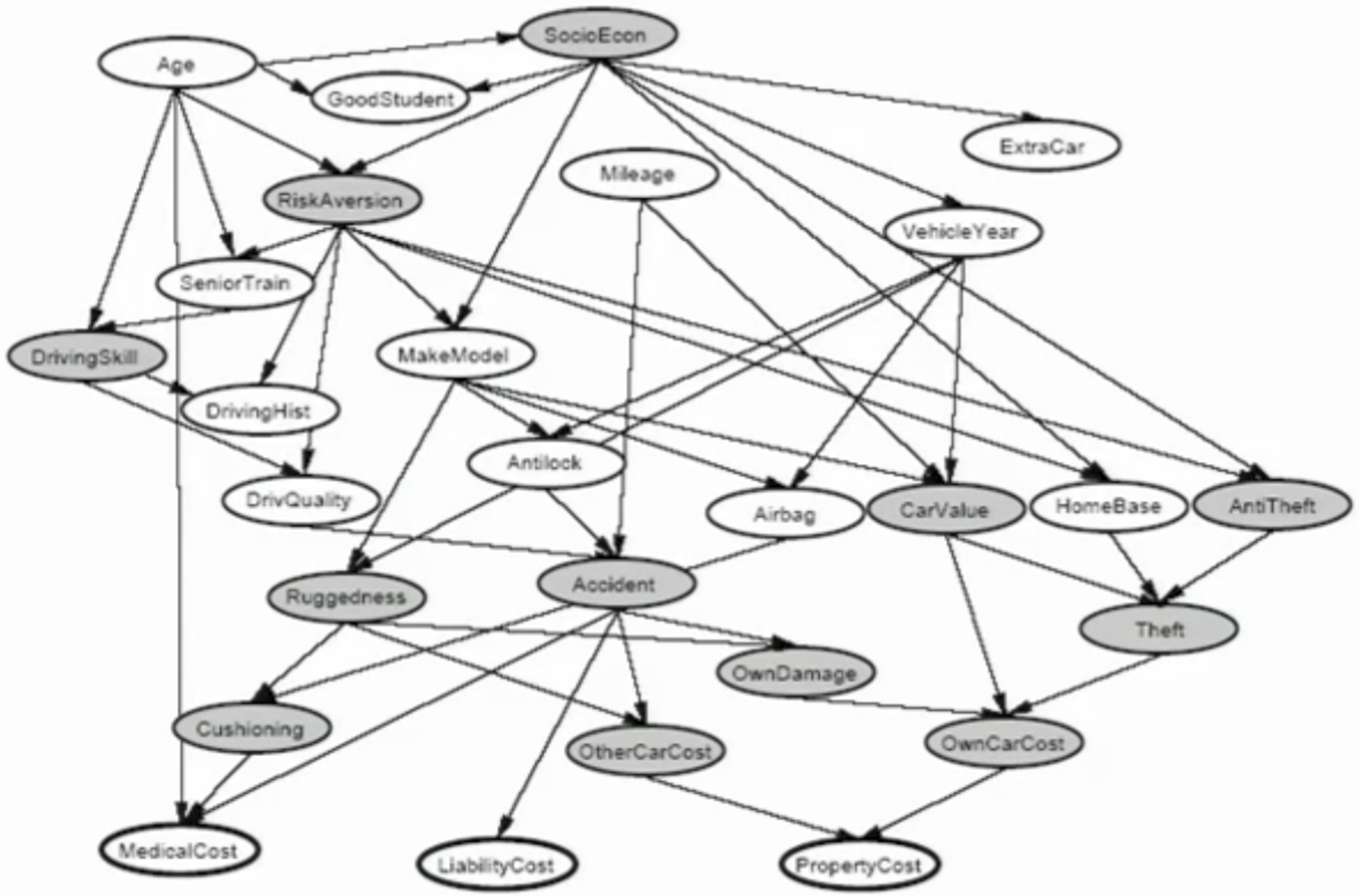
# Statistical Methods in AI (CSE 471)
# Bayesian networks

Vineet Gandhi
Centre for Visual Information Technology (CVIT)

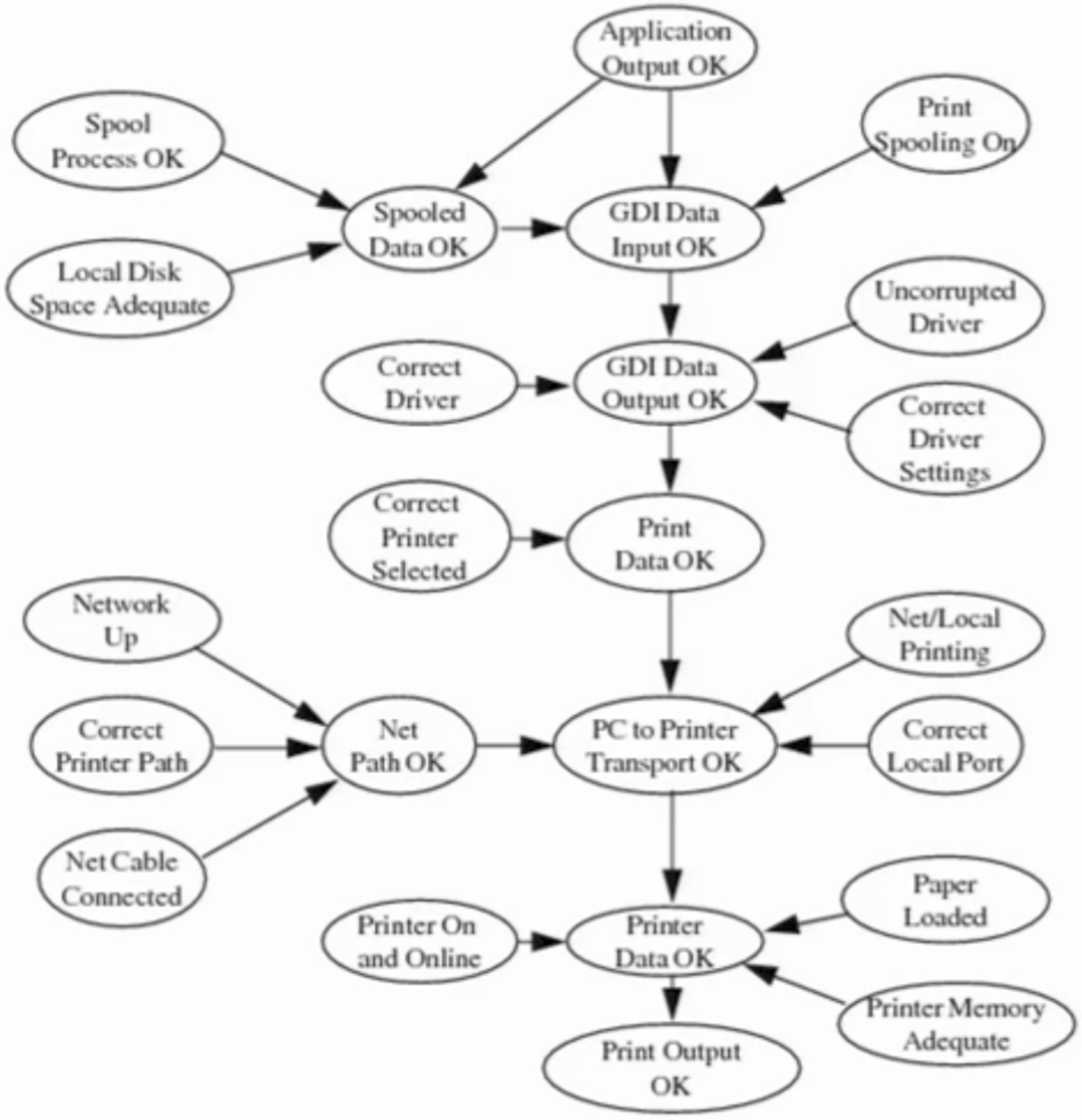
INTERNATIONAL INSTITUTE OF
INFORMATION TECHNOLOGY
HYDERABAD

# Motivation: Vehicle insurance

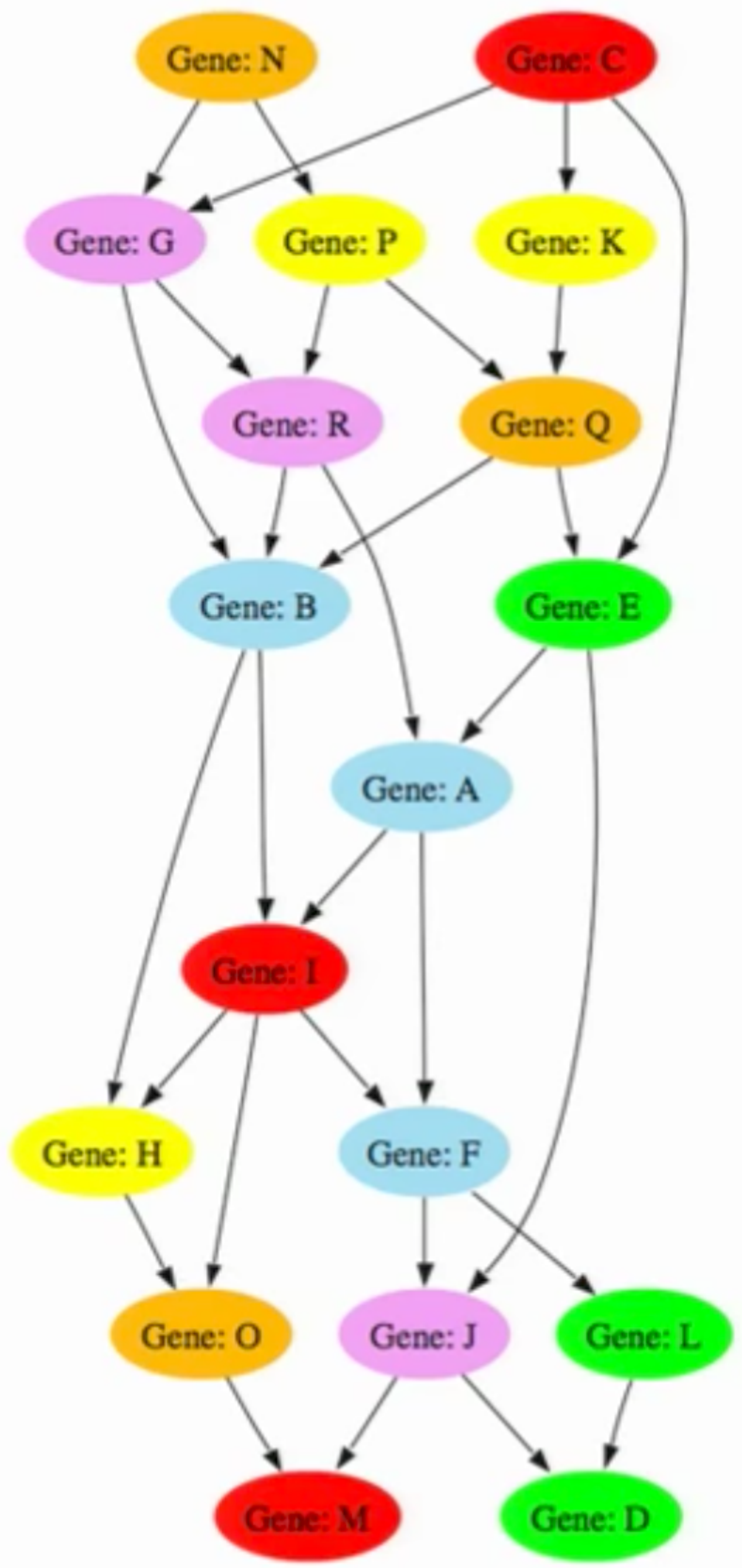



The 12 **shaded variables** are considered **hidden** or **unobservable**, while the other 15 are observable. The network has over 1400 parameters. An insurance company would be interested in predicting the bottom three "cost" variables, given all or a subset of the other observable variables.

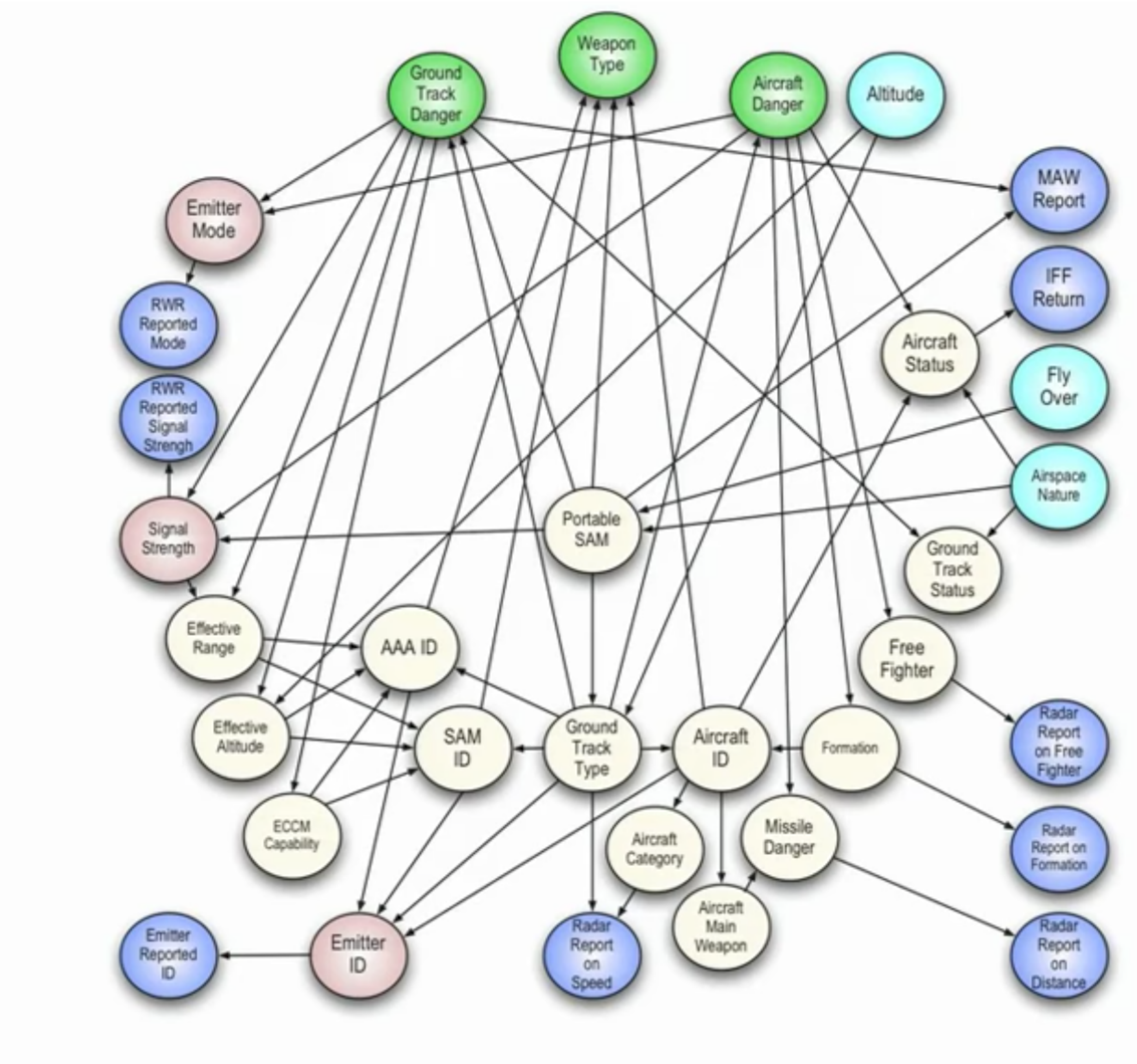
# Motivation: Windows printer troubleshooter



# Motivation: Gene expression



# Motivation: Radar and aircraft control



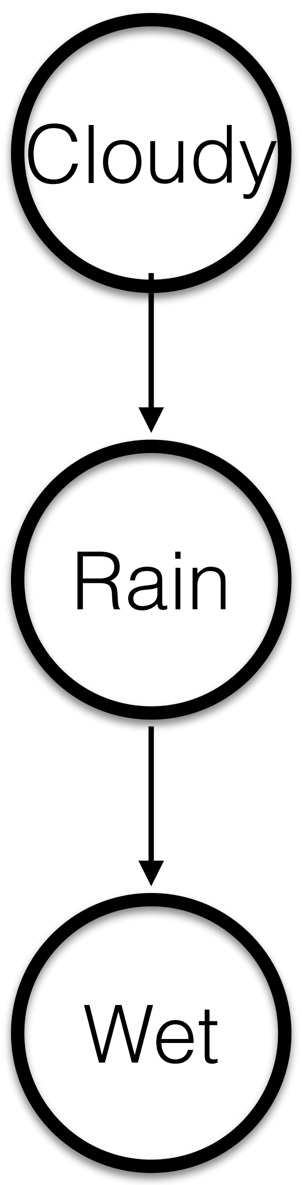
# Joint, Marginal, Conditional

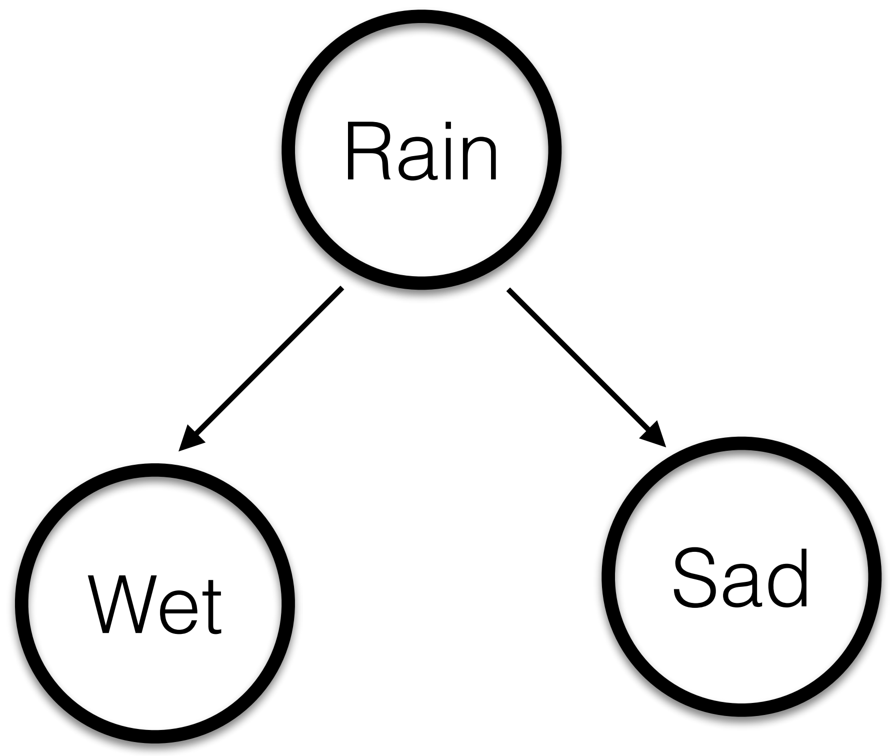
| | | Cheated on College Exam? | | |
|---|---|---|---|---|
| | | Yes | No | Total |
| Gender | Male | .32 | .22 | .54 |
| | Female | .28 | .18 | .46 |
| | Total | .60 | .40 | 1.00 |

Given the person has cheated, what is the probability he is male?

# Conditional independence

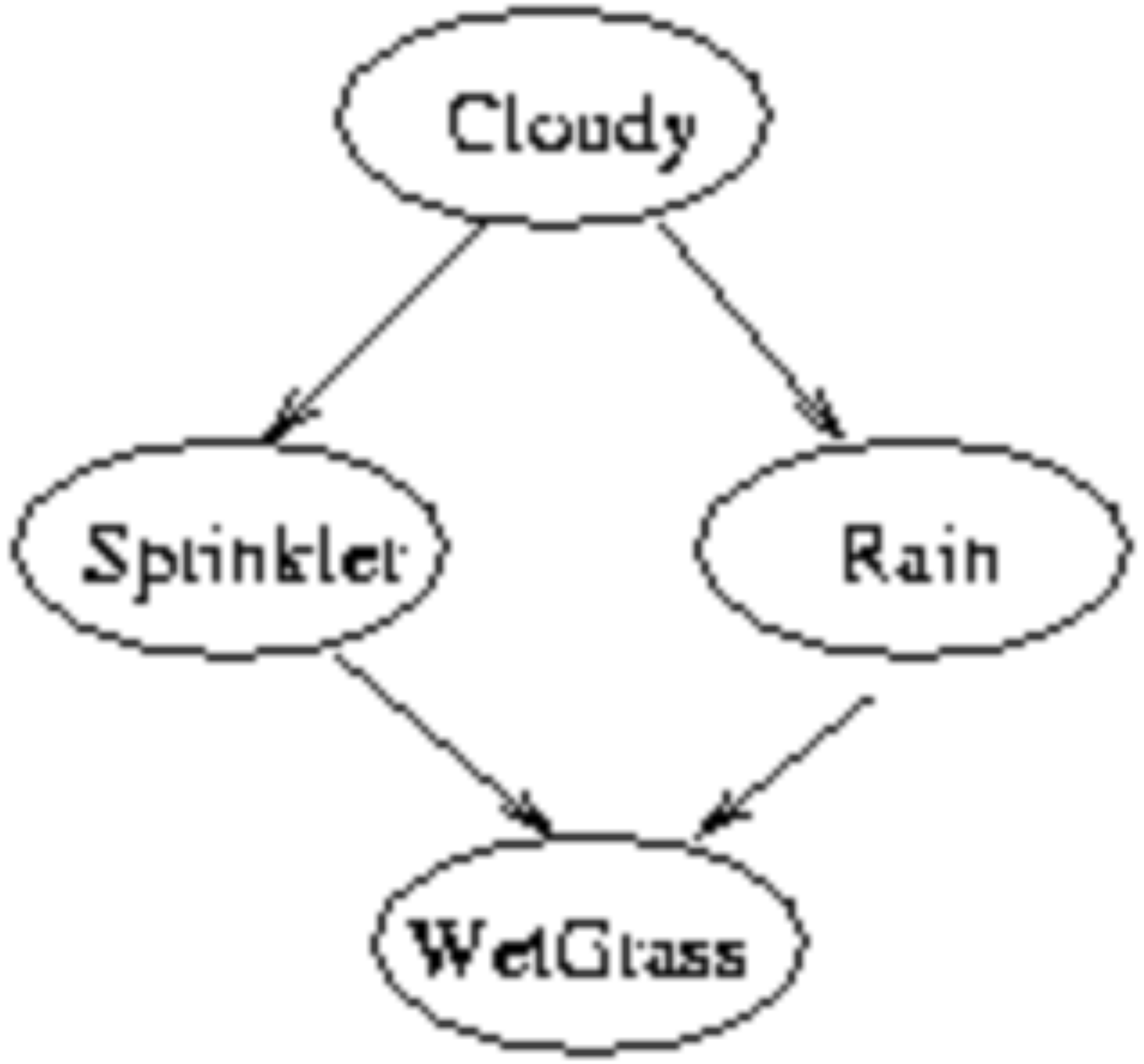
Conditional independence relationship encoded in a Bayesian network : a node is independent of its ancestors given its parents, where the ancestor/parent relationship is with respect to some fixed topological ordering of the nodes



# Conditional independence

Conditional independence relationship encoded in a Bayesian network : a node is independent of its ancestors given its parents, where the ancestor/parent relationship is with respect to some fixed topological ordering of the nodes




By the chain rule of probability, the joint probability of all the nodes in the graph above is

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

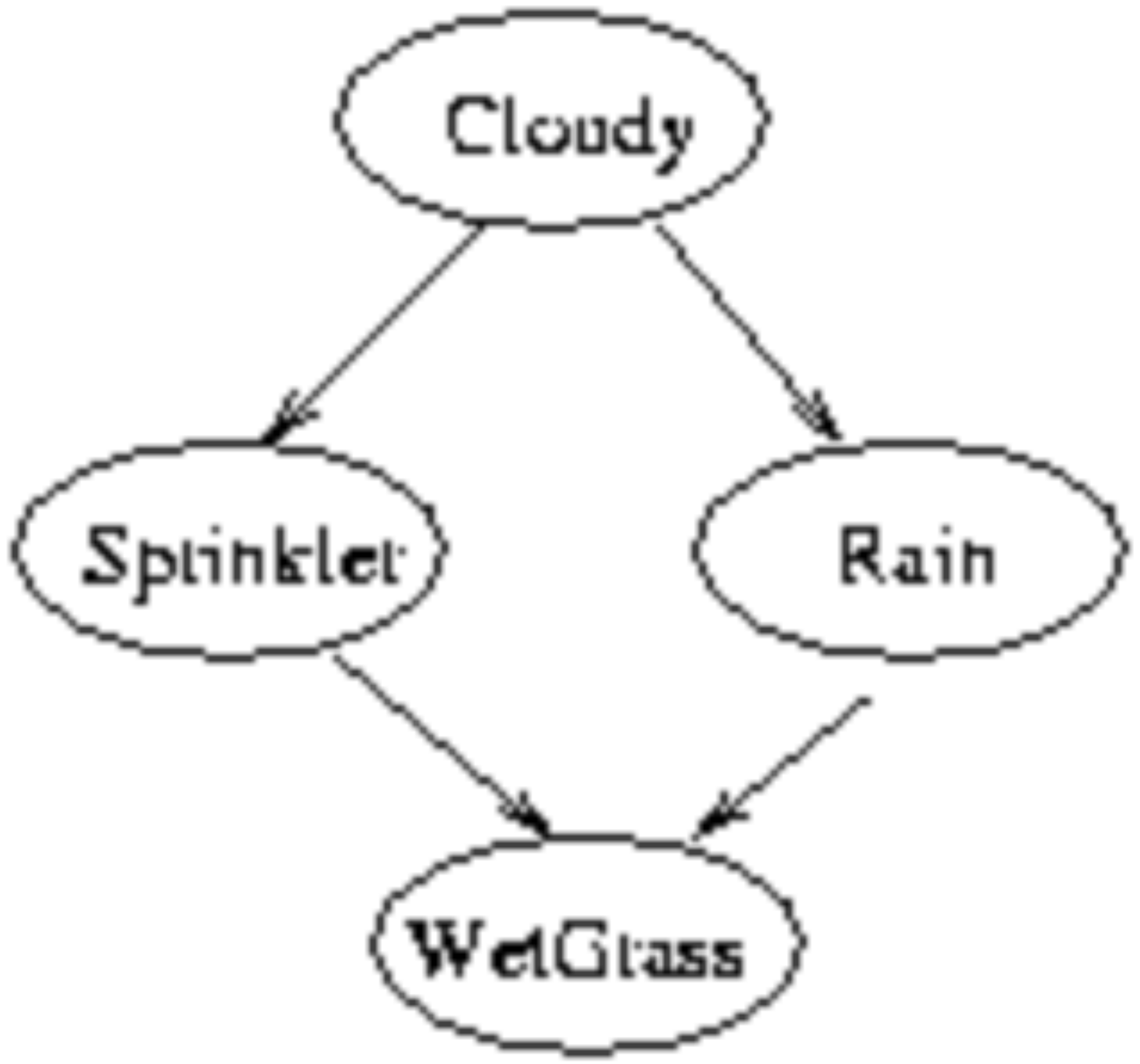
By using conditional independence relationships, we can rewrite this as

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

where we were allowed to simplify the third term because R is independent of S given its parent C, and the last term because W is independent of C given its parents S and R

# Conditional independence

Conditional independence relationship encoded in a Bayesian network : a node is independent of its ancestors given its parents, where the ancestor/parent relationship is with respect to some fixed topological ordering of the nodes

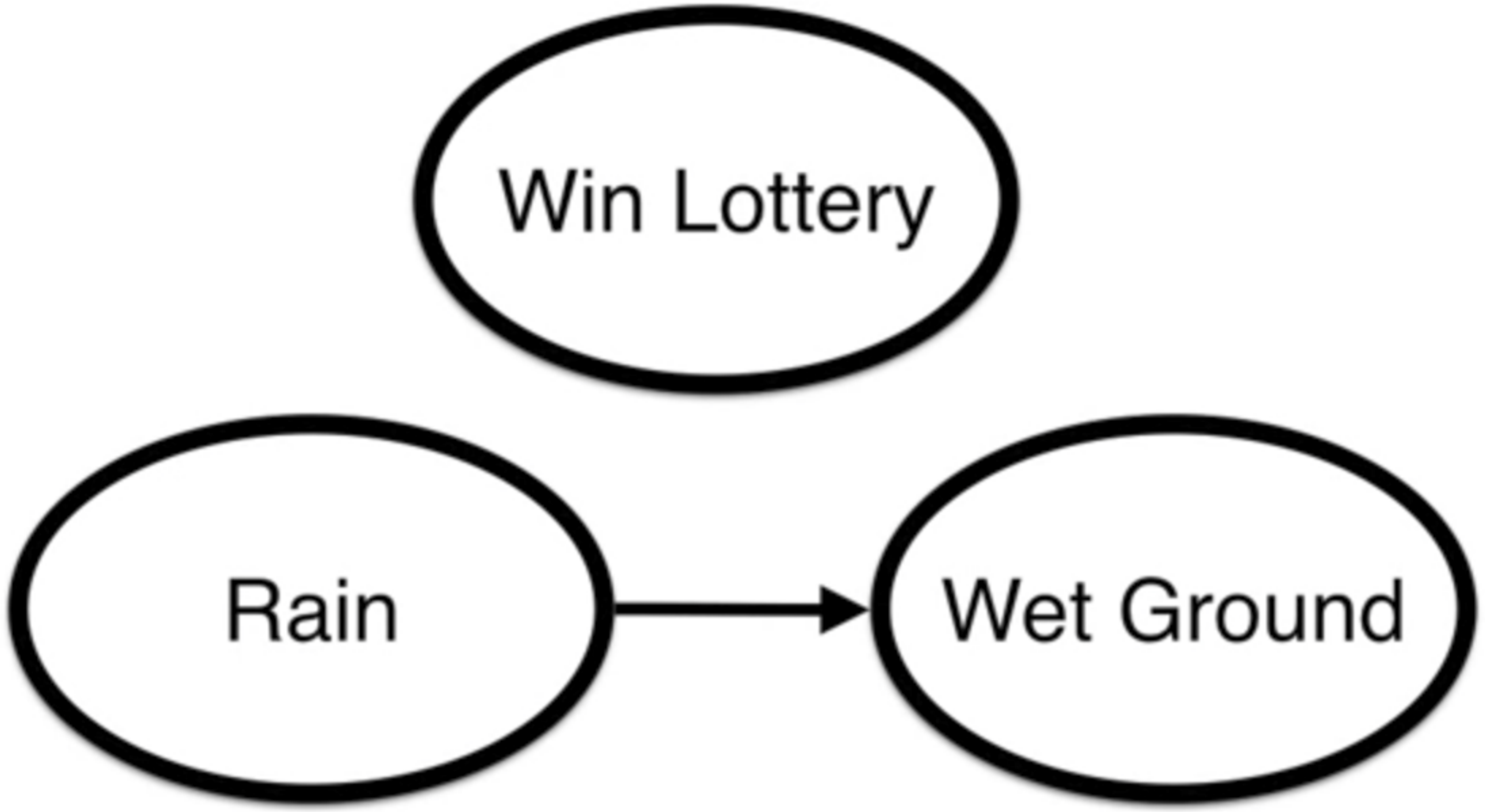



By using conditional independence relationships, we can rewrite this as

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

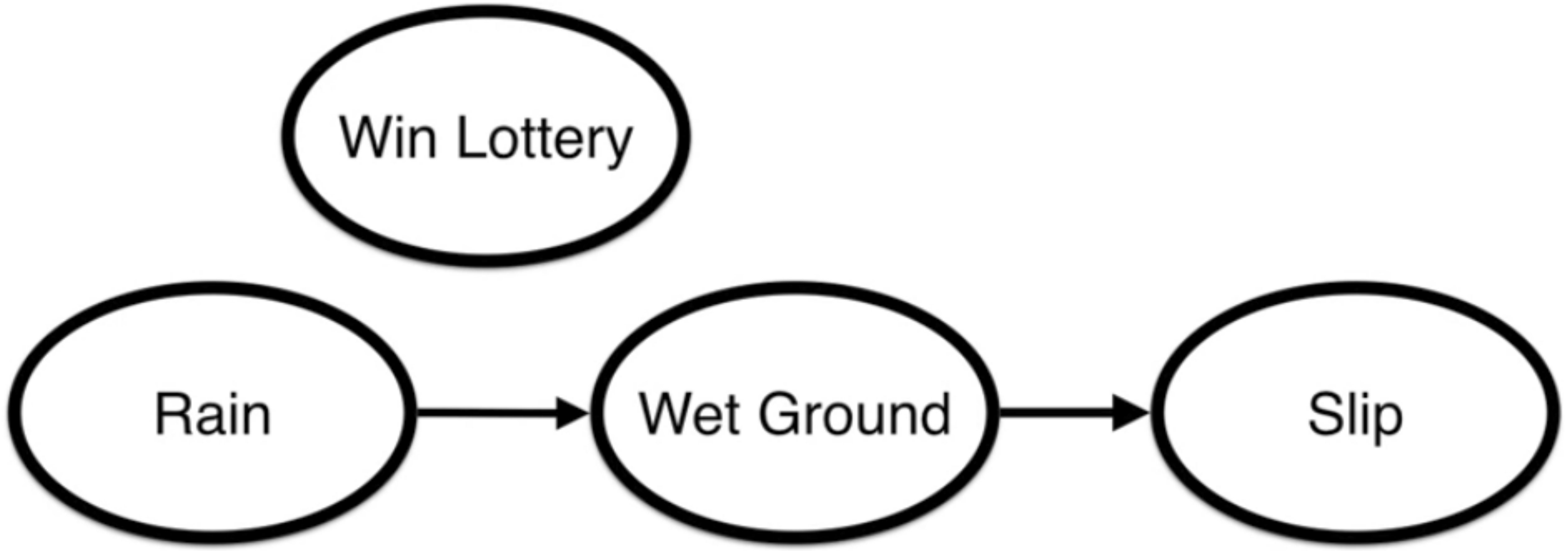
The absence of an arrow from Cloudy to Wet grass implies that the value of Cloudy has no direct influence on the value of Wet grass, only an indirect influence through Sprinkler and Rain. More formally, Wet grass is conditionally independent of Cloudy, given Sprinkler and Rain.

# Bayesian network



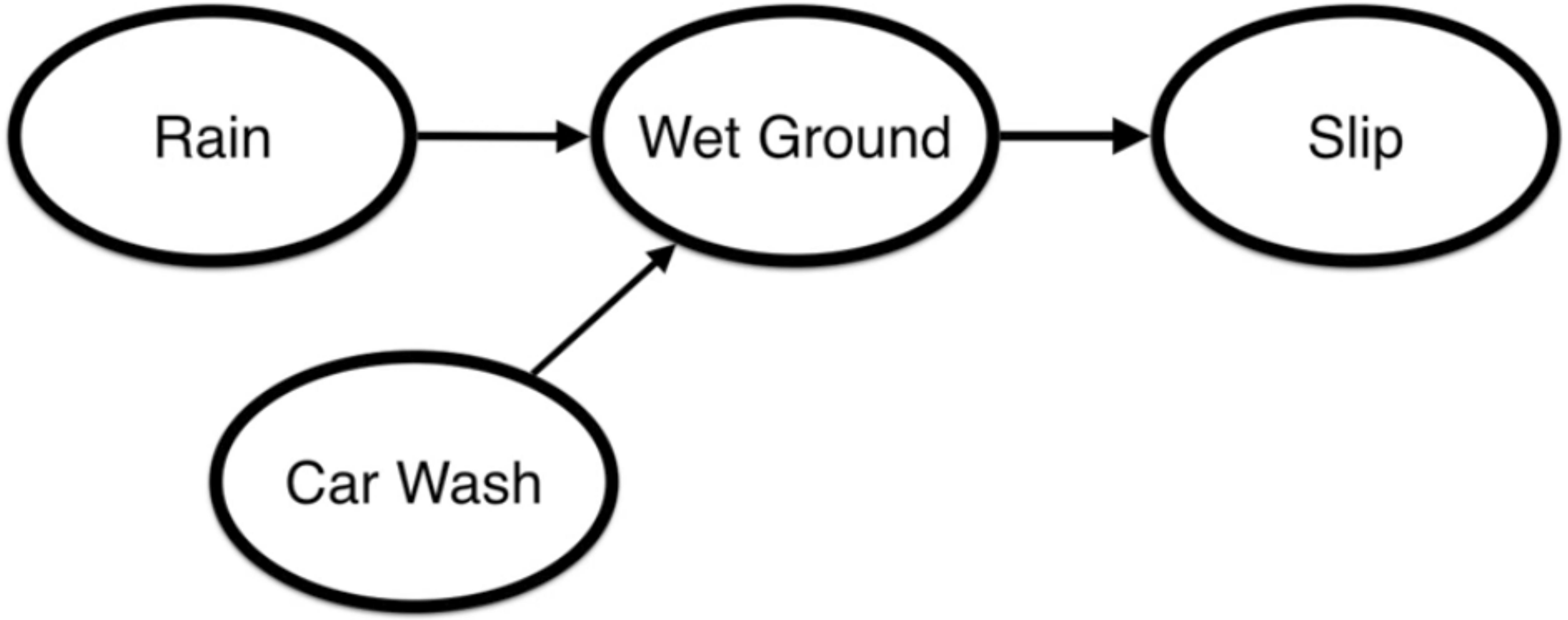

We are trying to express P(L,R,W)

# Bayesian network




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

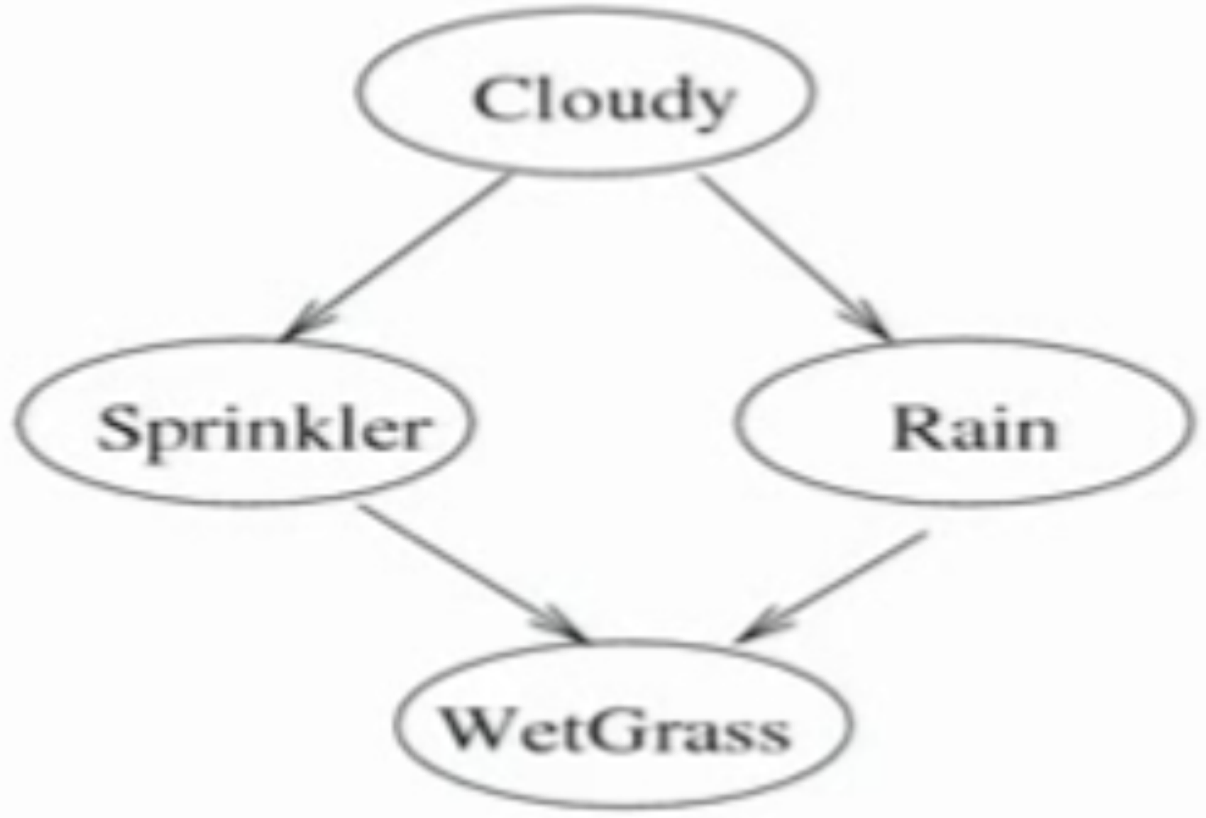
# Bayesian network




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

# Sprinkler network

| P(C=F) | P(C=T) |
|---|---|
| 0.5 | 0.5 |




| C | P(S=F) | P(S=T) |
|---|---|---|
| F | 0.5 | 0.5 |
| T | 0.9 | 0.1 |

| C | P(R=F) | P(R=T) |
|---|---|---|
| F | 0.8 | 0.2 |
| T | 0.2 | 0.8 |

| S | R | P(W=F) | P(W=T) |
|---|---|---|---|
| F | F | 1.0 | 0.0 |
| T | F | 0.1 | 0.9 |
| F | T | 0.1 | 0.9 |
| T | T | 0.01 | 0.99 |

P(C,S) = ?

P(W|S=0,R=1) = ?

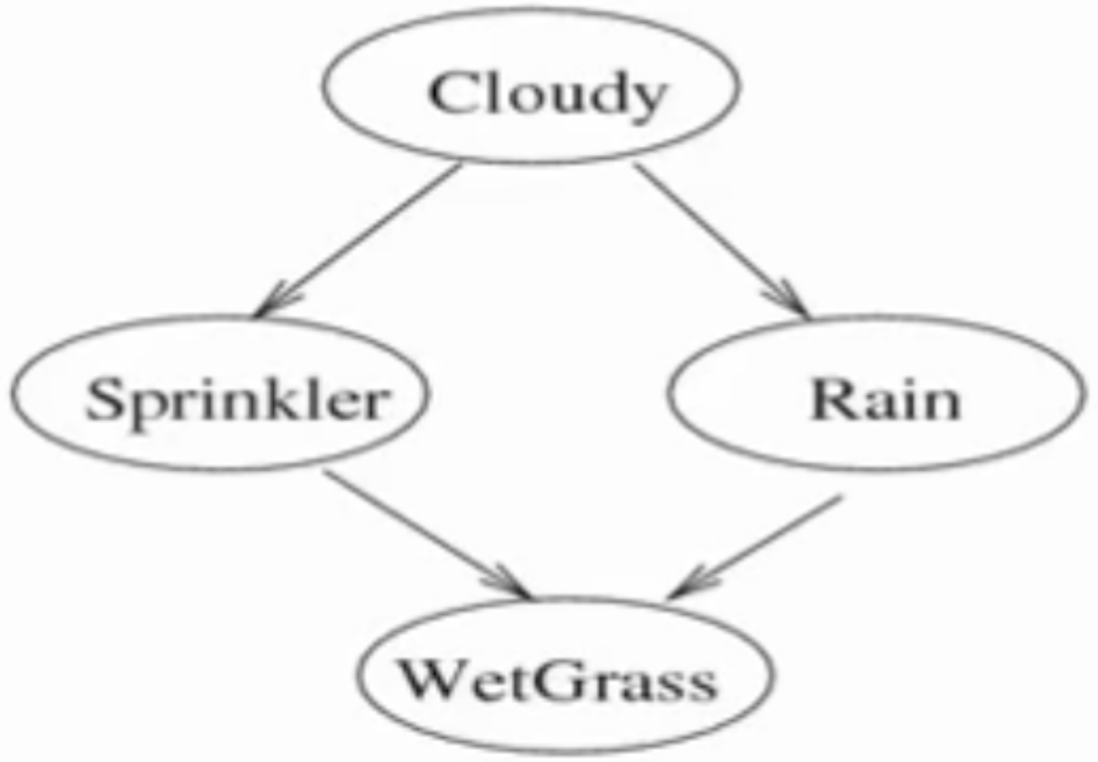
$$p(C, S, R, W) = p(C)p(S|C)p(R|C)p(W|S, R)$$

# Joint vs factored joint distributions

| c | s | r | w | prob |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0.200 |
| 0 | 0 | 0 | 1 | 0.000 |
| 0 | 0 | 1 | 0 | 0.005 |
| 0 | 0 | 1 | 1 | 0.045 |
| 0 | 1 | 0 | 0 | 0.020 |
| 0 | 1 | 0 | 1 | 0.180 |
| 0 | 1 | 1 | 0 | 0.001 |
| 0 | 1 | 1 | 1 | 0.050 |
| 1 | 0 | 0 | 0 | 0.090 |
| 1 | 0 | 0 | 1 | 0.000 |
| 1 | 0 | 1 | 0 | 0.036 |
| 1 | 0 | 1 | 1 | 0.324 |
| 1 | 1 | 0 | 0 | 0.001 |
| 1 | 1 | 0 | 1 | 0.009 |
| 1 | 1 | 1 | 0 | 0.000 |
| 1 | 1 | 1 | 1 | 0.040 |

$$p(C, S, R, W) = p(C)p(S|C)p(R|C)p(W|S, R)$$

| P(C=F) | P(C=T) |
|---|---|
| 0.5 | 0.5 |




| C | P(S=F) | P(S=T) |
|---|---|---|
| F | 0.5 | 0.5 |
| T | 0.9 | 0.1 |

| C | P(R=F) | P(R=T) |
|---|---|---|
| F | 0.8 | 0.2 |
| T | 0.2 | 0.8 |

| S | R | P(W=F) | P(W=T) |
|---|---|---|---|
| F | F | 1.0 | 0.0 |
| T | F | 0.1 | 0.9 |
| F | T | 0.1 | 0.9 |
| T | T | 0.01 | 0.99 |

# Inference

For example, consider the water sprinkler network, and suppose we observe the fact that the grass is wet. There are two possible causes for this: either it is raining, or the sprinkler is on. Which is more likely?

$$\Pr(S=1|W=1) = \frac{\Pr(S=1,W=1)}{\Pr(W=1)} = \frac{\sum_{c,r}\Pr(C=c,S=1,R=r,W=1)}{\Pr(W=1)} = 0.2781/0.6471 = 0.430$$

$$\Pr(R=1|W=1) = \frac{\Pr(R=1,W=1)}{\Pr(W=1)} = \frac{\sum_{c,s}\Pr(C=c,S=s,R=1,W=1)}{\Pr(W=1)} = 0.4581/0.6471 = 0.708$$

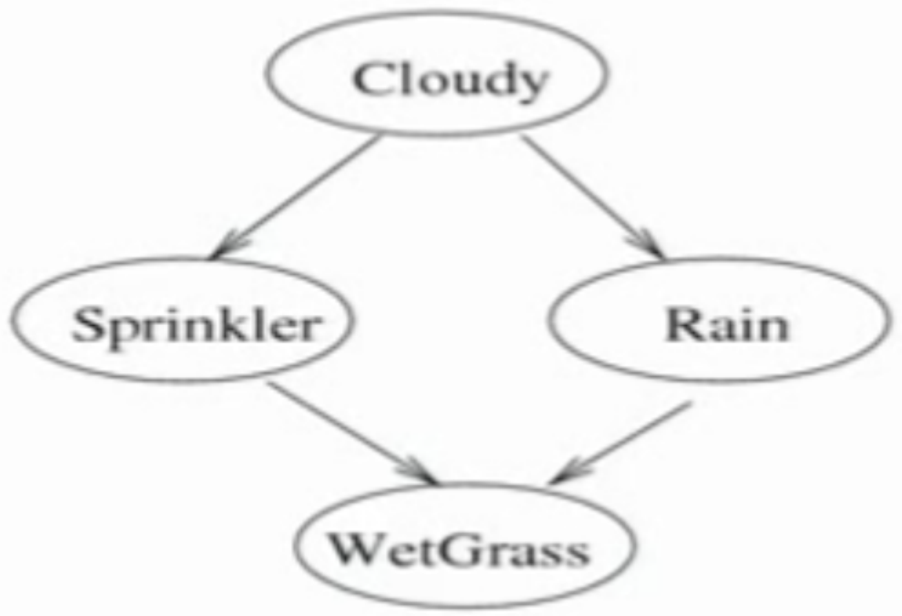
$$\Pr(W=1) = \sum_{c,r,s}\Pr(C=c,S=s,R=r,W=1) = 0.6471$$

# Inference

What is the marginal probability, P(S=1), that sprinkler is on?

Brute force?

| P(C=F) | P(C=T) |
|---|---|
| 0.5 | 0.5 |




| C | P(S=F) | P(S=T) |
|---|---|---|
| F | 0.5 | 0.5 |
| T | 0.9 | 0.1 |

| C | P(R=F) | P(R=T) |
|---|---|---|
| F | 0.8 | 0.2 |
| T | 0.2 | 0.8 |

| S | R | P(W=F) | P(W=T) |
|---|---|---|---|
| F | F | 1.0 | 0.0 |
| T | F | 0.1 | 0.9 |
| F | T | 0.1 | 0.9 |
| T | T | 0.01 | 0.99 |