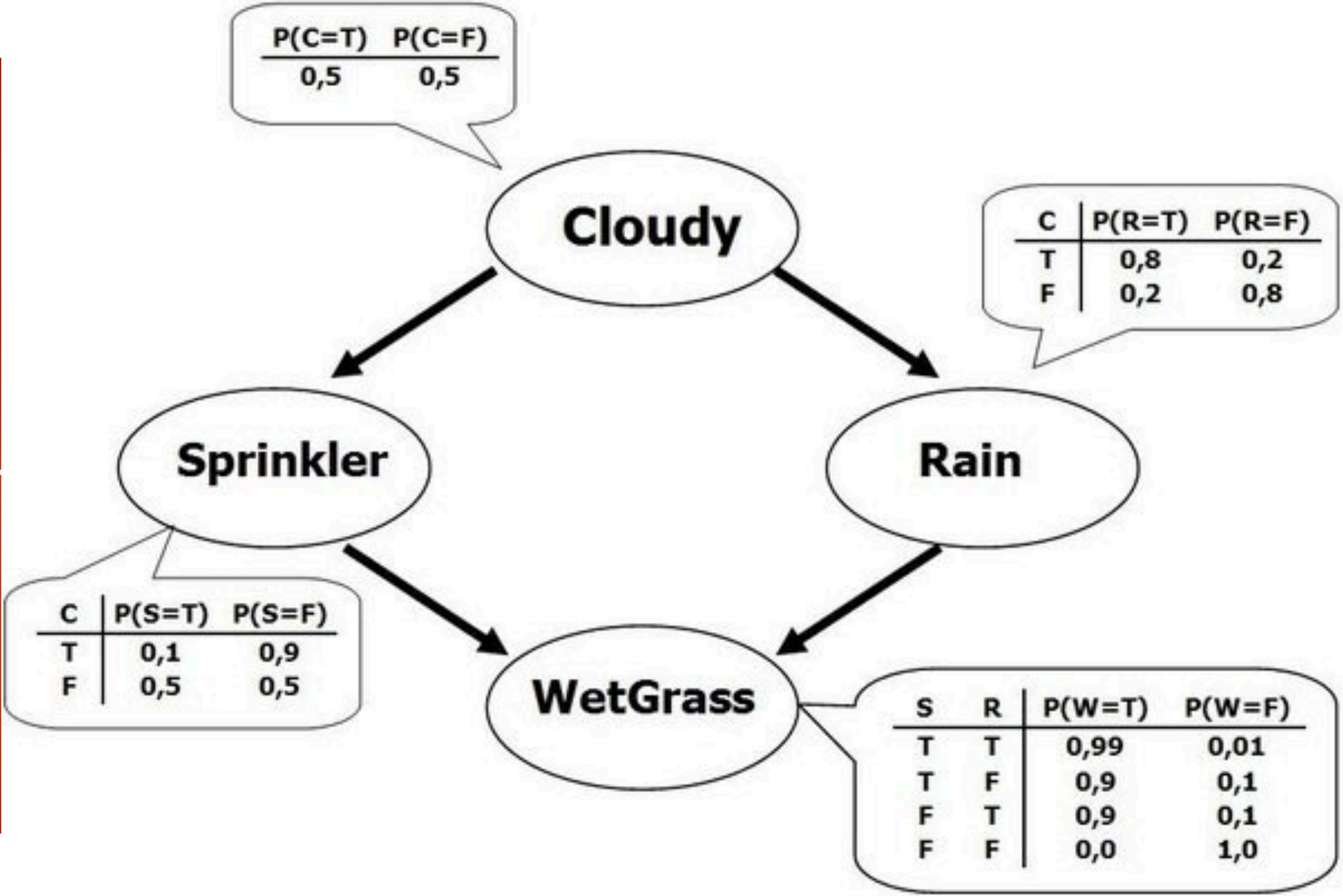
Causality

Bayesian Networks

Conditional Dependence

- Let A be an event
- Prob p(A) = fraction of instances recorded on which event A occurs
- Prob p(A | B) = fraction of instances recorded on which event A occurs, but counted only for those instances when B occurs

Bayes Network – Example

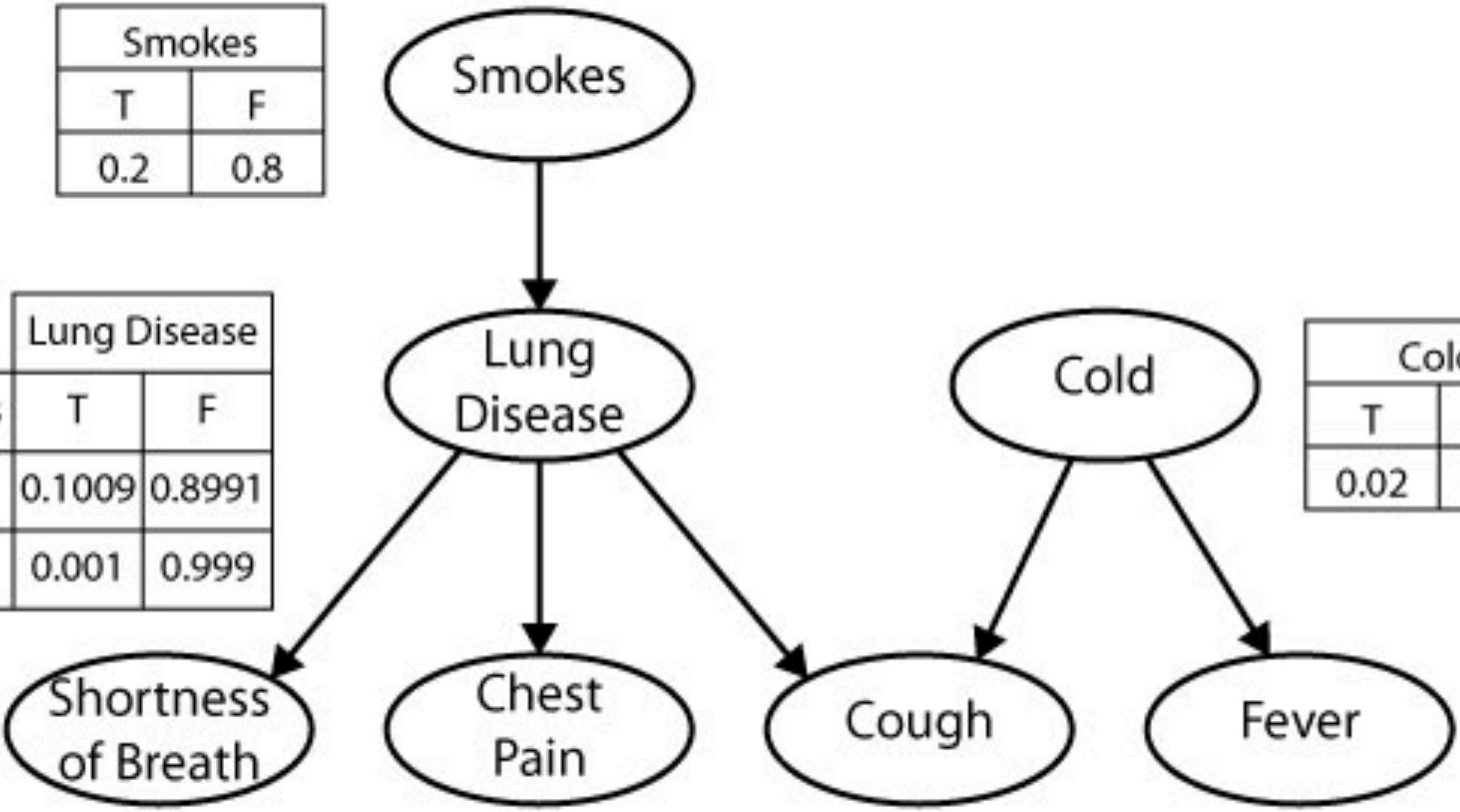


P(C=T)	P(C=F)
0,5	0,5

C	P(S=T)	P(S=F)
T	0,1	0,9
F	0,5	0,5

C	P(R=T)	P(R=F)
T	0,8	0,2
F	0,2	0,8

S	R	P(W=T)	P(W=F)
T	T	0,99	0,01
T	F	0,9	0,1
F	T	0,9	0,1
F	F	0,0	1,0



Smokes	
T	F
0.2	0.8

	Lung Disease	
Smokes	T	F
T	0.1009	0.8991
F	0.001	0.999

Cold	
T	F
0.02	0.98

	Shortness of Breath	
Lung Disease	T	F
T	0.208	0.792
F	0.01	0.99

	Chest Pain	
Lung Disease	T	F
T	0.208	0.792
F	0.01	0.99

		Cough	
Lung Disease	Cold	T	F
T	T	0.7525	0.2475
T	F	0.505	0.495
F	T	0.505	0.495
F	F	0.01	0.99

	Fever	
Cold	T	F
T	0.307	0.693
F	0.01	0.99

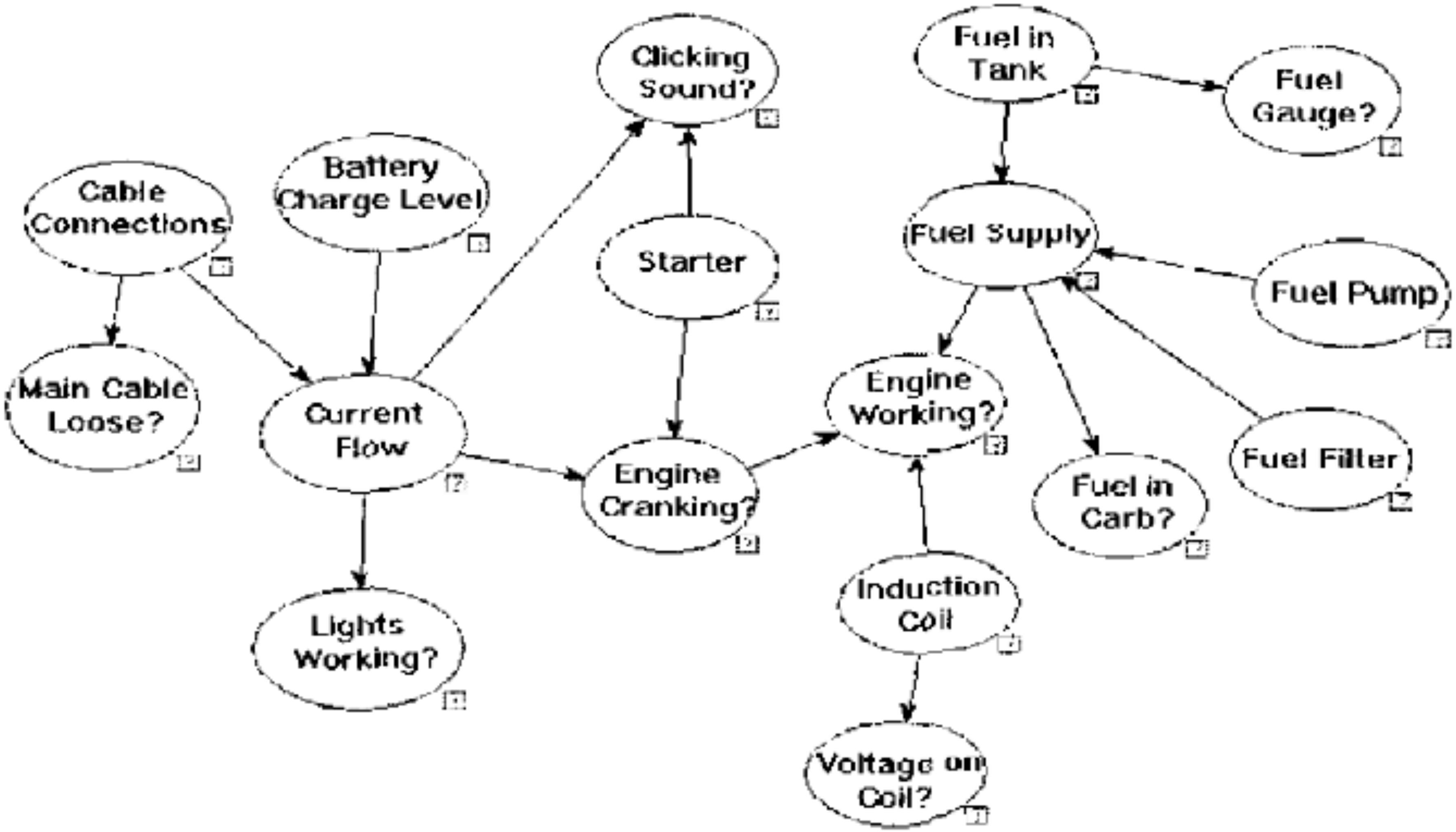
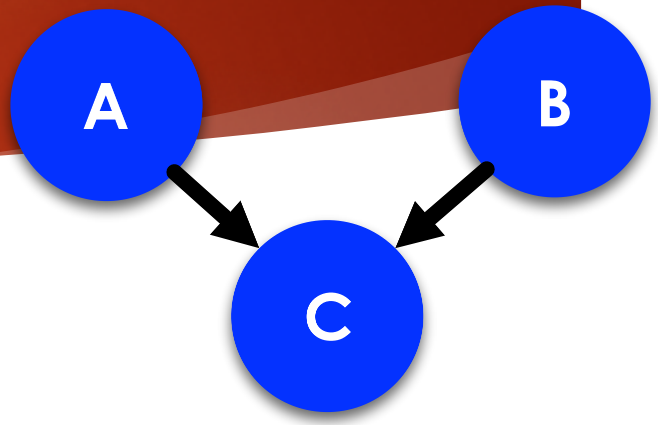


Figure 1. Bayesian Network for Example of Car Diagnostics

Bayesian Networks

- More often than not, two variables are independent or conditionally independent.
- Helps to cut down edges in a network of dependencies

Conditional Dependence in BN



- Consider situation shown here:
- We expect p(A | B) = p(A), i.e., A is independent of B
- What happens if C occurs?
  - If B occurs, the p(A) decreases since it is less critical to explain occurrence of C
  - I.e., p(A | B,C) < p(A | C) & p(B | A,C) < p(B | C)

E.g. of Conditional dependence

	1	2	3	4
A	0	1	0	1
B	0	0	1	1
C	0	1	1	1

- $p(A|B) = \frac{1}{2}$; $p(A) = 2/4 = \frac{1}{2}$;
- Since $p(A|B) = p(A)$, A is independent of B
- $p(A|B,C) = \frac{1}{2} < p(A|C) = 2/3$;

Causality

Causality

- Correlation doesn’t imply causation
- Examples: Drugs, Gene Regulatory
- Causal revolution in the last decade
- High impact in many domains
- Causality can shed light on Bioinformatics

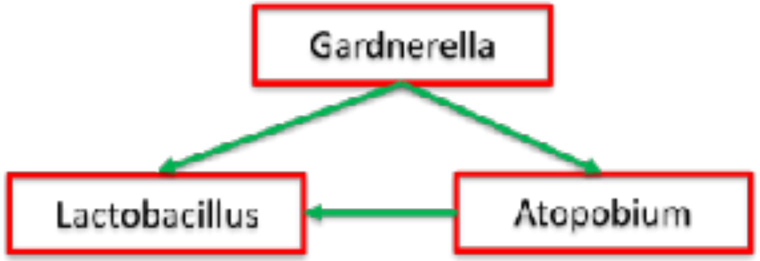
Judea Pearl

Steps in Causal Inference

- Causal Models
  - Causal structure / Causal Bayesian network
  - Casual parameters
- Causal Effects
  - Causal inference
  - Quantification of the causal influence

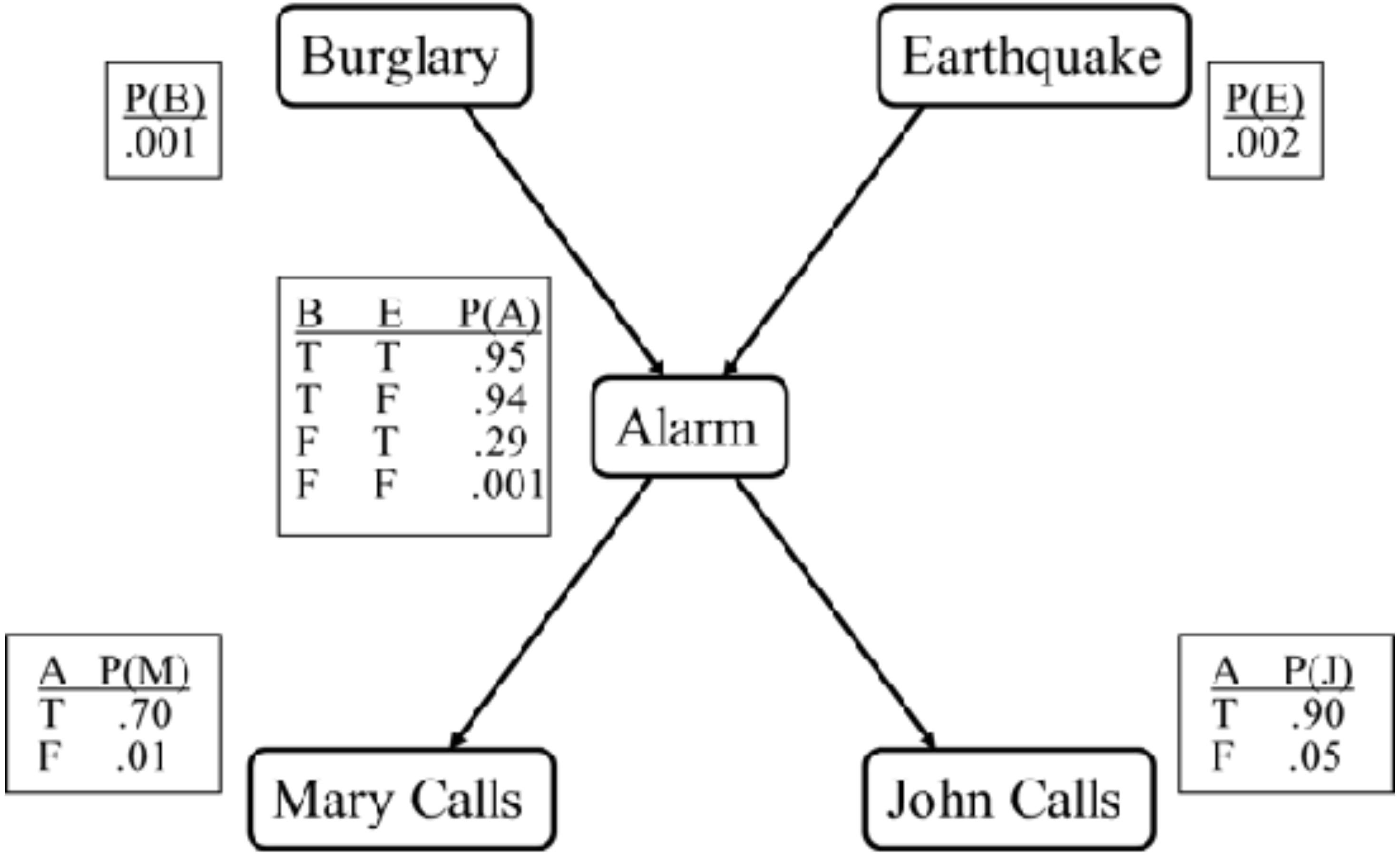
Causal Bayesian Network

- A class of Bayesian networks
  - Directed Acyclic Graph (DAG)
    - Set of nodes, set of directed edges, no cycle
  - Nodes represent random variables
  - Edges represent conditional relationships

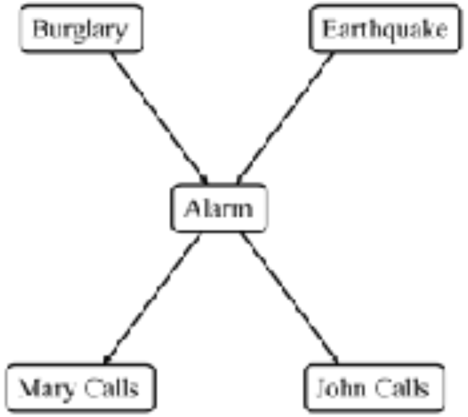


-

Joint Distribution



Complex Inferencing



-

- E.g.

  $$P(J \,\&\, M \,\&\, A \,\&\, \neg B \,\&\, \neg E)$$

  $$= P(J \mid A)P(M \mid A)P(A \mid \neg B, \neg E)P(\neg B)P(\neg E)$$

  $$= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998$$

  $$= 0.00062$$

Causal Chains

- This configuration is a **causal chain**

X: Low pressure Y: Rain Z: Traffic

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

- $X \not\perp Z$
- $X \perp Z \mid Y$

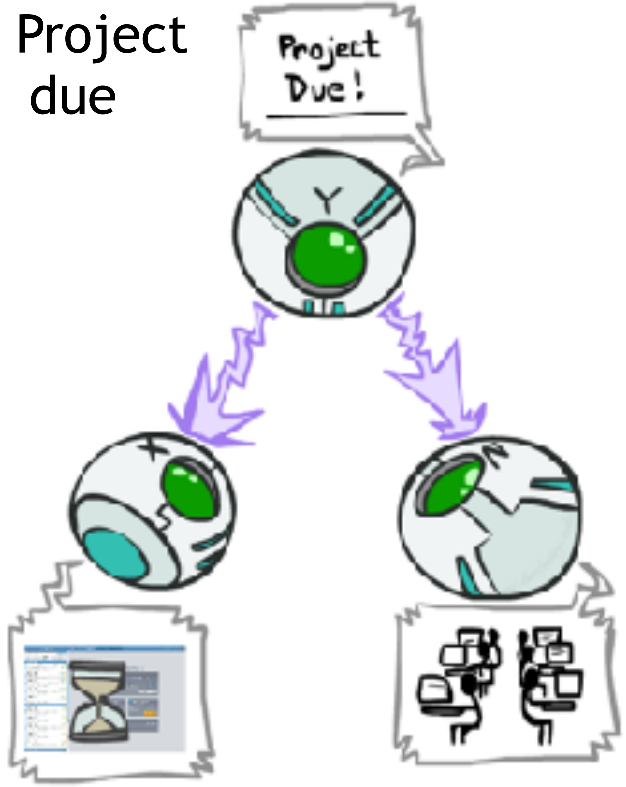
Common Cause

- This configuration is a "common cause"

$$P(x, y, z) = P(y)P(x|y)P(z|y)$$

- $X \not\perp Z$
- $X \perp Z \mid Y$

Y: Project due



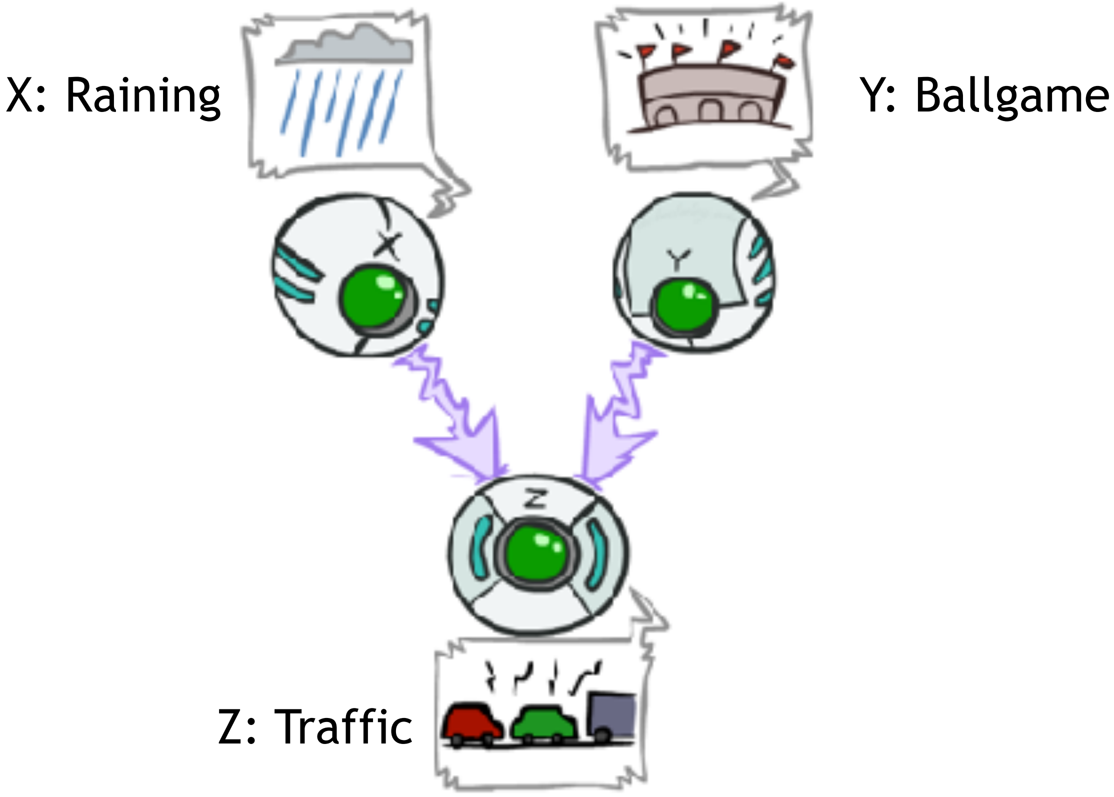
X: Forums busy

Z: Lab full

Common Effect

- Two causes of one effect (v-structures)

  - $X \perp Z$
  - $X \not\perp Z \mid Y$

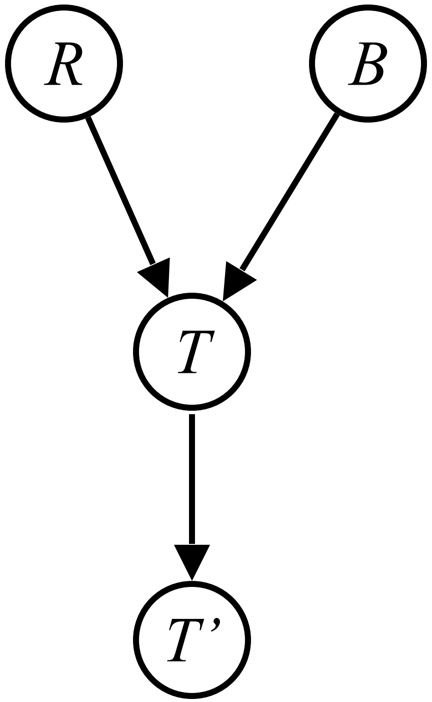


Example

$R \perp\!\!\!\perp B$ *Yes*

$R \perp\!\!\!\perp B | T$

$R \perp\!\!\!\perp B | T'$



Example

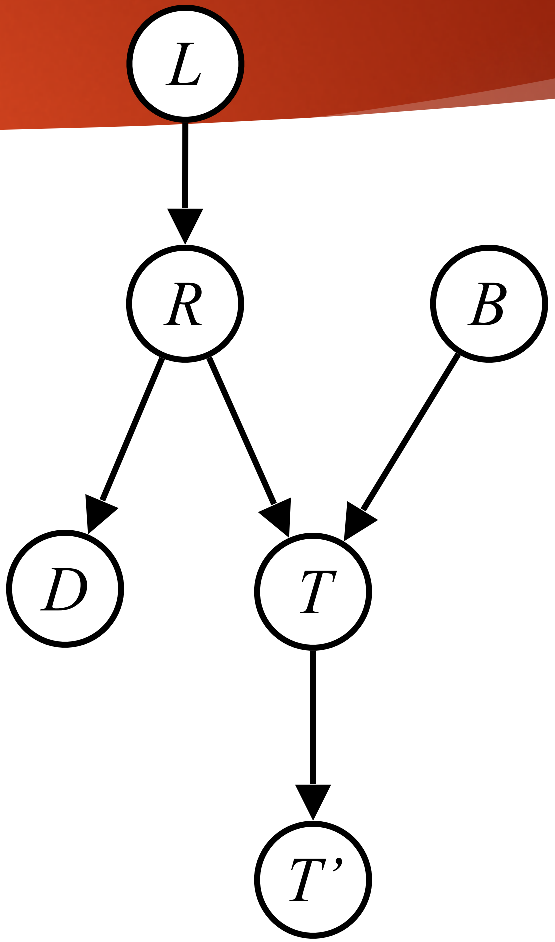
$L \perp\!\!\!\perp T' | T$ *Yes*

$L \perp\!\!\!\perp B$ *Yes*

$L \perp\!\!\!\perp B | T$

$L \perp\!\!\!\perp B | T'$

$L \perp\!\!\!\perp B | T, R$ *Yes*



Example

- Variables:
  - R: Raining
  - T: Traffic
  - D: Roof drips
  - S: I'm sad
- Questions:

$$T \perp\!\!\!\perp D$$

$$T \perp\!\!\!\perp D | R$$ *Yes*

$$T \perp\!\!\!\perp D | R, S$$