FIT5047 Module 4

Bayes Net

授课老师：Joe

Bayesian Network

FIT5047平时班 – Module 4

Decision making under uncertainty – what action to take when the state of the world is unknown

Bayesian answer –
Find the utility of each possible outcome (action-state pair), and take the action that maximizes the *expected utility*

FIT5047平时班 – Module 4

Action	Rain (p=0.4)	Shine (1-p=0.6)
Take umbrella	60	-10
Leave umbrella	-100	50

Expected utilities:

- E(Take umbrella) = $60 \times 0.4 + (-10) \times 0.6 = 18$
- E(Leave umbrella) = $-100 \times 0.4 + 50 \times 0.6 = -10$

FIT5047平时班 – Module 4

- X and Y are **independent** if

$$\forall x, y \; \Pr(x, y) = \Pr(x)\Pr(y)$$

$$\dashrightarrow \quad X \perp\!\!\!\perp Y$$

- X and Y are **conditionally independent** given Z

$$\forall x, y, z \; \Pr(x, y|z) = \Pr(x|z)\Pr(y|z)$$

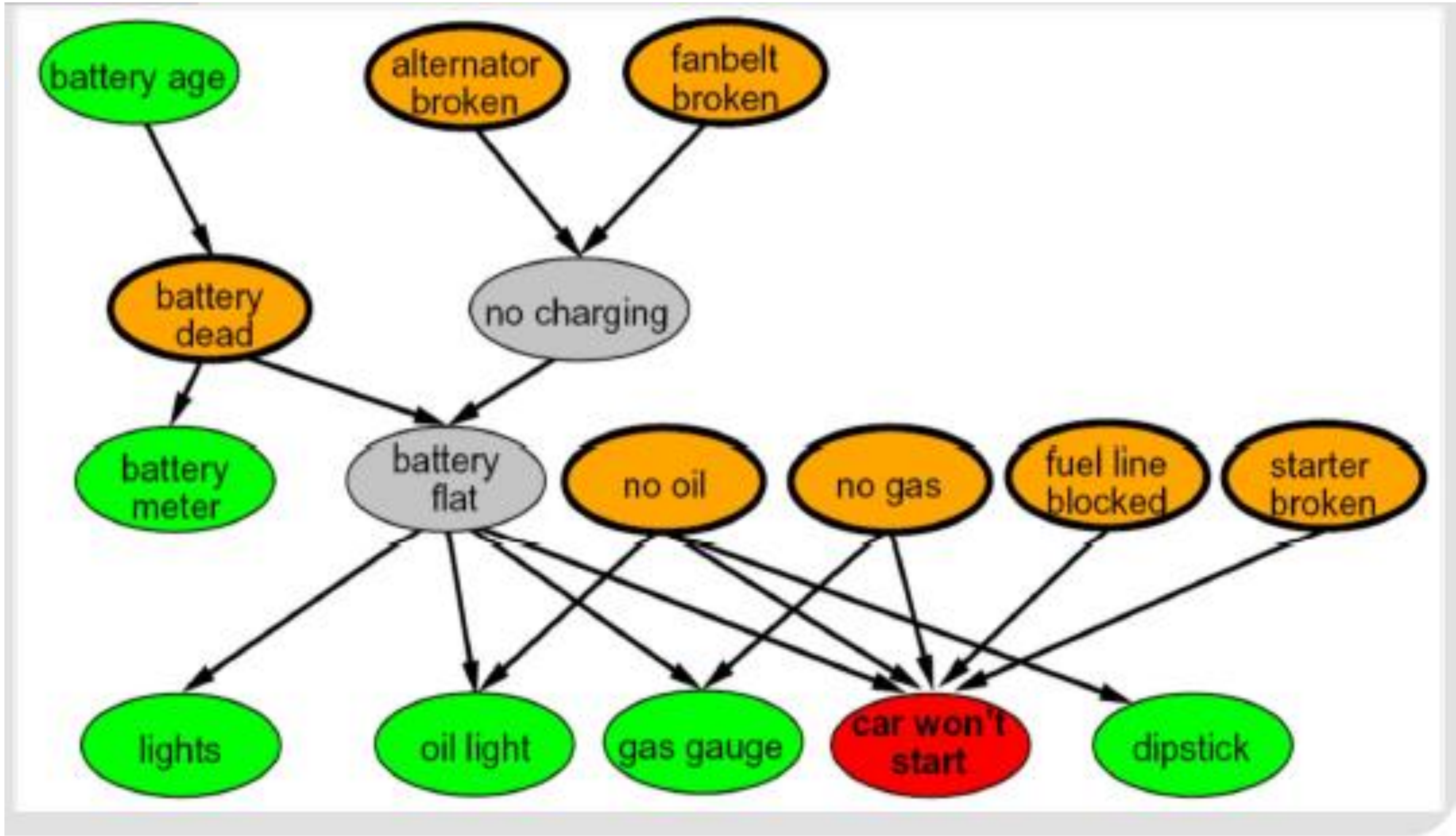
$$\dashrightarrow \quad X \perp\!\!\!\perp Y | Z$$

- (Conditional) independence is a property of a distribution

FIT5047平时班 – Module 4

- Bayes nets (aka graphical models): a technique for describing complex joint distributions (models) using simple, local distributions (conditional probabilities)
  - Describe how variables interact locally
    - > Local interactions chain together to give global, indirect interactions

FIT5047平时班 – Module 4



FIT5047平时班 – Module 4

Nodes: variables (with domains)

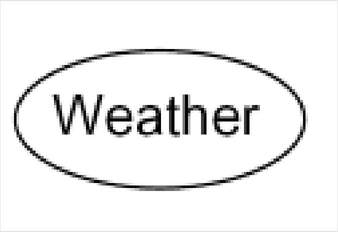
– Can be assigned (observed) or unassigned (unobserved)

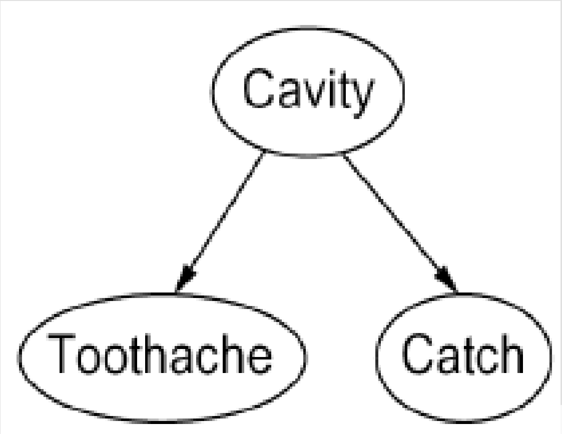
Arcs: interactions

– Indicate "direct influence" between variables

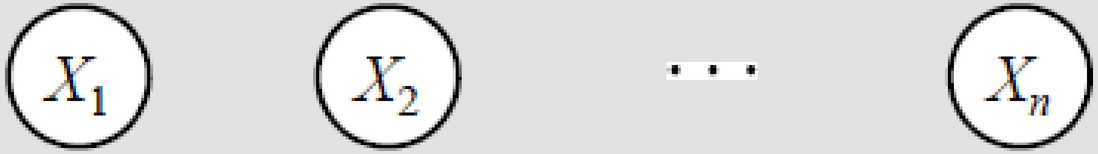
– Formally: encode conditional independence

For now, imagine that arrows mean direct causation





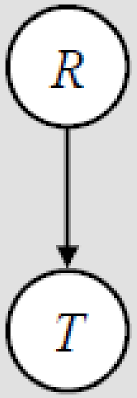
N independent coin flips



No interactions between variables: absolute independence

FIT5047平时班 – Module 4

- **Variables:**
  - R: It rains
  - T: There is traffic
- **Model 1: independence**
- **Model 2: rain causes traffic**
- **Why is model 2 better?**

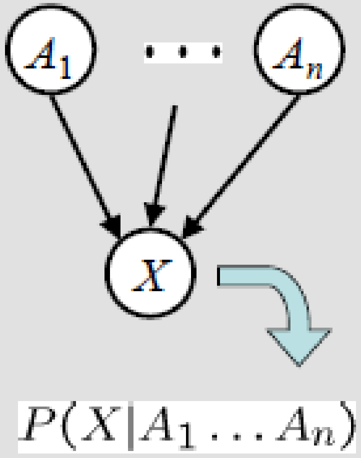


Bayesian Networks – Definition (II)

- **The probability distribution for each node X is a collection of distributions over X, one for each combination of its parents' values**

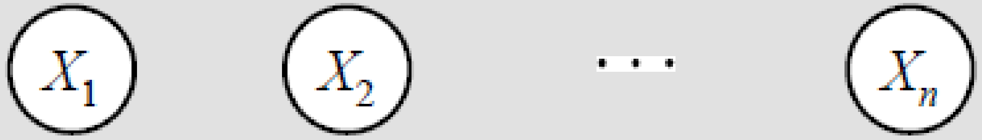
$$\mathbf{Pr}(X|a_1, \dots, a_n)$$

  - described by a ***Conditional Probability Table (CPT)***
  - describes a "noisy" causal process



Bayesian network = Topology (graph) + Local Conditional Probabilities

FIT5047平时班 – Module 4



$\Pr(X_1)$

h	0.5
t	0.5

$\Pr(X_2)$

h	0.5
t	0.5

$\cdots$

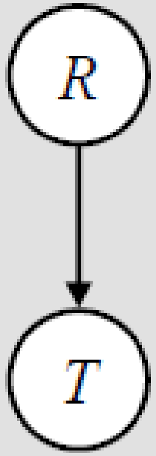
$\Pr(X_n)$

h	0.5
t	0.5

$$\Pr(h, t, t, h) = 0.5 \times 0.5 \times 0.5 \times 0.5$$

Only distributions whose variables are independent can be represented by a Bayes net with no arcs

FIT5047平时班 – Module 4



$\Pr(R)$

+r	1/4
¬r	3/4

$\Pr(T|R)$

+r	+t	3/4
+r	¬t	1/4
¬r	+t	1/2
¬r	¬t	1/2

$=$ $\Pr(T|R)$

	+t	¬t
+r	3/4	1/4
¬r	1/2	1/2

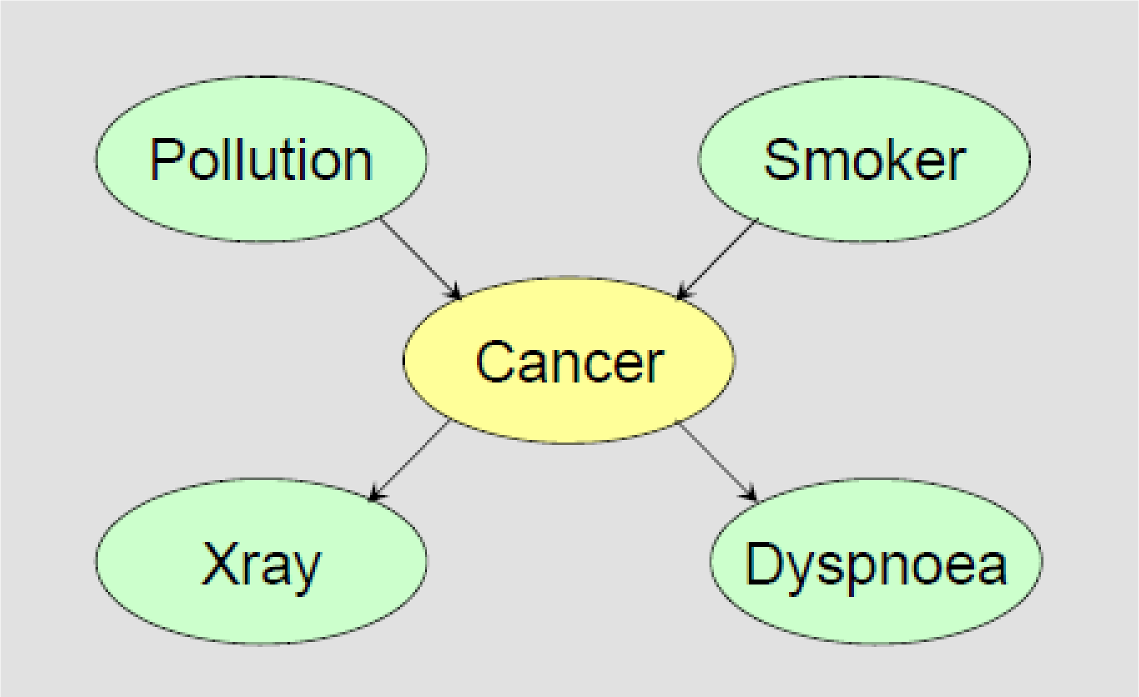
$$\Pr(+r, \neg t) = ?$$

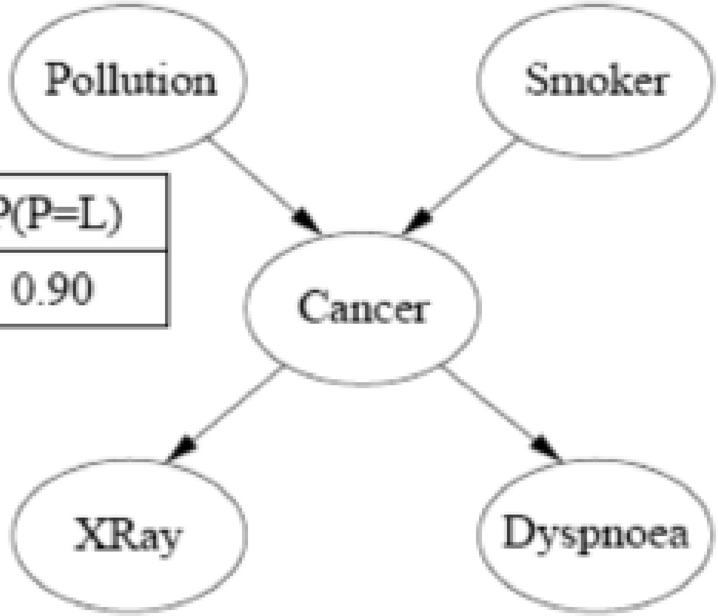
$$\Pr(+r, \neg t) = \Pr(\neg t|+r)\Pr(+r) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

FIT5047平时班 – Module 4

A patient has been suffering from shortness of breath (called dyspnoea) and visits the doctor, worried that he has lung cancer.

The doctor knows that other relevant information includes whether or not the patient is a smoker (increasing the chances of cancer) and what sort of air pollution he has been exposed to. A positive Xray would indicate lung cancer.





P(P=L)
0.90

P(S=T)
0.30

P	S	P(C=T|P,S)
H	T	0.05
H	F	0.02
L	T	0.03
L	F	0.001

C	P(X=pos|C)
T	0.90
F	0.20

C	P(D=T|C)
T	0.65
F	0.30

The value for =F is the complement

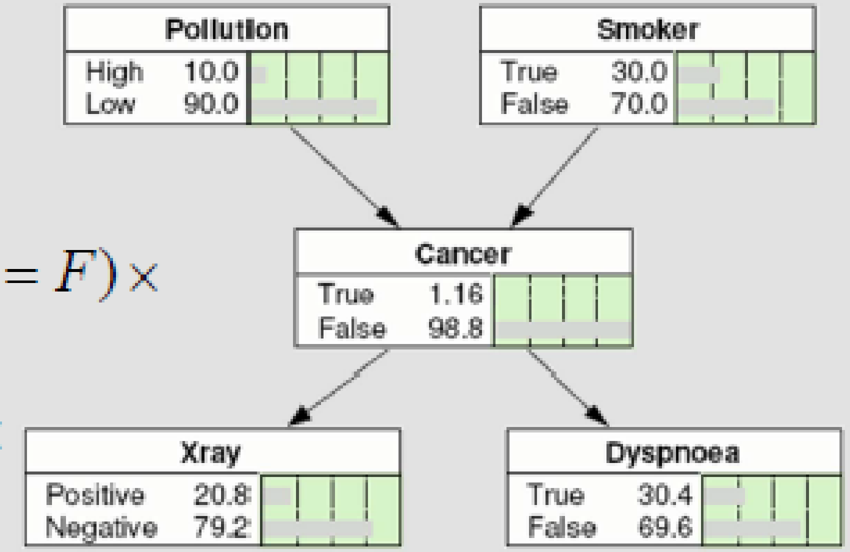
FIT5047平时班 – Module 4

$$\Pr(P = low \wedge S = F \wedge C = T \wedge X = pos \wedge D = T) =$$

$$\Pr(P = low) \times$$
$$\Pr(S = F) \times$$

$$\Pr(C = T \mid P = low, S = F) \times$$

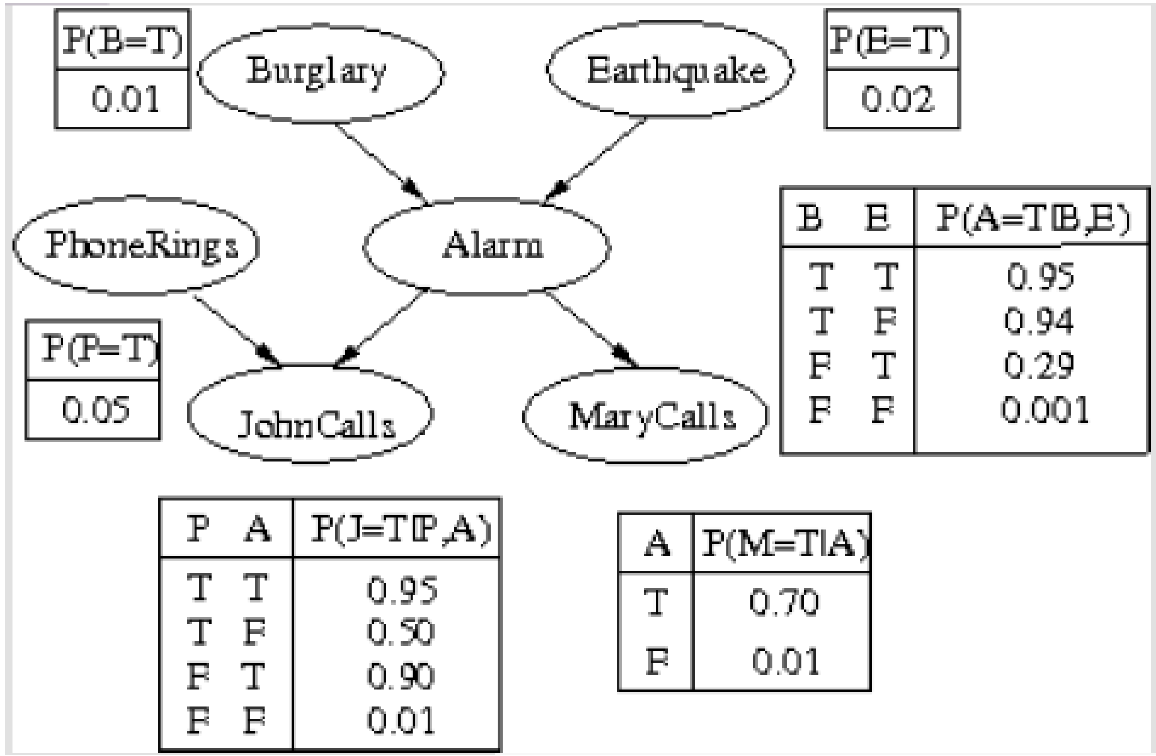
$$\Pr(X = pos \mid C = T) \times$$
$$\Pr(D = T \mid C = T)$$



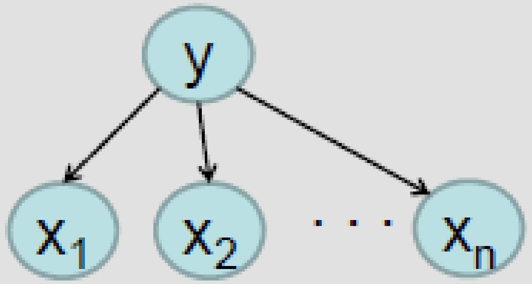
FIT5047平时班 – Module 4

You have a new burglar alarm installed. It reliably detects burglary, but also responds to minor earthquakes. Two neighbours, John and Mary, promise to call the police when they hear the alarm. John always calls when he hears the alarm, but sometimes confuses the alarm with the phone ringing and calls then also. On the other hand, Mary likes loud music and sometimes doesn't hear the alarm.

Given evidence about who has and hasn't called, you'd like to estimate the probability of a burglary.



Imagine we have one cause *y* and several effects *x*:



$$\Pr(y, x_1, x_2, \dots, x_n) = \Pr(y)\Pr(x_1|y)\Pr(x_2|y) \dots \Pr(x_n|y)$$

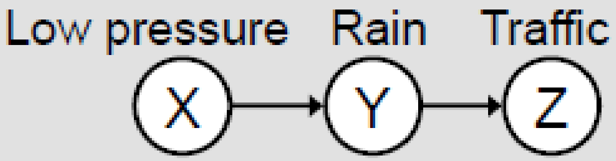
This is a *naïve Bayes* model

FIT5047平时班 – Module 4

- The relationship between conditional independence and BN structure is important for understanding how BNs work
- Factors that affect conditional independence
  - + Causal chains
  - + Common causes
  - – Common effects

Causal Chains

- A causal chain of events



$$\Pr(x, y, z) = \Pr(x)\Pr(y|x)\Pr(z|y)$$

- Is Z independent of X given Y? *Yes!*

$$\Pr(z|x, y) = \frac{\Pr(x, y, z)}{\Pr(x, y)} = \frac{\cancel{\Pr(x)}\cancel{\Pr(y|x)}\Pr(z|y)}{\cancel{\Pr(x)}\cancel{\Pr(y|x)}}$$

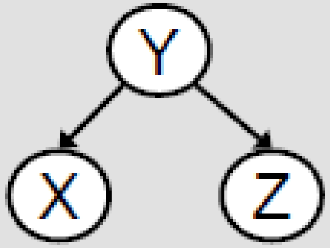
$$= \Pr(z|y)$$

- Evidence along the chain "blocks" the influence

Common Cause

- **Two effects of the same cause**
  - Are X and Z independent? ***No***
  - Are X and Z independent given Y? ***Yes!***

Y: Project due
X: Newsgroup busy
Z: Lab full



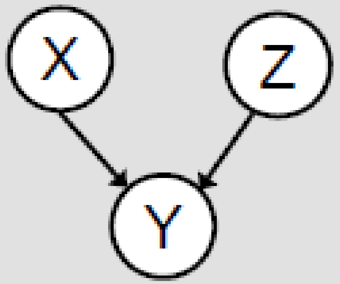
$$\Pr(z|x,y) = \frac{\Pr(x,y,z)}{\Pr(x,y)} = \frac{\Pr(\cancel{y})\Pr(\cancel{x|}y)\Pr(z|y)}{\Pr(\cancel{y})\Pr(\cancel{x|y})} = \Pr(z|y)$$

- **Observing the cause blocks influence between the effects**

Common Effect

- Two causes of one effect (v-structures)
- Are X and Z independent?
  - Yes: the ballgame and the rain cause traffic, but they are not correlated
- Are X and Z independent given Y?
  - No: seeing traffic puts the rain and the ballgame in competition as explanation
- This is different from the other cases
  - Observing an effect activates the influence between possible causes

X: Rain
Z: Ballgame
Y: Traffic

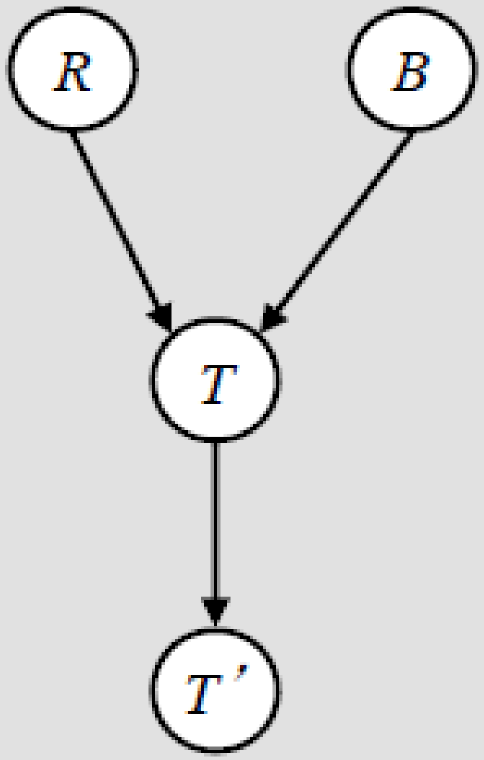


FIT5047平时班 – Module 4

$R \perp\!\!\!\perp B$

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

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



FIT5047平时班 – Module 4

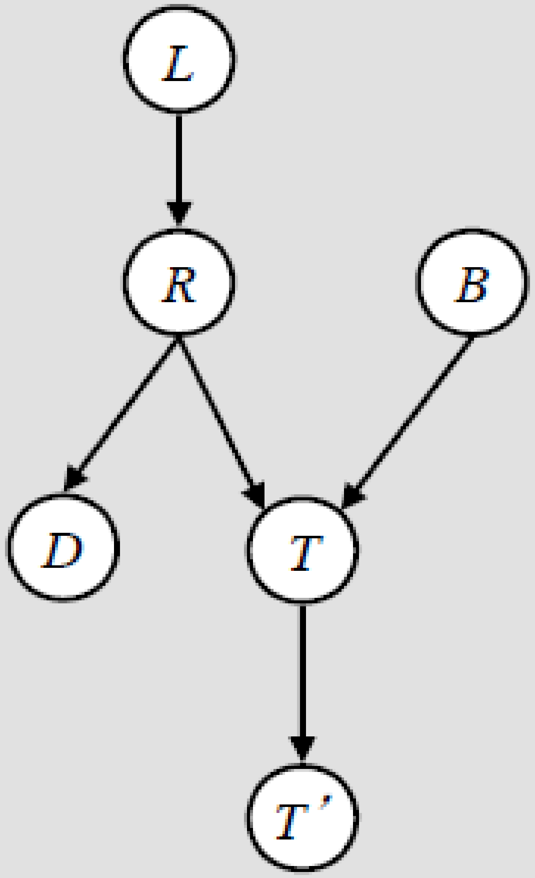
$L \perp\!\!\!\perp T' | T$

$L \perp\!\!\!\perp B$

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

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

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



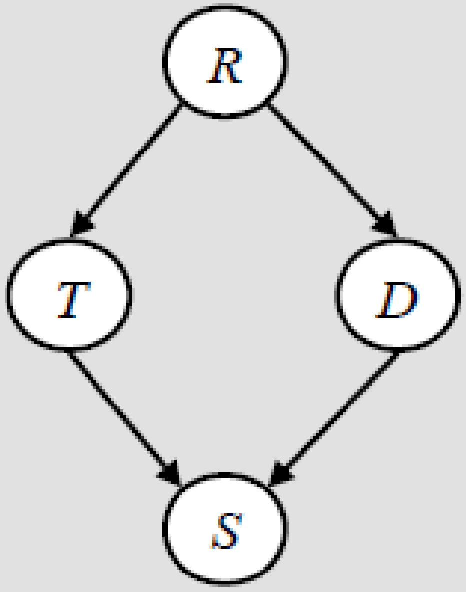
FIT5047平时班 – Module 4

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

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

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

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



Decision Networks

- Extension of BNs to support making decisions
- Utility theory represents preferences between different outcomes of various plans
- Decision theory =
  Utility theory + Probability theory

Types of Nodes

- ***Chance nodes*** **– (ovals) random variables**
  - Have an associated CPT
  - Parents can be decision nodes and other chance nodes
- ***Decision nodes*** **– (rectangles) points where the decision maker has a choice of actions**
  - The table is the decision with the highest computed EU for each combination of evidence in the ***information link*** parents
- ***Utility nodes (Value nodes)*** **– (diamonds) the agent's utility function**
  - The table represents a multi-attribute utility function
  - Parents are variables describing the outcome states that directly affect utility

FIT5047平时班 – Module 4

- *Informational Links* – indicate when a chance node needs to be observed before a decision is made
  - Any link entering a decision node is an informational link
- *Conditioning links* – indicate the variables on which the probability assignment to a chance node will be conditioned

Suppose that you know that a fever can be caused by the flu. You can use a thermometer, which is fairly reliable, to test whether or not you have a fever.

Suppose you also know that if you take aspirin it will almost certainly lower a fever to normal. Some people (about 5% of the population) have a negative reaction to aspirin. You'll be happy to get rid of your fever, so long as you don't suffer an adverse reaction if you take aspirin.