Reasoning under Uncertainty

CPSC 470 – Artificial Intelligence
Brian Scassellati

STRIPS planner
STanford Research Institute Problem Solver

- "Holy Roman Empire" naming
- Represent states and goals in first-order logic

  At(Home) ∧ Have(Milk) ∧ Have(Drill) ∧
     Have(Banana)

- Assume existential quantification of variables

  At(x) ∧ Sells(x, Milk)

Why STRIPS is Insufficient for many Domains

- **Hierarchical plans**
  - Allow for more complex plans by varying level of abstraction
- **Resource Limitations**
  - Consumption and generation of resources
  - Time as a resource
    - Based on situation calculus, assumes all actions take place simultaneously and in one unit of time
    - Actions in a plan may have durations, deadlines, and time windows
- **Complex conditions**
  - No conditionals in STRIPS
  - No universals in STRIPS
- **Dealing with incomplete or inaccurate information**
  - Conditional planning
  - Execution monitoring

Failures of Logic

- First-order logic represents a certainty
  - $\forall p$ Symptom(p,Toothache) $\Rightarrow$ Disease(p,Cavity)
- To make the rule true, we must add an almost unlimited set of causes
  - $\forall p$ Symptom(p,Toothache) $\Rightarrow$ Disease(p,Cavity) $\vee$ Disease(p, GumDisease) $\vee$ Disease(p, SinusInfection) $\vee$ …
- Conversion to a causal rule does not help
 - $\forall p$ Disease(p,Cavity) $\Rightarrow$ Symptom(p,Toothache)

Why does Logic Fail?

- **Laziness**
  - Too much work to list entire sets of consequents or antecedents
  - List all possible causes for a toothache
- **Theoretical Ignorance**
  - No complete theory for the domain exists
  - Describe precisely the conditions that cause cancer
- **Practical Ignorance**
  - Even if we know all the rules, we may be uncertain about a particular event
  - What was the white blood cell count of the patient two years ago?

Probability provides a way of summarizing the uncertainty that comes from our laziness and ignorance

Basic Probability

- Assume a basic understanding of probability theory, but as a quick review:
- Unconditional (prior) probability:
  - P(Cavity) = 0.1
- Random Variables
  - P(*Weather* = snow) = 0.05
- Probability distribution
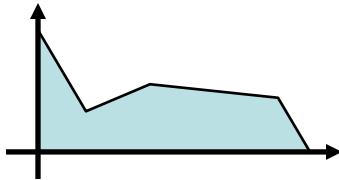
- Conditional (posterior) probability:
  - P(Cavity|Toothache) = 0.8

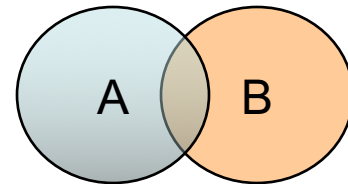# Basic Probability II

- Axioms

  $0 \leq P(A) \leq 1$

  $P(\text{True}) = 1$

  $P(\text{False}) = 0$

  $P(A \lor B) = P(A) + P(B) - P(A \land B)$

  $P(\neg A) = 1 - P(A)$

  $P(A \land B) = P(A|B)\, P(B)$ (the product rule)

Bayes' Rule

- From the product rule

  $P(A \land B) = P(A|B)\ P(B)$

  $P(A \land B) = P(B|A)\ P(A)$

- Bayes' Rule:

  $P(A|B)\ P(B) = P(B|A)\ P(A)$

  $$P(B|A) = \frac{P(A|B)\ P(B)}{P(A)}$$

Application of Bayes' Rule

- **Medical example:**
  - 1 in 20 patients reports a stiff neck
  - 1 in 50,000 patients has meningitis
  - Meningitis causes a stiff neck 50% of the time
  - If I have a stiff neck, what is the chance that I have meningitis?

- **Apply Bayes' Rule:**

  P(S) = 1/20

  P(M) = 1/50000

 P(S|M) = 0.50

  $$P(M|S) = \frac{P(S|M)P(M)}{P(S)}$$

  $$P(M|S) = \frac{0.5 \times 1/50000}{1/20}$$

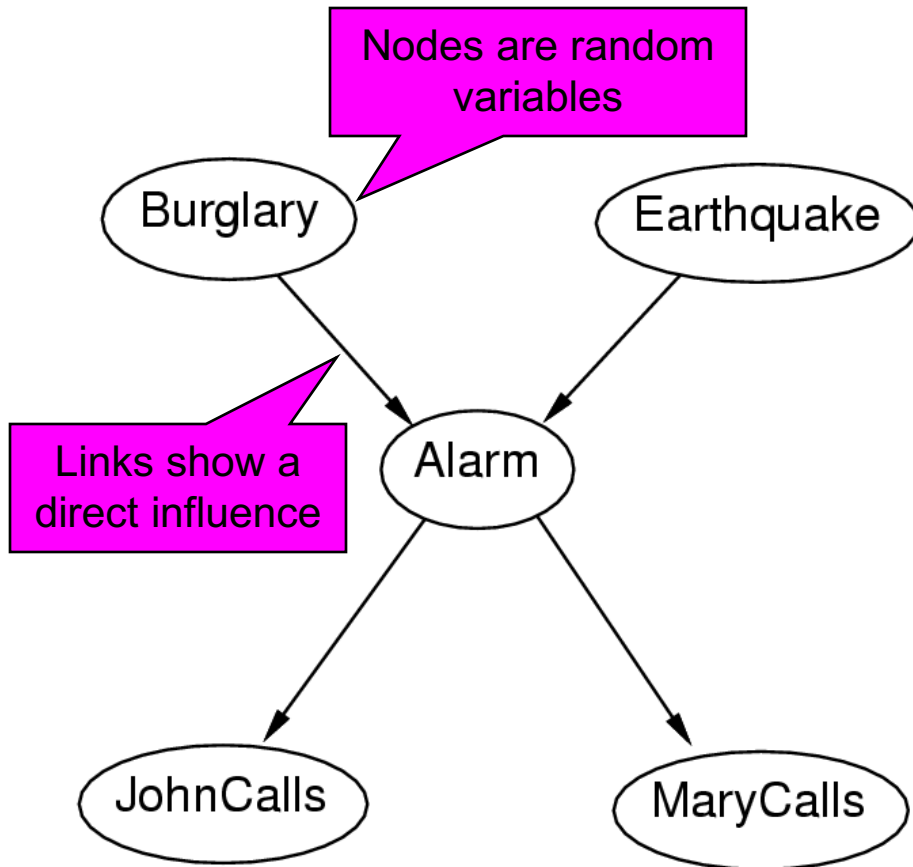 P(M|S) = 0.0002

Uncertainty and Rational Decisions

- Consider these plans:
  - Plan #1: Pay $10 for a 75% chance of event X
  - Plan #2: Pay $30 for a 10% chance of event X
  - Plan #3: Pay $90 for a 79% chance of event X
- Which plan should you choose?
  - If the event X is "receive 2 bonus points on PS5"
  - If the event X is "surviving an operation"
- An agent must have a set of preferences in order to make decisions in an uncertain world
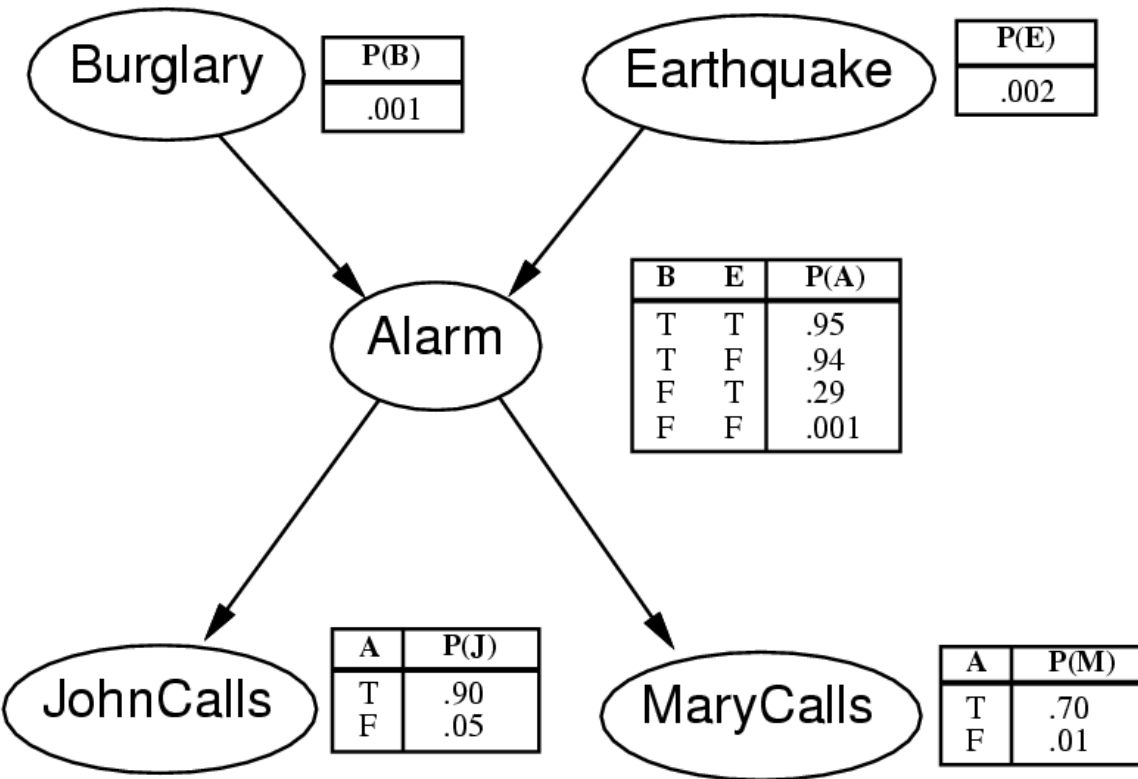
Probabilistic Reasoning Systems
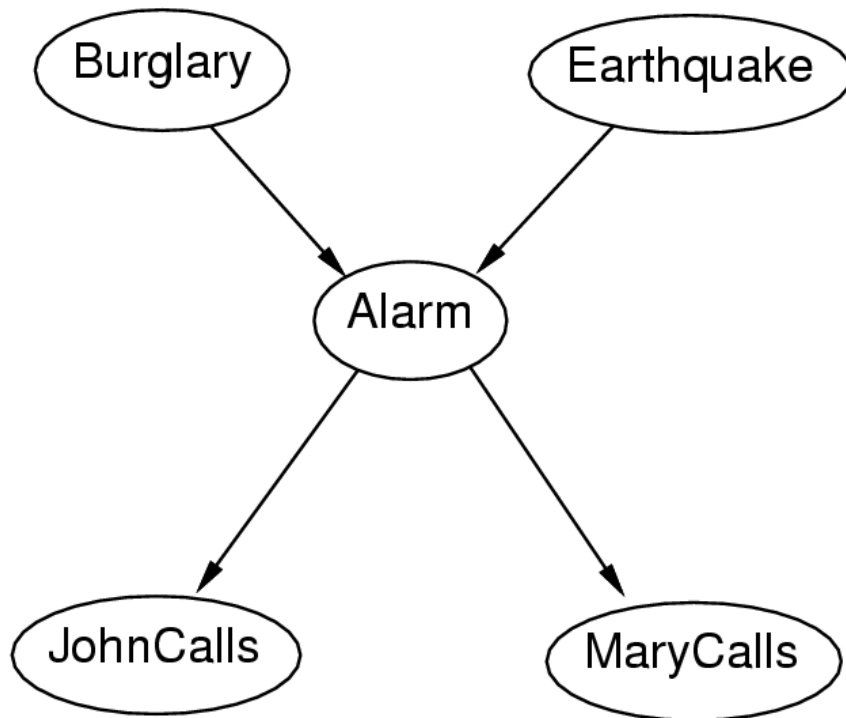
- We've seen the syntax and semantics of probability
- Now we look at an inference mechanism: Belief networks

A Basic Belief Network

Nodes are random variables

Burglary

Earthquake

Links show a direct influence

Alarm

JohnCalls

MaryCalls

- You have a new home alarm that responds
  - Accurately to burglaries
  - Occasionally responds to earthquakes
- When the alarm rings, your neighbors call you at work
  - John always calls, but sometimes confuses the telephone for the alarm
 - Mary sometimes misses the alarm, but only calls when the alarm actually rings

A Basic Belief Network

Burglary

P(B)
.001

Earthquake

P(E)
.002

Alarm

B	E	P(A)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

JohnCalls

A	P(J)
T	.90
F	.05

MaryCalls

A	P(M)
T	.70
F	.01

A conditional probability table
gives the likelihood of a
particular combination of values

- You have a new home alarm that responds
  - Accurately to burglaries
  - Occasionally responds to earthquakes
- When the alarm rings, your neighbors call you at work
  - John always calls, but sometimes confuses the telephone for the alarm
  - Mary sometimes misses the alarm, but only calls when the alarm actually rings

Conditional Independence Relations in Belief Nets
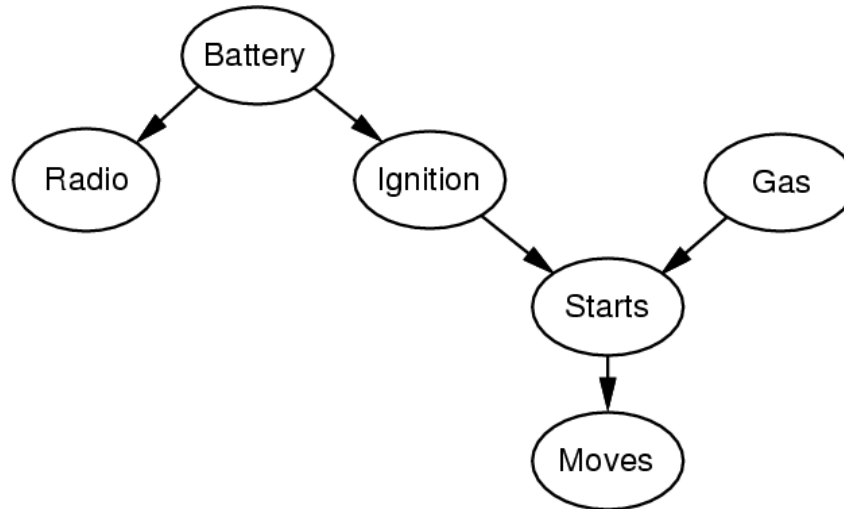
- Each variable is conditionally independent of its non-descendants, given its parents
- Presence of Gas and a working Radio are
  - Independent given evidence about the ignition
  - Independent given evidence about the battery
  - Independent given no evidence
  - Dependent given evidence that the car starts
    - If the car does not start, but the radio plays, then the chance of being out of gas is increased

Incremental Construction of Belief Nets

- Conditional Independence Property

  $P(X \mid A, B, C\ldots) = P(X \mid Parents(X))$

  where Parents(X) gives those nodes (A, B, etc.) that are the parent nodes of X

- Incremental belief net construction:

  1. Choose the set of relevant variables
  2. Choose an ordering for the variables
  3. While there are variables left:
     a. Pick a variable *a1* and add a node to the network for it.
     b. Set *Parents(a1)* to the minimal set of nodes that satisfies the conditional independence property
     c. Define the conditional probability table for node *a1*

Incremental Construction
of Belief Networks
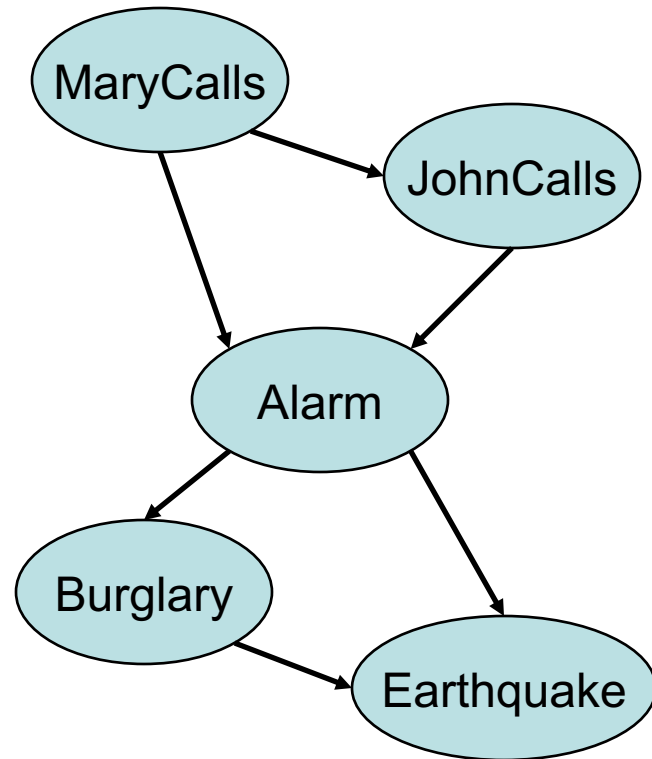
- Add *MaryCalls*
- Add *JohnCalls*
  - Dependence with *MaryCalls* since P(John|Mary)≠P(John)
- Add Alarm
  - More likely if both calls are made
- Add Burglary
  - Phone calls don't tell us anything about the chance of a burglar, but the alarm does
- Add Earthquake
  - Alarm acts as earthquake predictor
  - Presence of a burglar helps determine whether or not an earthquake occurred

# Incremental Construction
# of Belief Networks
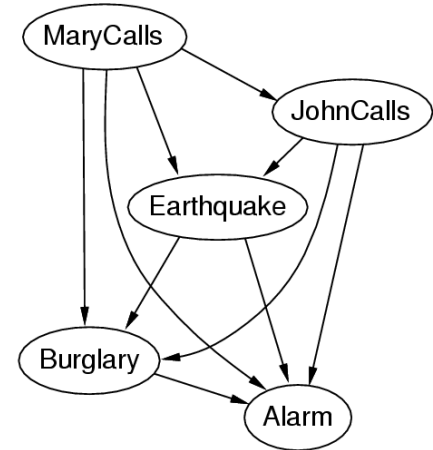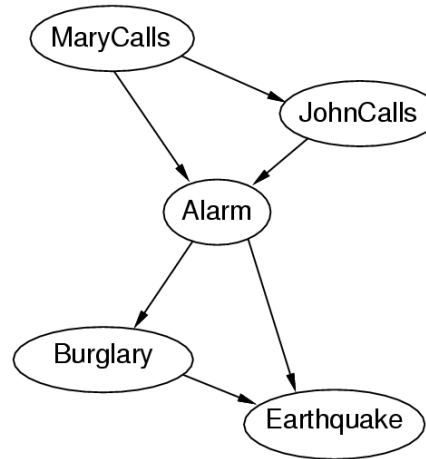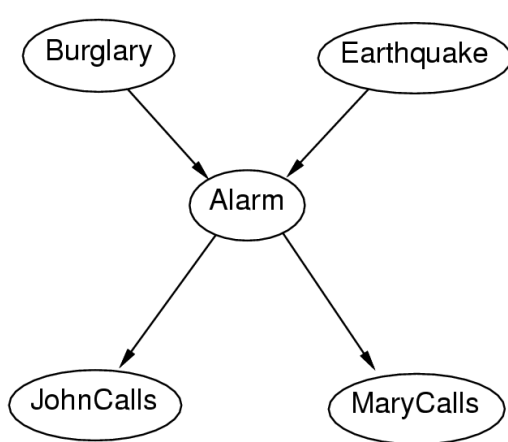
- Order in which you add nodes can make a difference on the number of links
- "Correct order" to add nodes is to add the "root causes" first and then the variables they influence, and so on…
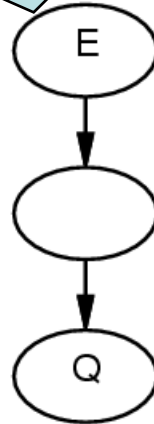- If we stick to a causal model, we need fewer probabilities and these probabilities will be easier to create

Types of Inference

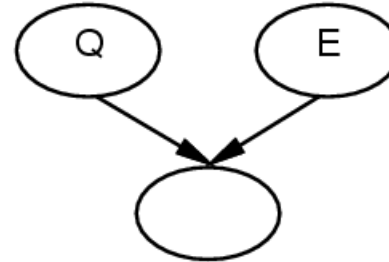Given that a burglary occurred, compute the chance that John calls

Given that John calls and that there was an earthquake, compute the chance of the alarm going off

Given that John calls, infer the chance of a burglary
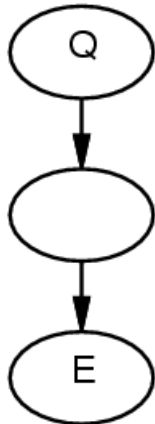
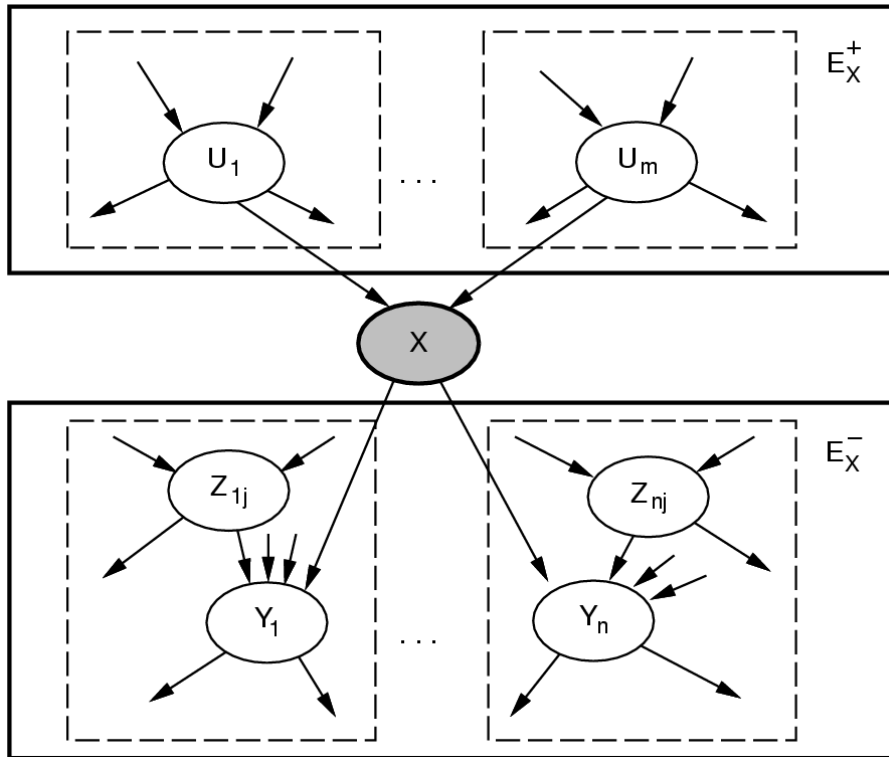**Diagnostic**   **Causal**   (Explaining Away) **Intercausal**   **Mixed**

Given that the alarm rang and there was an earthquake, what is the chance that there was a burglary?

E = evidence
Q = query

Inference with Belief Nets

- Singly connected graph
  - only one path between any two nodes

- Inference in singly-connected graphs
  - Causal support for X:
    - Evidence "above" X (from the direction of X's parents)
  - Evidential support for X:
    - Evidence "below" X (from the direction of X's children)

- Multiply connected graph
  - multiple paths between nodes

# Three Techniques for Inference in Multiply Connected Belief Networks

- ## Clustering
  - Transform the network by merging nodes
    - Probabilistically equivalent
    - Topologically different

- ## Cutset Conditioning
  - Transform by instantiating variables to values and re-evaluating the network

- ## Stochastic Simulation
  - Generate many consistent concrete models and approximate an exact evaluation

Clustering in Belief Networks
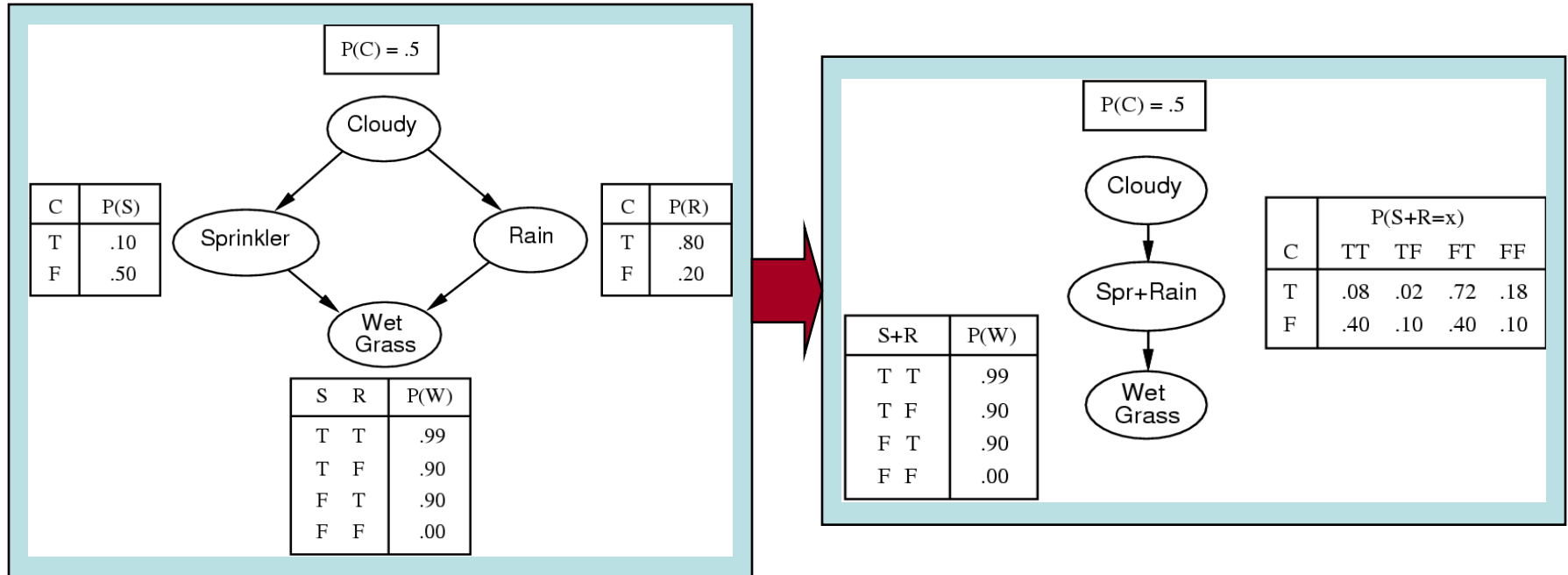
- Try to change a multiply-connected graph into a singly-connected graph by merging nodes
- Introduces complexity into each merged node, but the payoff is in the use of simpler inference mechanisms

Cutset Conditioning
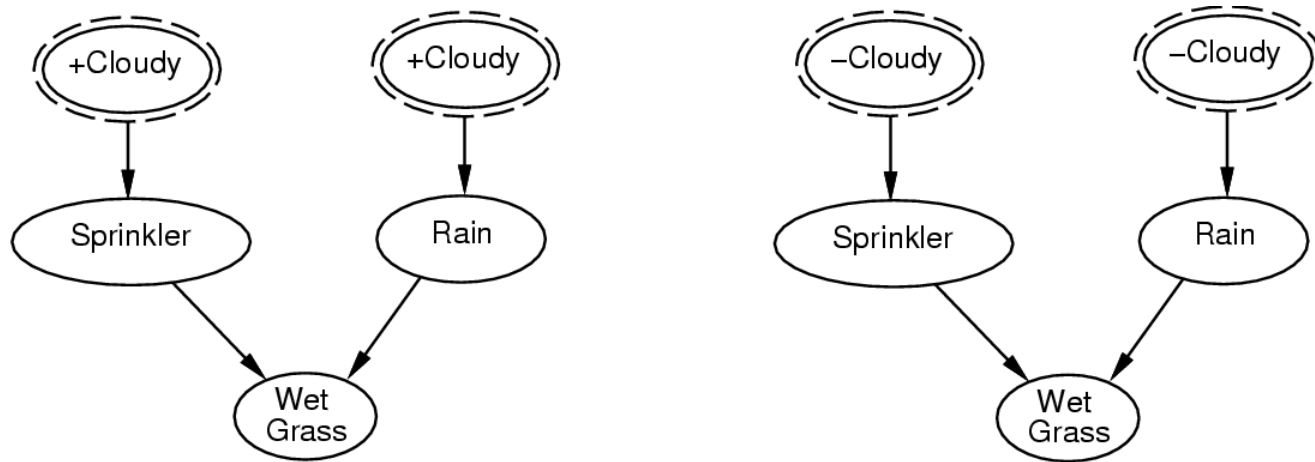
- "Opposite" of clustering: transforms the network into several simpler graphs

  - Number of graphs is exponential in the size of the cutset

- Transform by instantiating variables to values and re-evaluating the network

Stochastic Simulation Methods
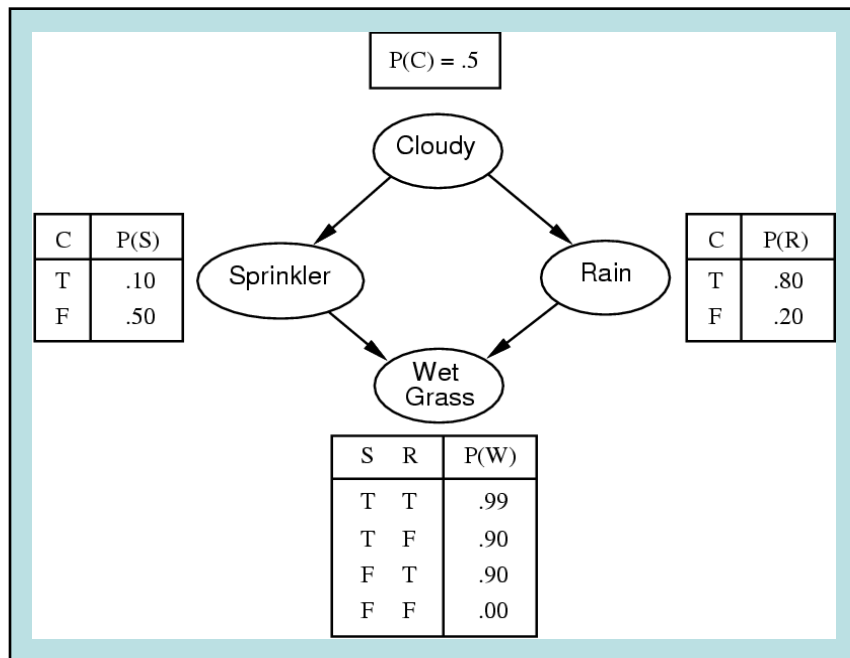
- If we want to estimate P(WetGrass | Rain) then we just start at a root node and simulate a large number of trials
  - Choose a value for Cloudy based on the probability
  - Propagate this value through the network
  - Count the number of instances of WetGrass and Rain to estimate the probability

Where do Probabilities come from?

- **Frequentist**
  - Probabilities come from experiments
  - "9 out of 10 dentists agree"

- **Objectivist**
  - Probabilities are real aspects of the universe
  - Propensities of objects to act in certain ways

- **Subjectivist**
  - Probabilities characterize an agent's beliefs
  - "In my opinion, there is a 30% chance of success"

Computing Probabilities and Reference Classes

- Probability that the sun will exist tomorrow

Undefined	There has never been an experiment that tested the existence of the sun tomorrow
1	In all previous experiments (previous days), the sun has continued to exist
1-ε	Where ε is the proportion of stars in the universe that go nova every day
$\frac{(d+1)}{(d+2)}$	Where d is the number of days that the sun has existed so far (formula due to Laplace)
???	Can be derived from the type, age, size, and temperature of the sun

Administrivia

- Coming up next…learning!