

# 第五章 不确定性推理

- 概述
- 概率论基础
- 主观Bayes方法
- 确定性方法（可信度方法）
- 证据理论
- Bayes网络

# 贝叶斯网络

- 80' s Bayes网络成功地应用于专家系统，成为表示不确定性专家知识和推理的一种流行的方法。
- 数学理论基础坚实。综合了先验信息（领域知识）和样本信息，可避免只用先验信息可能带来的主观偏见。
- 虽然很多Bayes网络问题是NP难的。但可以利用近似法求解。
- Bayes网络方法的不确定性表示和计算保持了概率的方式，只是实现时，根据应用背景采用各种近似计算方法。推理过程称为概率推理。因此，Bayes网络没有其它不确定性推理方法拥有的不确定性表示、计算、语义解释等问题。
- 只介绍Bayes网络的基本概念和简单的推理方法。

# 贝叶斯网络（事件的独立性）

- 独立：如果X与Y相互独立，则

$$P(X, Y) = P(X)P(Y)$$

$$P(X|Y) = P(X)$$

- 条件独立：如果在给定Z的条件下，X与Y相互独立，则

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

实际中，条件独立比完全独立更重要

# 贝叶斯网络（因果关系网络）

命题S(smoker)：患者是一个吸烟者

命题C(coal Miner)：患者是一个煤矿矿井工人

命题L(lung Cancer)：他患了肺癌

命题E(emphysema)：他患了肺气肿

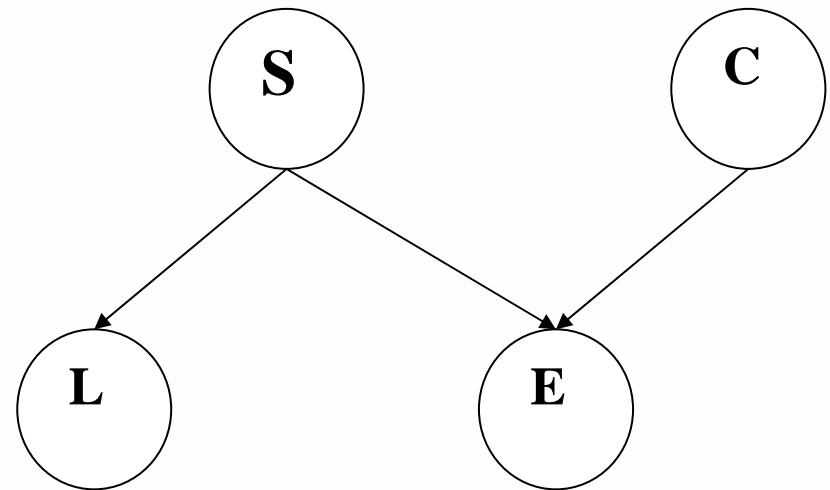
- S对L和E有因果影响，而C对E也有因果影响。

# 贝叶斯网络（因果关系图例）

命题间的关系可以描绘成因果关系网。节点代表命题，弧代表节点间的直接因果关系。

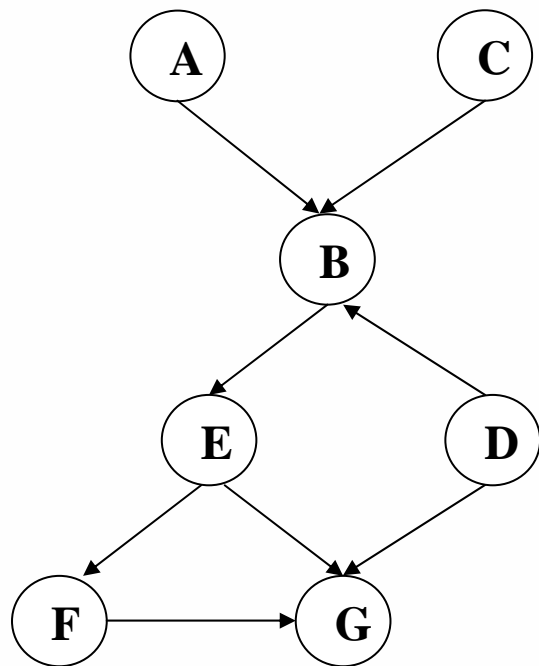
节点S，C是节点L和E的父节点或称双亲节点；

L，E也称为是S和C的子节点或称后代节点。



因果关系图例

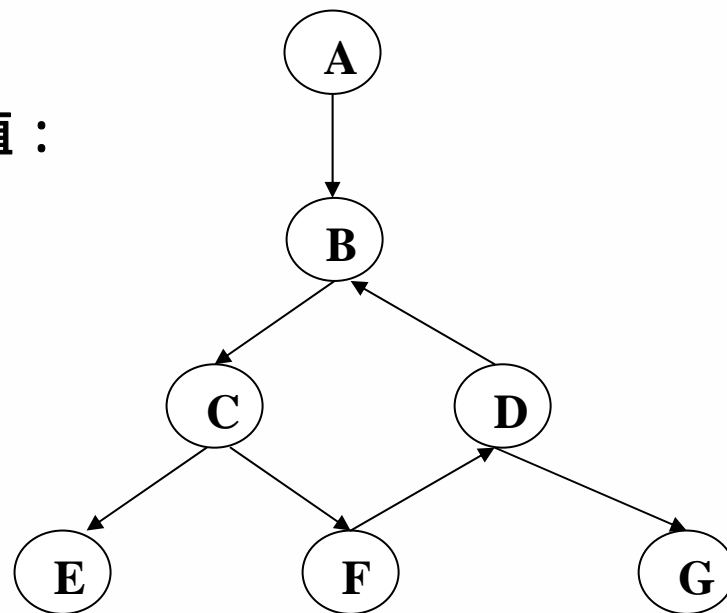
# 贝叶斯网络（图例）



无环图和指定概率值：

$P(A)$ ,  
 $P(B)$ ,  
 $P(B|AC)$ ,  
 $P(E|C)$ ,  
 $P(D|C)$ ,  
 $P(F|E)$ ,  
 $P(G|DEF)$

贝叶斯网络图例



非贝叶斯网络图例

# 贝叶斯网络（定义）

- 贝叶斯网就是一个在弧的连接关系上加入连接强度的因果关系网络。
- 两个部分
  - 贝叶斯网络结构图是一个有向无环图（DAG: Directed Acyclic Graph）。
  - 节点之间关系的条件概率表（Conditional Probability Table, CPT）
- 目的：由证据得出结论发生的概率。 即观察到  $P(Y)$ ，求  $P(X|Y)$

# 贝叶斯网络（如何构造）

- 确定建模的目标
- 选择变量，生成节点（ mutually exclusive and collectively exhaustive states ）
- 对节点排序
- 填充网络连接弧，表示节点之间的关系
- 得到条件概率关系表

条件概率表示的概率网络有时叫“Belief Nets”  
（信念网）

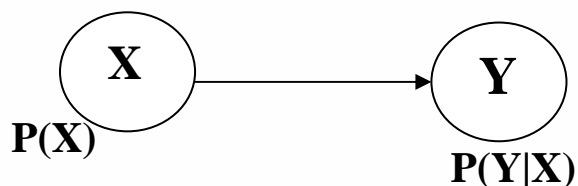


## 贝叶斯网络（计算）

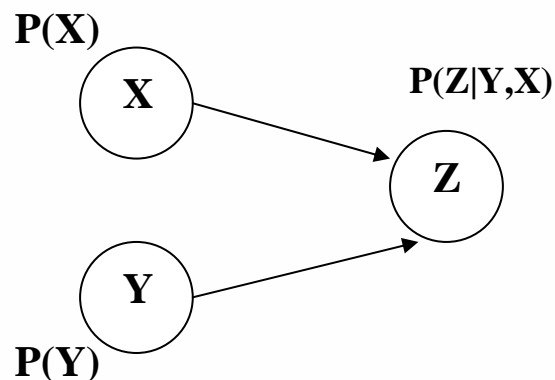
- 有向无环图是各个节点变量关系传递的合理表达形式。
- 条件独立的引入使得计算较之全连接网络有了大大的简化。
- CPT表相对比较容易得到。

## 贝叶斯网络（计算续）

- 简单的联合概率可以直接从网络关系上得到



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

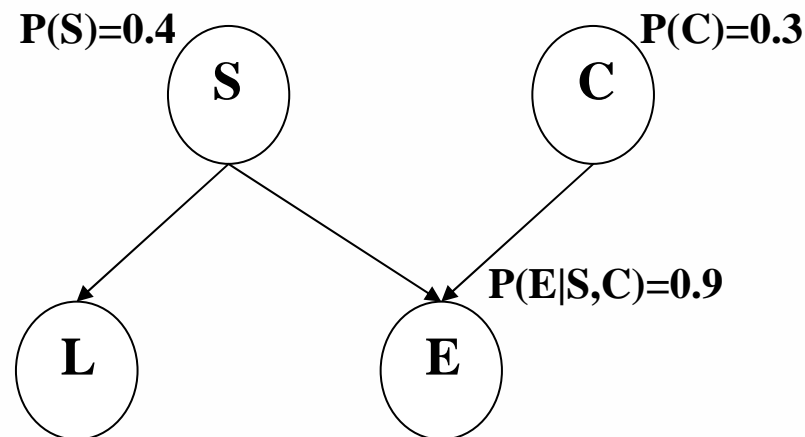


$$P(X, Y, Z) = P(X)P(Y)P(Z|X, Y)$$

# 贝叶斯网络（例）

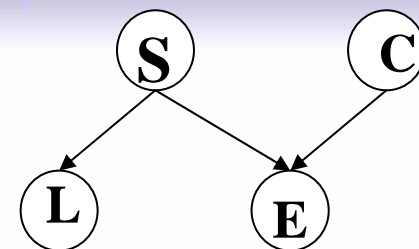
CPT表为：

- $P(S)=0.4$
- $P(C)=0.3$
- $P(E|S, C)=0.9$
- $P(E|S, \neg C)=0.3$
- $P(E|\neg S, C)=0.5$
- $P(E|\neg S, \neg C)=0.1$



贝叶斯网络实例图

# 贝叶斯网络（例续）



- 联合概率密度：

$$P(S, C, L, E) = P(E | S, C, L) * P(L | S, C) * P(C | S) * P(S)$$

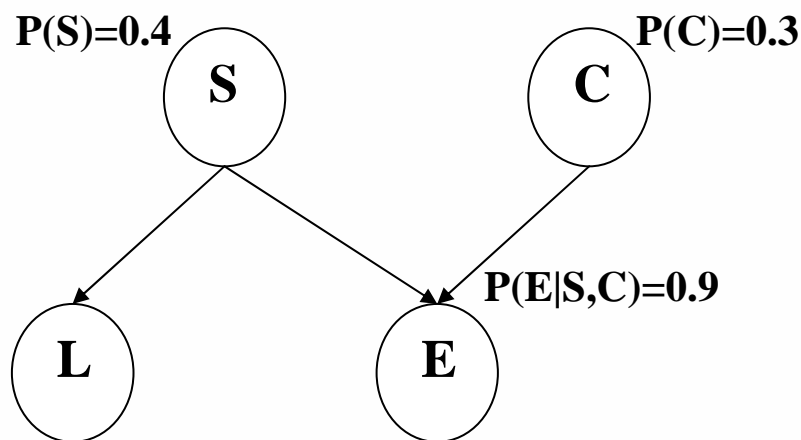
- E与L在S条件下独立， $P(E | S, C, L) = P(E | S, C)$ ,
- L与C在S，E条件下独立， $P(L | S, C) = P(L | S)$
- C与S在E条件下独立， $P(C | S) = P(C)$
- 以上三条等式可从Bayes网的条件独立性和后面的D分离定义中得到。
- 简化后的联合概率密度为，

$$P(S, C, L, E) = P(E | S, C) * P(L | S) * P(C) * P(S)$$

若原Bayes网中条件独立语义较多，计算量减少更明显。

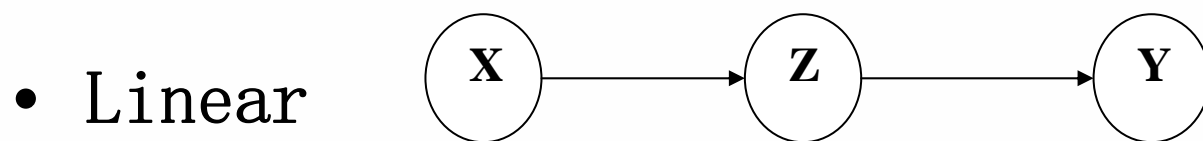
- 如何判定个节点之间的条件独立性？
- D分离！！**

# 贝叶斯网络（D分离定义思路）



- 考虑S, L, E
- L（结果）影响S（起因），S影响E（另一个结果）。
- 如果给定原因S后，L并不能告诉我们有关E的更多事情。即对于S，L和E是相对独立的，减少计算复杂度。
- 称S能D分离L和E。（定义D分离的思路）
- D分离是一种寻找条件独立的有效方法。

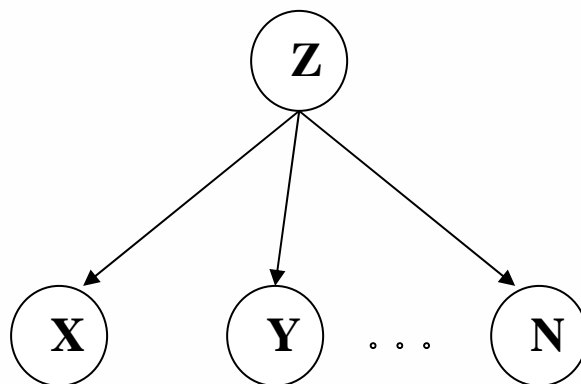
# 贝叶斯网络（D分离之一-串行）



- 事件X通过Z影响Y，反之事件Y也是通过Z影响X。
- 但若原因证据Z给定，X并不能给Y更多的东西。通道被Z阻塞，X和Y就是独立的了。则称X和Y是被Z节点D分离的。

# 贝叶斯网络（D分离之二-分叉连接）

- Diverging

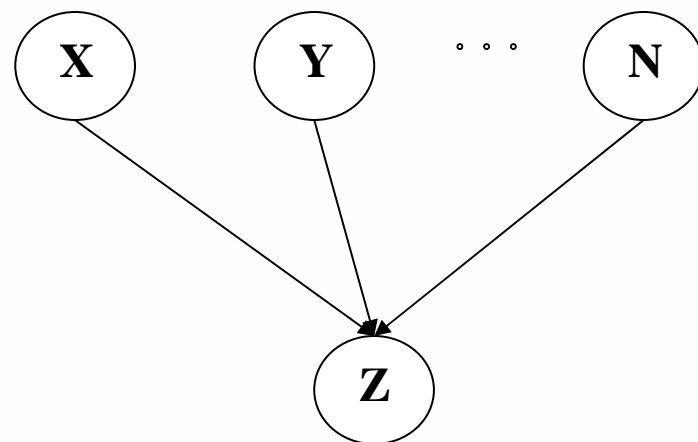


- 子节点X, ..., N通过父节点Z可以相互影响
- 如果父节点Z是已知的，没有更多的信息能够通过Z影响到其它子节点。
- 父节点Z是已知时，子节点X, ..., N是相互独立的。称子节点X, ..., N是被Z节点D分离的。

# 贝叶斯网络（D分离之三(汇集连接)）

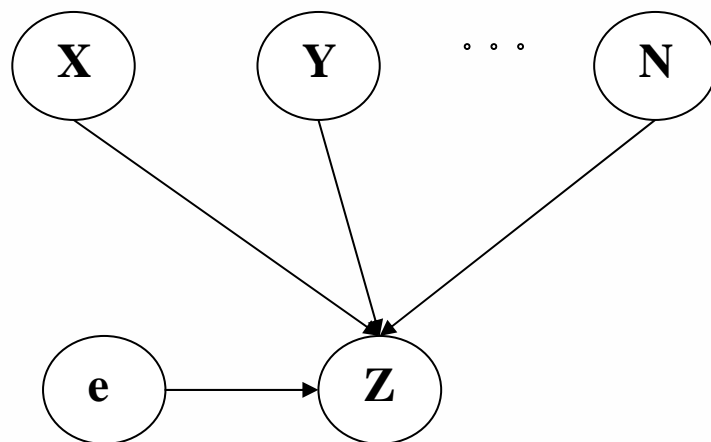
汇集(Converging)（独立、不独立之转化）

- 如果不从父节点得到推断，子节点Z就一无所知，父节点是相互独立的。
- 若某事件影响了Z，则各个父节点就不是相互独立的了。该事件可直接影响Z，也可通过它的后代节点影响Z。这种现象称作**条件依存**。
- 总之，如果子节点有了变化，或子节点的后代节点发生变化，信息可通过汇集连接传播，此时X, Y, .....N不独立。

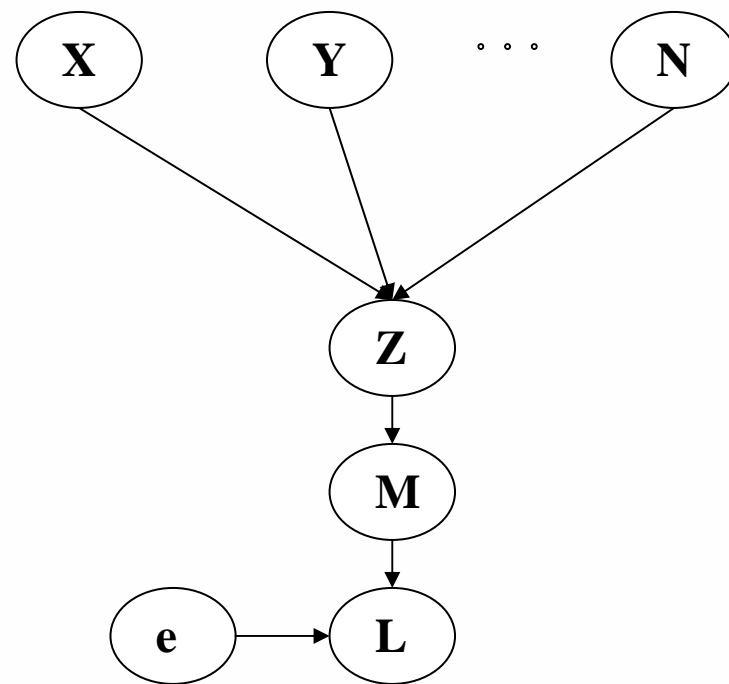




# 贝叶斯网络（D分离(条件依存)）



事件e直接影响节点Z



事件e影响节点Z的后代

# 贝叶斯网络（D分离(定义)）

- 给定结点集  $\varepsilon$ ，若对Bayes网中的结点  $V_i$  和  $V_j$  之间的每个无向路径（即不考虑弧的方向），都有某个结点  $V_b$  且有属性：
    - $V_b$  在  $\varepsilon$  中，且路径上的两条弧都以  $V_b$  为尾（ $V_b$  是分叉连接的父节点）
    - $V_b$  在  $\varepsilon$  中，路径上的一条弧以  $V_b$  为头，一条以  $V_b$  为尾（ $V_b$  为串行连接的中节点）
    - $V_b$  和它的任何后继都不在  $\varepsilon$  中，路径上的两条弧都以  $V_b$  为头(箭头)（ $V_b$  为汇集连接的子节点，但没有后代节点）
- 则称  $V_i$  和  $V_j$  被  $V_b$  结点阻塞。

# 贝叶斯网络（D分离(定义)）

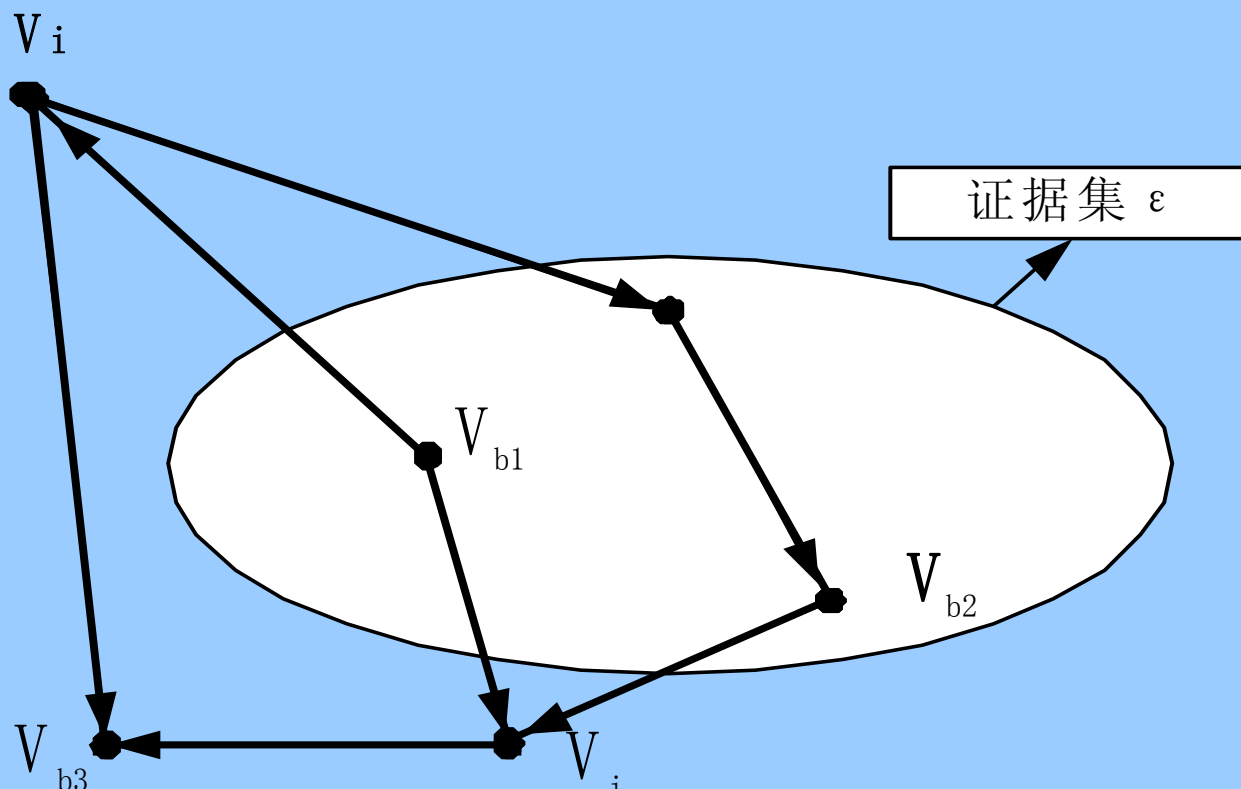
- 如果 $V_i$ 和 $V_j$ 被证据集合  $\varepsilon$  中的任意结点阻塞，则称 $V_i$ 和 $V_j$ 是被  $\varepsilon$  集合D分离，结点 $V_i$ 和 $V_j$ 条件独立于给定的证据集合  $\varepsilon$ ，记为：

$$P(V_i | V_j, \varepsilon) = P(V_i | \varepsilon) \quad P(V_j | V_i, \varepsilon) = P(V_j | \varepsilon)$$

或

$$I(V_i, V_j | \varepsilon) \quad I(V_j, V_i | \varepsilon)$$

## 贝叶斯网络 (D分离(图示))



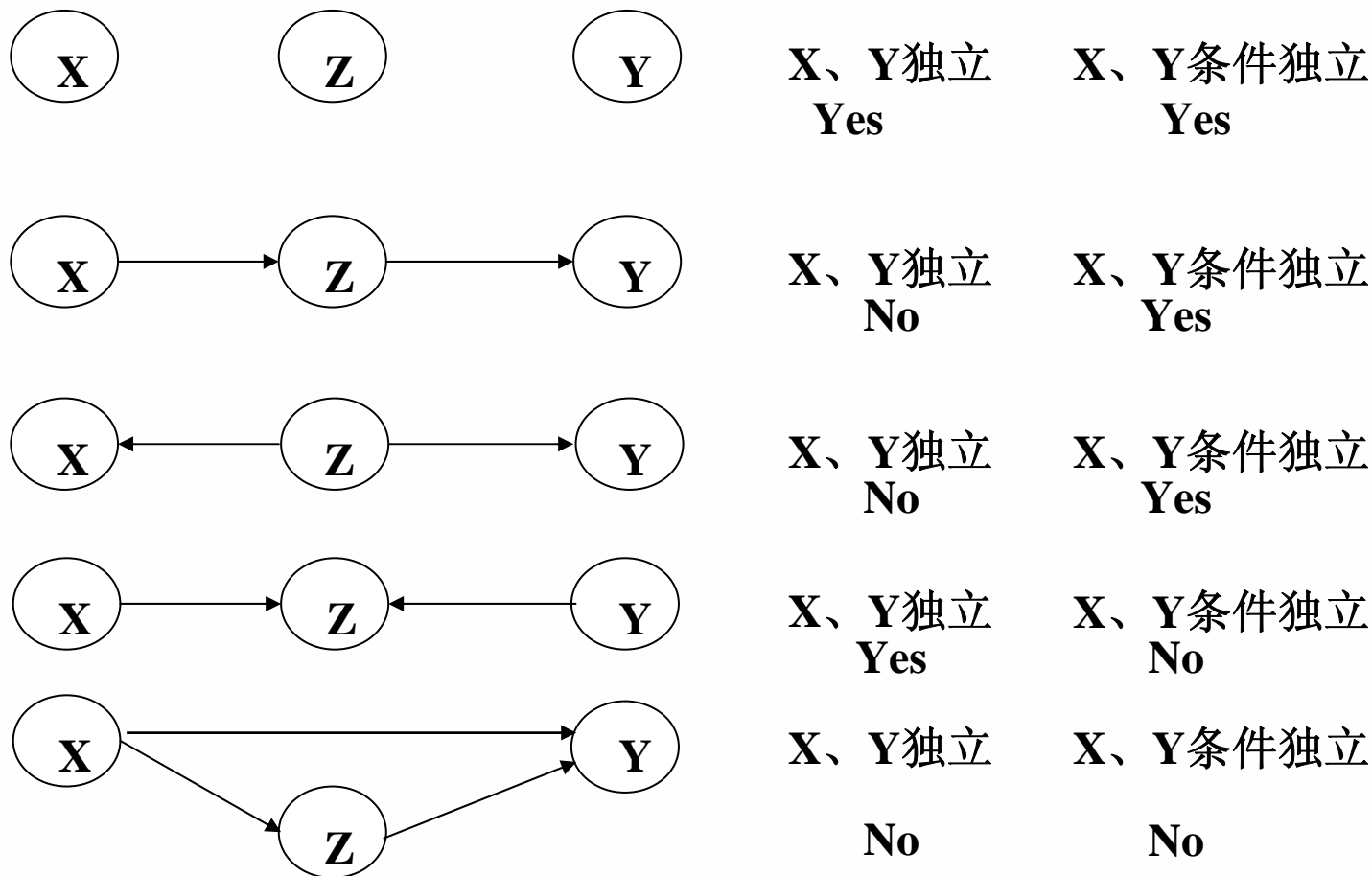
给定证据结点集  $\epsilon$  ,  $V_i$  独立  $V_j$ :  $V_i$  到  $V_j$  的所有三条路径都被阻塞

- a)  $V_{b1}$  是证据结点, 两条弧都以  $V_{b1}$  为尾
- b)  $V_{b2}$  是证据结点, 一条以  $V_{b2}$  为头, 一条以  $V_{b2}$  为尾
- c)  $V_{b3}$  及其任何后继都不在证据结点集中, 两条弧都以  $V_{b3}$  为头

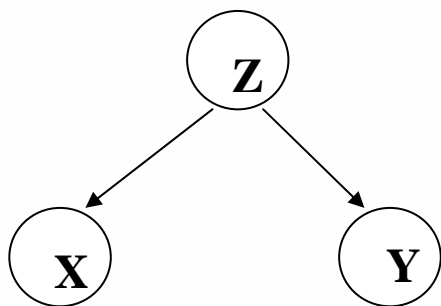
# 贝叶斯网络（定义）

- 条件独立：
  - 如具有以上三个属性之一，就说结点 $V_i$ 和 $V_j$ 条件独立于给定的结点集 $\varepsilon$ 。
- 阻塞：
  - 给定证据集合 $\varepsilon$ ，当上述条件中的任何一个满足时，就说 $V_b$ 阻塞相应的那条路径。
- D分离：
  - 如果 $V_i$ 和 $V_j$ 之间所有的路径被阻塞，就叫证据集合 $\varepsilon$ 可以D分离 $V_i$ 和 $V_j$

# 贝叶斯网络 (D分离(例1))

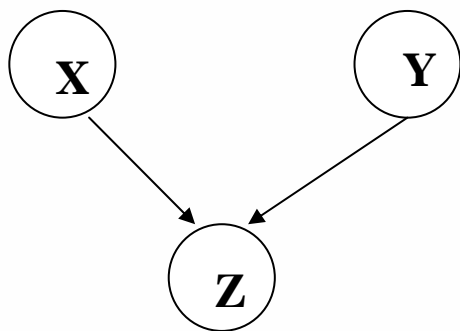


# 贝叶斯网络（D分离(例2)）



X—草湿  
Y—彩虹  
Z—下雨

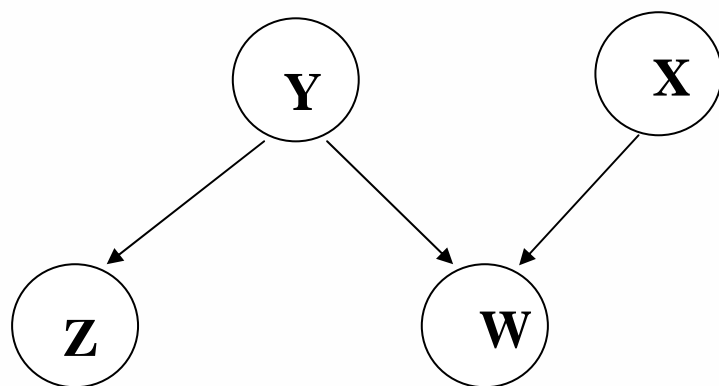
$$P(X,Y) \neq P(X)P(Y)$$
$$P(X|Y,Z) = P(X|Z)$$



X—下雨  
Y—洒水  
Z—草湿

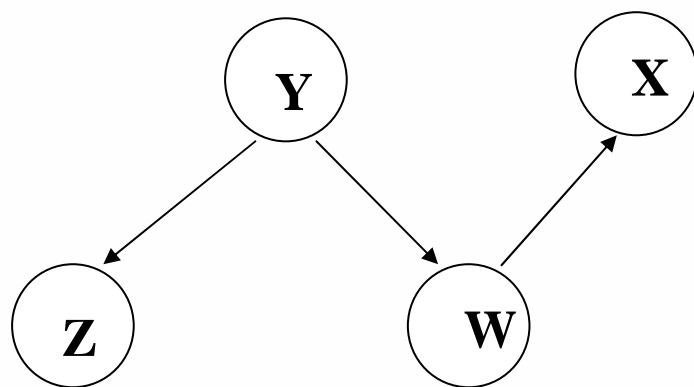
$$P(X,Y) = P(X)P(Y)$$
$$P(X|Y,Z) \neq P(X|Z)$$

# 贝叶斯网络（D分离(例3)）



X—草湿  
Y—洒水  
Z—彩虹  
W—长虫

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



X—草湿  
Y—洒水  
Z—彩虹  
W—长虫

$$P(X, Y) \neq P(X)P(Y)$$
$$P(X|Y, Z) \neq P(X|Z)$$



# 贝叶斯网络（D分离(例4)）

Radio and Ignition, given Battery?

Yes

Radio and Start, given Ignition?

Yes

Gas and Radio, given Battery?

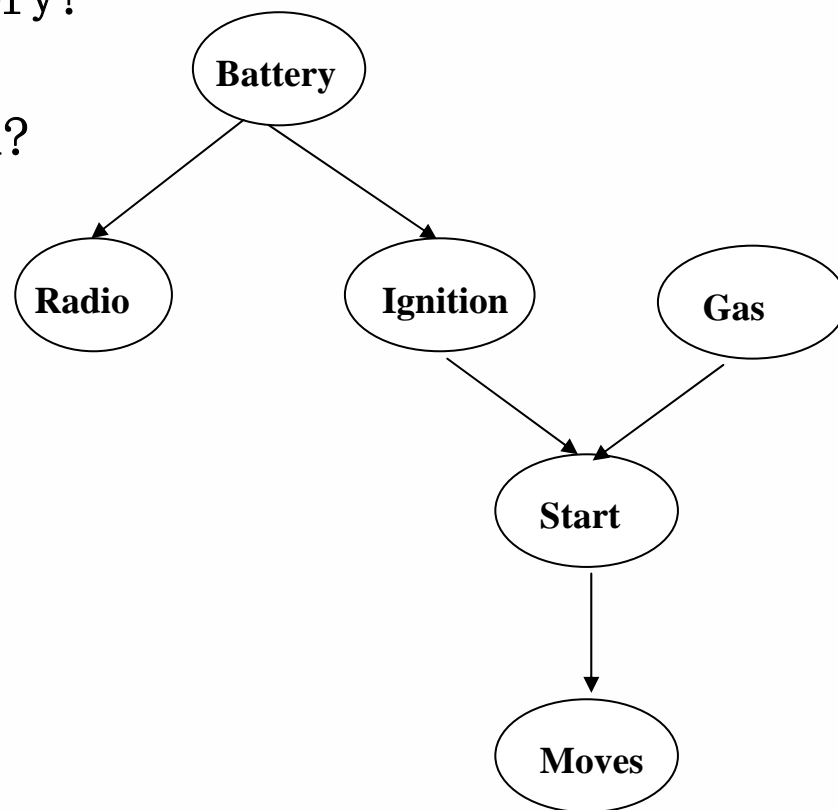
Yes

Gas and Radio, given Start?

No

Gas and Battery, given Moves?

N



# 贝叶斯网络（推理）

- 建立贝叶斯网络的目的
  - 有了网络。可以提出问题：
    - $P(\text{问题}|\text{证据})$ ， 如：  $P(\text{吸烟}|\text{肺癌})$
  - 进行概率推理
  - 与谓词逻辑有相似之处 。
- 在某些场合下有有效的推理方法。有一些工具包。
- 一般情况下是很困难的， 原因
  - 不是所有的CPT表都能够得到
  - 网络结构大且复杂
  - NP-hard推理
- 我们要做的是：将问题正确的表示为合理的网络形式，选用适合的算法。

## 贝叶斯网络（推理续）

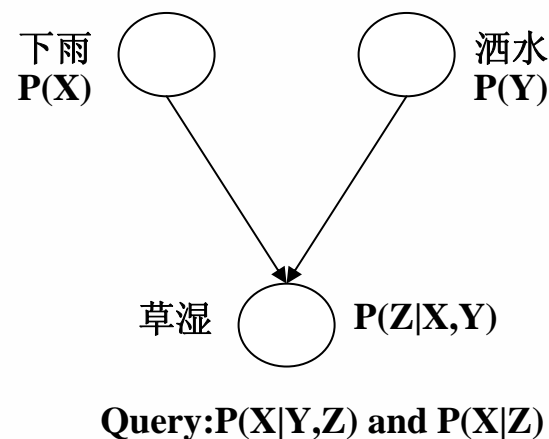
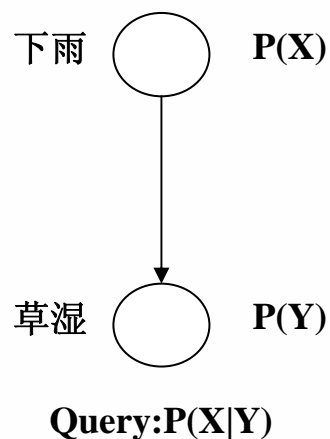
- 贝叶斯网络通常使用因果或诊断规则与推理
  - 因果规则：X Cause Y with some probability
  - 诊断规则：Y is evidence of X with some probability
  - 因果推理：Given cause C, determine  $P(\text{Query} | C)$
  - 诊断推理：Given evidence E, determine  $P(\text{Query} | E)$

# 贝叶斯网络（推理续）

- 推理需求：  $P(X|Y)$ 
  - 诊断推理是从效果到起因  
证据是一些征兆：  $X$ 是起因，  $Y$ 是征兆
  - 因果推理是从起因到效果  
证据是一些起因：  $X$ 是征兆，  $Y$ 是起因
  - 解释历史  
 $X$ 和 $Y$ 是起因，  $Z$ 是两个起因的征兆。这时可以用一个起因 $Y$ 解释另一个起因 $X$ 。

# 贝叶斯网络（推理例）

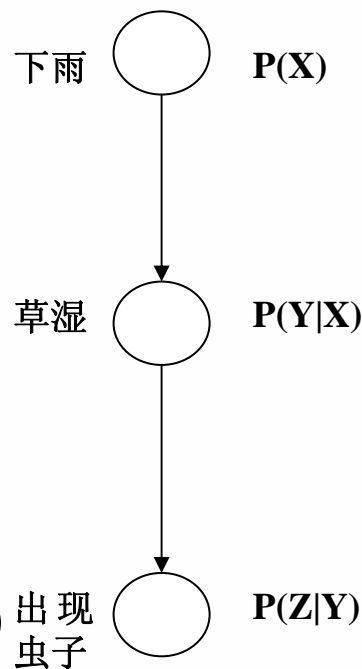
- 下雨、草湿、洒水



# 贝叶斯网络（推理例续）

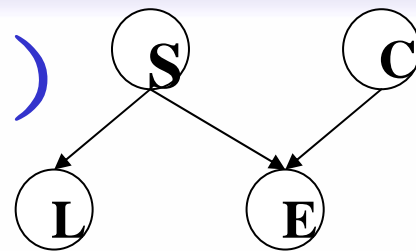
- 条件：
  - 下雨
  - 草湿
  - 出现虫子

- 求：
  - $P(\text{Raining} | \text{Worm Sighting})$



Query:  $P(X|Z)$

# 贝叶斯网络（因果推理例）



计算概率 $P(E|S)$ 。 $S$ 称作推理的证据， $E$ 叫询问结点。

首先， $E$ 的另一个父结点（ $C$ ），

$$P(E|S)=P(E,C|S)+P(E,\sim C|S);$$

$$\begin{aligned} P(E,C|S) &= P(E,C,S)/P(S) = P(E|C,S)*P(C,S)/P(S) \\ &= P(E|C,S)*P(C|S) \end{aligned}$$

$$P(E,\sim C|S) = P(E|\sim C,S)*P(\sim C)。$$

$$\begin{aligned} P(E|S) &= P(E|C,S)*P(C)+P(E|\sim C,S)*P(\sim C) \\ &= 0.9*0.3+0.3*(1-0.3)=0.48 \end{aligned}$$

主要操作：

按照给定证据 $V$ 和它的所有双亲的联合概率，  
重新表达给定证据的询问结点的所求条件概率。  
直到所有的概率值可从CPT表中得到，推理完成。

CPT表为：

$$P(S)=0.04$$

$$P(C)=0.3$$

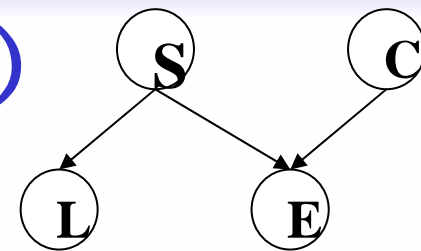
$$P(E|S,C)=0.9$$

$$P(E|S,\sim C)=0.3$$

$$P(E|\sim S,C)=0.5$$

$$P(E|\sim S,\sim C)=0.1$$

# 贝叶斯网络（诊断推理例）



$P(\sim C|\sim E)$ , 结果推断起因, 诊断推理

$P(\sim C|\sim E) = P(\sim E|\sim C) * P(\sim C) / P(\sim E)$  (因果)

$P(\sim E|\sim C) = P(\sim E, S|\sim C) + P(\sim E, \sim S|\sim C)$

$= P(\sim E|S, \sim C) * P(S) + P(\sim E|\sim S, \sim C) * P(\sim S) = 0.82$

$P(\sim C|\sim E) = 0.82 * (1 - 0.3) / P(\sim E) = 0.574 / P(\sim E)$

$P(C|\sim E) = 0.102 / P(\sim E)$

$P(\sim C|\sim E) + P(C|\sim E) = 1, \quad P(\sim E) = 0.676$

$P(\sim C|\sim E) = 0.849$

主要操作:

将诊断推理转化为因果推理

CPT表为:

$P(S) = 0.04$

$P(C) = 0.3$

$P(E|S, C) = 0.9$

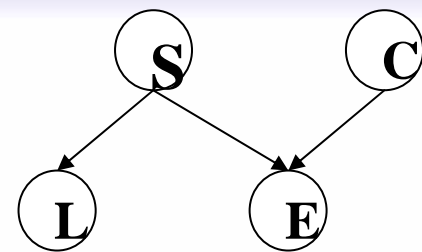
$P(E|S, \sim C) = 0.3$

$P(E|\sim S, C) = 0.5$

$P(E|\sim S, \sim C) = 0.1$



# 贝叶斯网络（辩解推理例）



$P(\sim C | \sim E, \sim S)$ , 嵌入在一个诊断推理中的因果推理  
**=0.807**