课程《人工智能原理与技术》

同济大学
TONGJI UNIVERSITY

第二章 知识表达与推理

概率推理与贝叶斯网络

同济大学
TONGJI UNIVERSITY

目录

为什么要进行不确定性推理？--刻画信息的不确定性

- 试图使用逻辑会失败，有以下三个原因:
  - 惰性: 无法列举出前提和结论的完整集合.
  - 无知: 缺乏相关事实、初始条件, etc.
    - 理论的无知: 对于该领域，缺少完整的理论.
    - 实践的无知: 即使我们知道所有的规则, 对于一个特定的对象（病人）我们也无法确定。
  - 不完备性:逻辑系统本身的表达能力不足，无法解决特定问题.

处理不确定性的方法

- **概率推理**：基于概率图进行的推理，概率推理反映了推理过程中存在的不确定性特点。

➢ 概率图模型一般分为贝叶斯网络（Bayesian Network）和马尔可夫网络（Markov Network）两大类。

- **因果推理**：基于因果图进行的推理，因果图是一种有向无环图(directed acyclic graphs, DAG)。

➢ 有向无环图指的是一个无回路的有向图，即从图中任意一个节点出发经过任意条边，均无法回到该节点。

➢ 有向无环图刻画了图中所有节点之间的依赖关系。DAG 可用于描述数据的生成机制。

➢ 这样描述变量联合分布或者数据生成机制的模型，被称为 "贝叶斯网络"（Bayesian network）

我们重点介绍基于贝叶斯网络的推理

相关知识点

同濟大學 TONGJI UNIVERSITY

- 不确定环境下，智能体的知识提供相关语句的信念度(degree of belief)，处理信念度的主要工具是概率理论(probability theory)，概率提供了一种方法以概括由我们的惰性和无知产生的不确定性

- 概率理论相关概念：随机变量、事件；概率公理；无条件概率（先验概率）、条件概率（后验概率）；完全联合概率分布；乘法法则、链式法则

相关知识点

- **随机变量**： 表示可能世界中的不确定性。随机变量以大写字母开头。
- 基本命题通过单个随机变量的赋值进行构造：
  - e.g., $Cavity = false$ (abbreviated as $\neg cavity$)
  - $Weather = sunny$ (abbreviated as $sunny$)
- 复合命题由基本命题的逻辑连接构造
  - e.g., $Weather = sunny \vee Cavity = false$

概率分布

考虑随机变量$Weather$，其定义域为 $<sunny, rainy, cloudy, snow>$，每个可能取值的概率，可以写成：

$$P(Weather = sunny) = 0.6$$

$$P(Weather = rain) = 0.1$$

$$P(Weather = cloudy) = 0.29$$

$$P(Weather = snow) = 0.01$$

相关知识点

- 先验概率 或 无条件概率

  e.g., P($Cavity$ = true) = 0.1

  P($Weather$ = sunny) = 0.72

- 联合概率分布：多个变量取值的所有组合的概率**P**($Weather, Cavity$) 是一个4×2的概率表:

$Weather$ =	sunny	rainy	cloudy	snow
$Cavity$ = true	0.144	0.02	0.016	0.02
$Cavity$ = false	0.576	0.08	0.064	0.08

  - e.g., **P**($sunny, Cavity$)

- 完全联合概率分布：
  - e.g., **P**($Toothache, Weather, Cavity$)
  - 一个完全联合分布基本满足计算任何命题的概率的需求

相关知识点

- 条件概率 或 后验概率
  - e.g., P(*cavity* | *toothache*) = 0.8
- 额外条件很重要
  - 如果获得进一步信息，*cavity*为真，条件概率将更新为：

  P(*cavity* | *toothache,cavity*) = 1
- 无关的证据，可以简化
  - e.g.,P(*cavity* | *toothache, sunny*) = P(*cavity* | *toothache*) = 0.8
- 条件概率是由无条件概率定义的：

$$P(a|b) = \frac{P(a,b)}{P(b)}$$ 要求：$P(b)>0$

●乘法法则和链式法则

- 乘法法则:

$$P(a \wedge b) = P(a \mid b)\, P(b) = P(b \mid a)\, P(a)$$

$$\text{e.g., } \mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather \mid Cavity)\, \mathbf{P}(Cavity)$$

- 考虑有n个变量的联合分布：

$$\begin{aligned} \mathbf{P}(X_1, \ldots, X_n) &= \mathbf{P}(X_1,\ldots,X_{n-1})\, \mathbf{P}(X_n \mid X_1,\ldots,X_{n-1}) && \text{(乘法法则)} \\ &= \mathbf{P}(X_1,\ldots,X_{n-2})\, \mathbf{P}(X_{n-1} \mid X_1,\ldots,X_{n-2})\, \mathbf{P}(X_n \mid X_1,\ldots,X_{n-1}) && \text{(乘法法则)} \\ &= \ldots \\ &= \Pi_{\mathrm{i}=1\sim\mathrm{n}}\, \mathbf{P}(X_i \mid X_1, \ldots, X_{i-1}) && \text{(乘法法则)} \end{aligned}$$

- 链式法则： $\mathbf{P}(X_1, \ldots, X_n) = \Pi_{\mathrm{i}=1\sim\mathrm{n}}\, \mathbf{P}(X_i \mid X_1, \ldots, X_{i-1})$

相关知识点

●使用完全联合分布进行推理

- 使用完全联合概率分布作为“知识库”，从中可以导出所有问题的答案。
- 概率推理: 根据已观察到的证据，计算查询命题的后验概率 。

- 例如，计算条件概率
  - P(A_{100} on time | no reported accidents) = 0.90
  - 表示给定证据下的信念度

- 随着新的证据的出现，概率会发生变化
  - P(A_{100} on time | no accidents, 5 a.m.) = 0.95
  - P(A_{100} on time | no accidents, 5 a.m., raining) = 0.80
  - 观察新的证据，更新信念度

●使用完全联合分布进行推理

- 一个简单的例子: 诊断牙病患者的牙痛
  - 问题域：由三个布尔变量*Toothache*，*Cavity*和*Catch*组成
  - *Catch*表示由于牙医的钢探针不洁而导致的牙龈感染
  - 给定完全联合分布，一个2×2×2的表格

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

Figure 13.3 A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- 给定完全联合分布:

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Figure 13.3 A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- 对于任意命题φ, 其概率是使得该命题成立的可能世界的概率之和：$P(\varphi) = \Sigma_{\omega:\omega \models \varphi} P(\omega)$

- 一种计算任何命题概率的方法：
  - 识别命题为真的可能世界，然后把它们的概率加起来
  - 例如：$P(cavity \lor toothache) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$

相关知识点

- 给定完全联合分布:

	toothache		*¬toothache*	
	catch	*¬catch*	*catch*	*¬catch*
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

Figure 13.3 A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- 一个特别常见的任务：提取某个变量的概率分布（无条件概率/边缘概率）
- 边缘化规则，或者称为求和消元：
    - P(*toothache*) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2

- 给定完全联合分布:

	toothache		$\neg$*toothache*	
	catch	$\neg$*catch*	*catch*	$\neg$*catch*
cavity	0.108	0.012	0.072	0.008
$\neg$*cavity*	0.016	0.064	0.144	0.576

Figure 13.3 A full joint distribution for the *Toothache*, *Cavity*, *Catch* world.

- 条件概率是由无条件概率定义的，可以计算条件概率:

$$P(\neg cavity \mid toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016+0.064}{0.108+0.012+0.016+0.064}$$

$$= 0.4$$

相关知识点

$$P(\neg cavity \mid toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016+0.064}{0.108 + 0.012 + 0.016 + 0.064}$$

$$= 0.4$$

这两个计算出来的值相加等于1。

分母是不变的，是常数。

$$P(cavity \mid toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064}$$

$$= 0.6$$

相关知识点

- 例如：$\mathbf{P}(Cavity \mid toothache)$

$= \alpha\ \mathbf{P}(Cavity, toothache)$

$= \alpha\ [\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$

$= \alpha\ [<0.108, 0.016> + <0.012, 0.064>]$

$= \alpha\ <0.12, 0.08>$

$= <0.6, 0.4>$

归一化方法

基本思想：计算查询变量的概率分布，可以固定证据变量 (evidence variables) ，然后在隐变量 (hidden variables) 上求和并归一化

归一化方法:

假设查询变量为X; 证据变量集合为 **E**，**e**表示其观察值；假设其余的未观测变量为**Y**。

计算查询变量:

$$\mathbf{P}(X \mid e) = \alpha\, \mathbf{P}(X, e) = \alpha \sum_y \mathbf{P}(X, e, y)$$

- **问题**：规模扩展性不好

对于一个由n个布尔变量所描述的问题域，最坏情况下的时间复杂性$O(2^n)$，空间复杂性 $O(2^n)$

01 概率推理知识点

02 贝叶斯规则及应用

03 贝叶斯网络推理

04 贝叶斯网络推理案例

贝叶斯规则

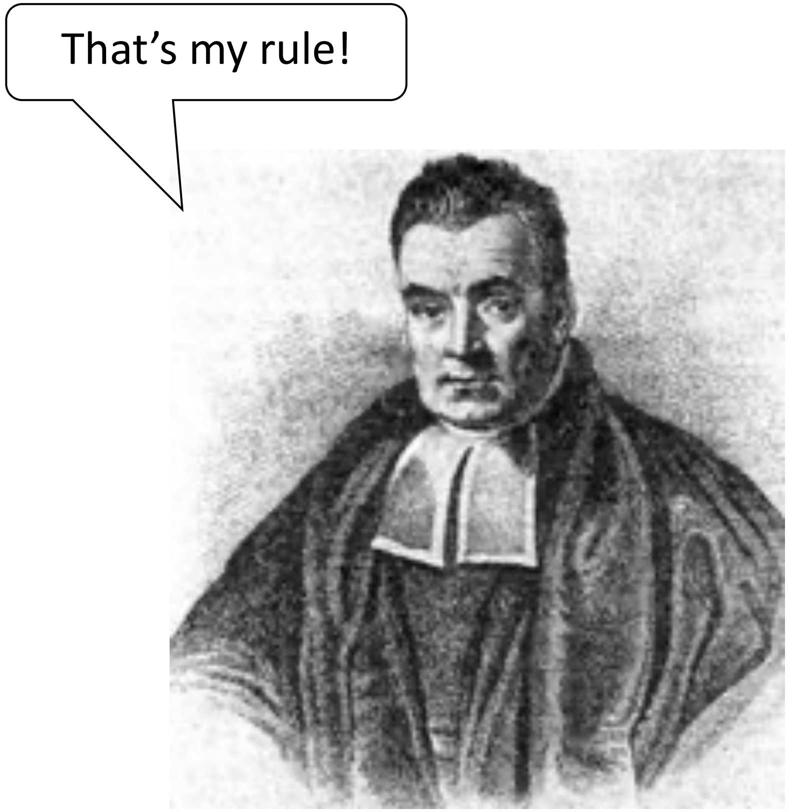
- 根据乘法法则，联合分布可以表示为:

$$P(a,b) = P(a|b)P(b) \qquad P(a,b) = P(b|a)P(a)$$

- 同时除以P(a)，得到贝叶斯规则:

$$P(b|a) = \frac{P(a|b)P(b)}{P(a)}$$

- 是大多数进行概率推理的人工智能系统的基础
- 贝叶斯规则在实践中很有用：
  - 很多情况下，前三项有很好的估计，而需要计算第4项



贝叶斯规则应用

- 简单示例: 医疗诊断

- 结果effect看作是证据，确定造成这一结果的未知因素cause ，贝叶斯规则:

$$P(\text{cause}|\text{effect}) = \frac{P(\text{effect}|\text{cause})P(\text{cause})}{P(\text{effect})}$$

  - 条件概率P(effect|cause) 量化了因果方向上的关系
  - 条件概率P(cause|effect) 描述诊断方向上的关系
  - 实际中，经常有因果关系的条件概率，而想得出诊断关系。

贝叶斯规则应用

- 简单实例: 医疗诊断

  - 医生知道脑膜炎有70%的可能会引起病人脖子僵硬。医生还了解一些无条件的事实：病人患脑膜炎的先验概率为1/50,000，而任何一个病人脖子僵硬的概率为 1%。问：脖子僵硬的病人患脑膜炎的概率？

  - 令m表示”病人患有脑膜炎”的命题, s表示”病人脖子僵硬”的命题

  - 则有：$P(m) = 1/50000 \qquad P(s|m) = 0.7 \qquad P(s) = 0.01$

  - $P(m|s) = \frac{P(s|m) * P(m)}{P(s)} = (0.7 * 1/50000) / 0.01 = 0.0014$

贝叶斯规则应用

- 简单实例: 医疗诊断

  - 医生知道脑膜炎有70%的可能会引起病人脖子僵硬。医生还了解一些无条件的事实：病人脖子僵硬的先验概率为1/50,000, 而任何一个病人脖子僵硬的概率为 1%。
  - 问题：计算$\mathbf{P}(M|s)$？
  - 归一化的方法：

$$\mathbf{P}(M|s) = \langle P(m|s), P(\neg m|s) \rangle$$

$$= \alpha \langle P(s|m)P(m), P(s|\neg m)P(\neg m) \rangle$$

贝叶斯规则应用

- 简单示例2:　明天要举行户外运动会。近年来，每年仅下雨5天（5/365=0.014）。不幸的是，天气预报员预测明天会下雨。当真的下雨时，天气预报员准确地预测了90%的降雨。当不下雨时，他错误地预测了10%的降雨。明天下雨和不下雨的可能性分别有多大？

- 令rain表示明天下雨的概率，predict表示预测明天下雨的概率
- 则有：
  - $P(rain) = 0.014$；$P(\neg rain) = 0.986$；$P(predict|rain) = 0.9$；$P(predict|\neg rain) = 0.1$
- 问题：计算$P(Rain|predict)$？

贝叶斯规则应用

- 简单示例2: 明天要举行户外运动会。近年来，每年仅下雨5天（5/365=0.014）。不幸的是，天气预报员预测明天会下雨。当真的下雨时，天气预报员准确地预测了90%的降雨。当不下雨时，他错误地预测了10%的降雨。明天下雨和不下雨的可能性分别有多大？

$$\mathbf{P}(\mathrm{Rain}|\mathrm{predict}) = \langle P(\mathrm{rain}|\mathrm{predict}), P(\neg \mathrm{rain}|\mathrm{predict}) \rangle$$

$$= \alpha \langle P(\mathrm{predict}|\mathrm{rain}) * P(\mathrm{rain}), P(\mathrm{predict}|\neg \mathrm{rain}) * P(\neg \mathrm{rain}) \rangle$$

$$= \alpha \langle 0.9 * 0.014, 0.1 * 0.986 \rangle$$

$$= \langle 0.111 ，0.889 \rangle$$

概率推理知识点

01

02

贝叶斯规则及应用

03

04

贝叶斯网络推理

贝叶斯网络推理案例

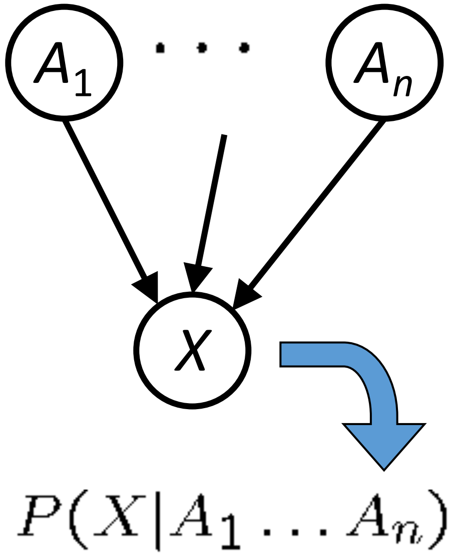
贝叶斯网络

- 完全概率分布能够回答关于问题域的任何问题，但是:
  - 随着变量数目增多会增大到不可操作的程度，计算复杂性十分巨大
  - 为每个可能世界逐个指定概率是不自然的
- 贝叶斯网络: 一种使用简单局部分布（条件概率）描述复杂联合分布的技术
  - 是一个有向无环图，也称之为图模型
  - 用于表示变量之间的依赖关系
  - 本质上，表示任何完全联合概率分布

贝叶斯网络模型

- 语法：
  - 一个结点对应一个随机变量X_i，可以是已知的或者是待观测的
  - 有向边表示连接的变量间的“直接影响”
  - 条件概率分布，量化父结点对该结点的影响

$$\mathbf{P}(X_i|Parents(X_i))$$

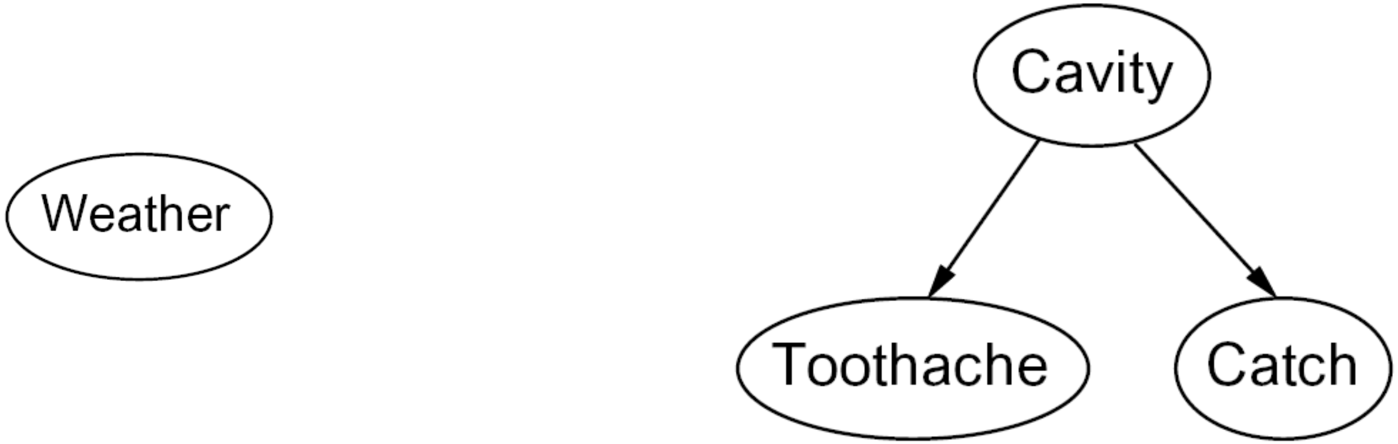


- 条件概率表(conditional probability table, CPT), 为每个变量指定其相对于父结点的条件概率

贝叶斯网络= 拓扑结构(图) + 条件概率

贝叶斯网络模型

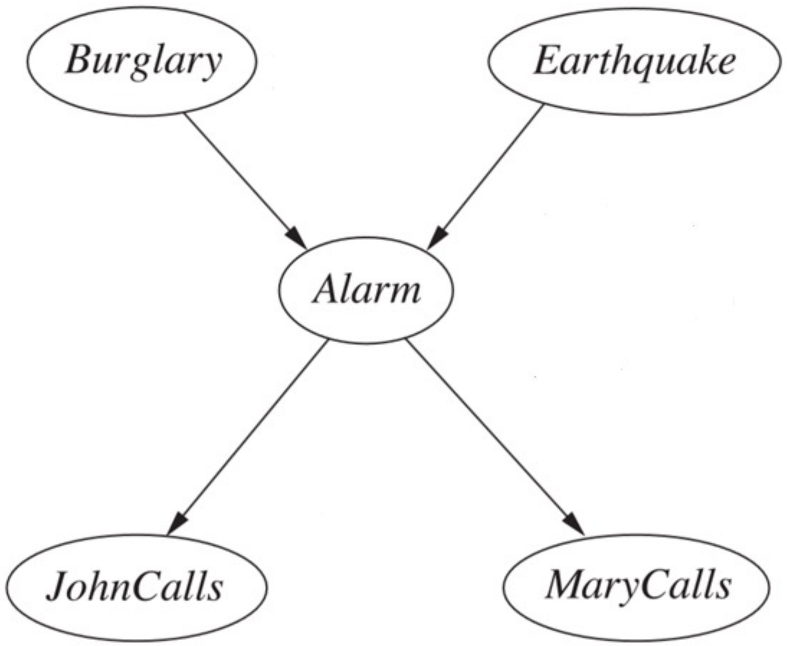
- 贝叶斯网络的拓扑结构
  - 用一种精确简洁的方式描述了在问题域中成立的独立和条件独立关系



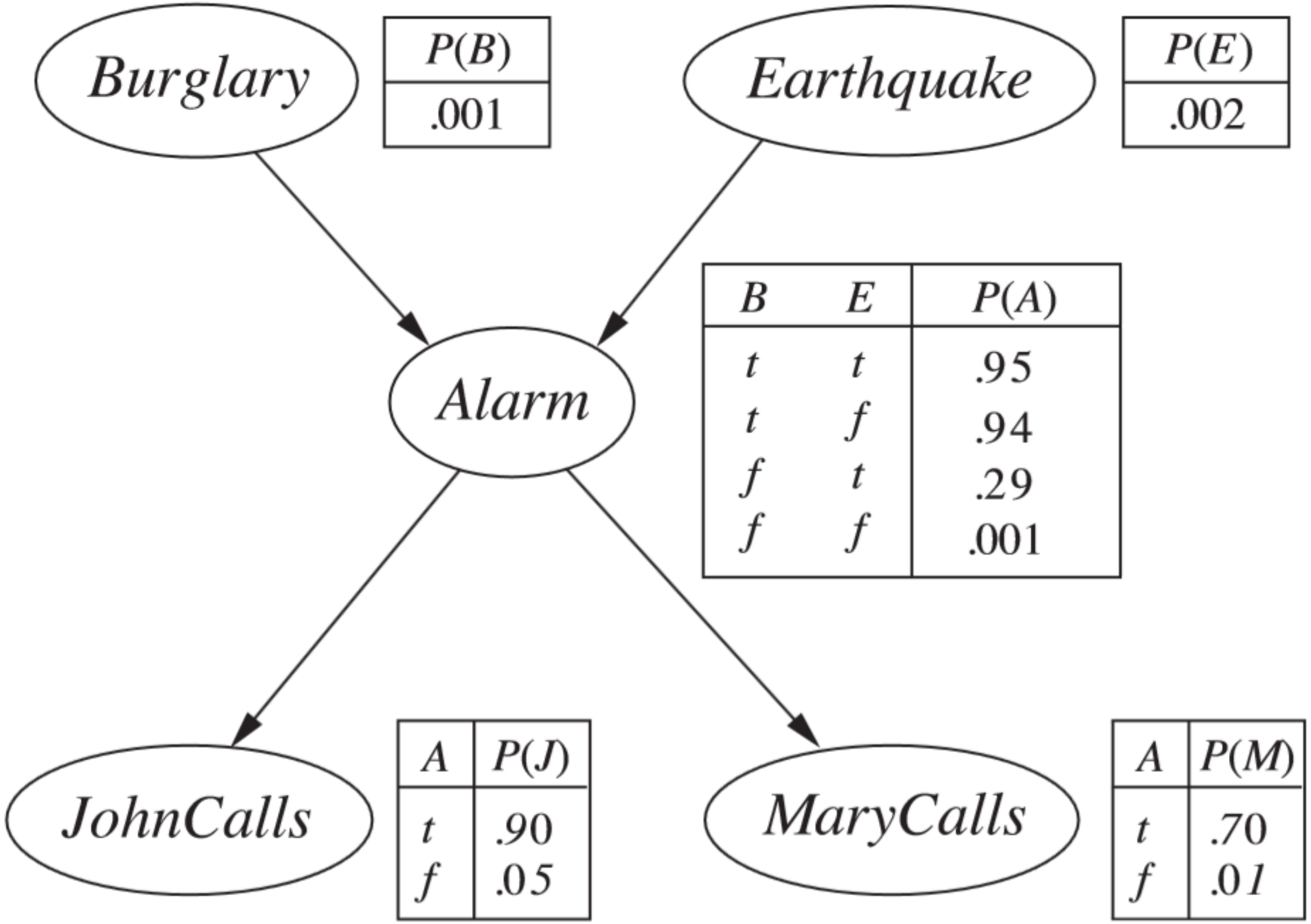
  - *Weather* 与其他变量是独立的
  - 给定 *Cavity* 条件下，*Toothache* 和 *Catch* 是条件独立的

示例: Alarm Network

- 一个贝叶斯网络： Alarm Network
  - 我在上班，邻居约翰打电话说我的报警器响了，但邻居玛丽不打电话。有时，它会受到轻微地震的影响。有窃贼吗？
  - 布尔变量: *Burglary*, *Earthquake*, *Alarm*, *JohnCalls*, *MaryCalls*
  - 网络拓扑反应了因果关系
    - 窃贼闯入，报警器会响 *Burglary -> Alarm*
    - 地震偶尔报警器也会有反应 *Earthquake -> Alarm*
    - 听到警报声，John会打来电话 *Alarm -> JohnCalls*
    - 听到警报声，Mary会打来电话 *Alarm -> MaryCalls*



示例: Alarm Network



$P(B)$
.001

$P(E)$
.002

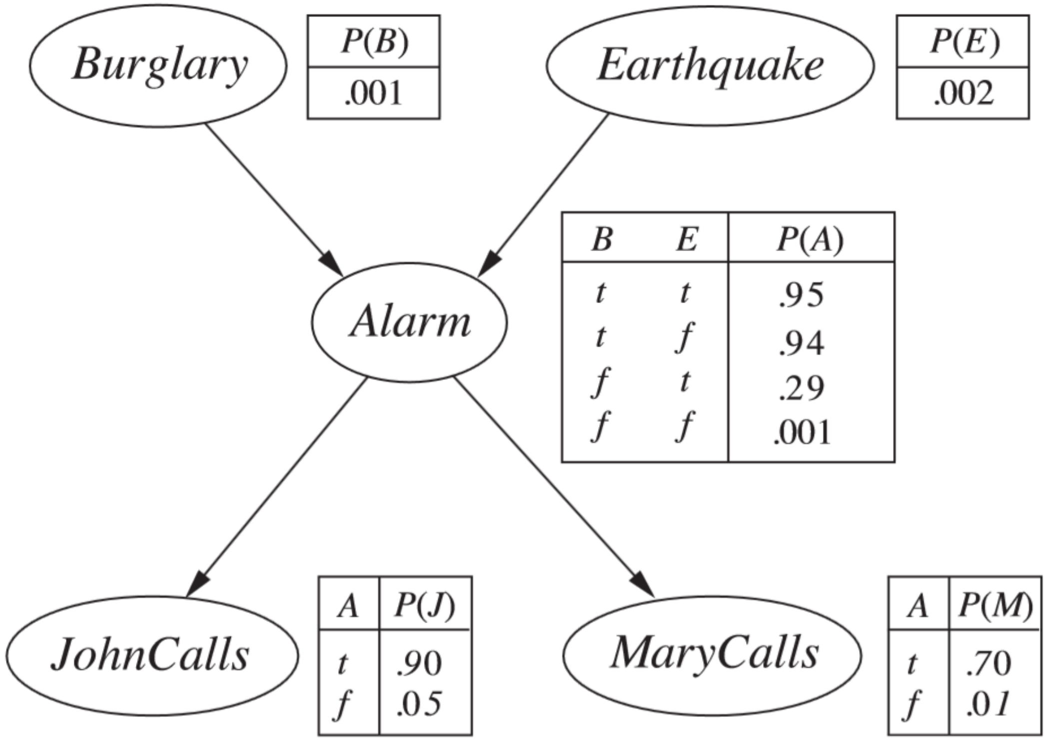
B	E	$P(A)$
t	t	.95
t	f	.94
f	t	.29
f	f	.001

A	$P(J)$
t	.90
f	.05

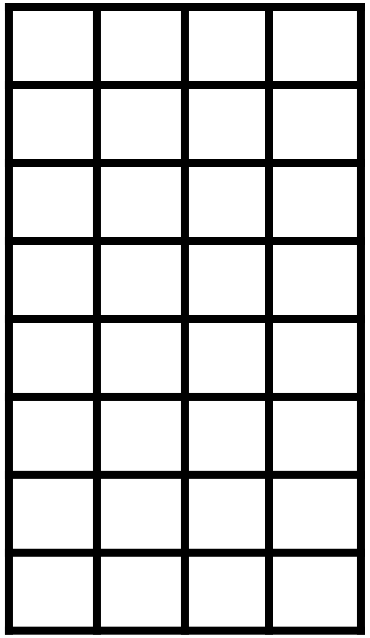
A	$P(M)$
t	.70
f	.01

贝叶斯网络的语义

- 贝叶斯网络是完全联合概率分布的一种表示



贝叶斯网络



完全联合概率分布

贝叶斯网络的语义

TONGJI UNIVERSITY

- 贝叶斯网络是完全联合概率分布的一种表示
  - 完全联合概率分布由局部条件概率分布的乘积进行定义（链式法则）

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid parents(X_i))$$

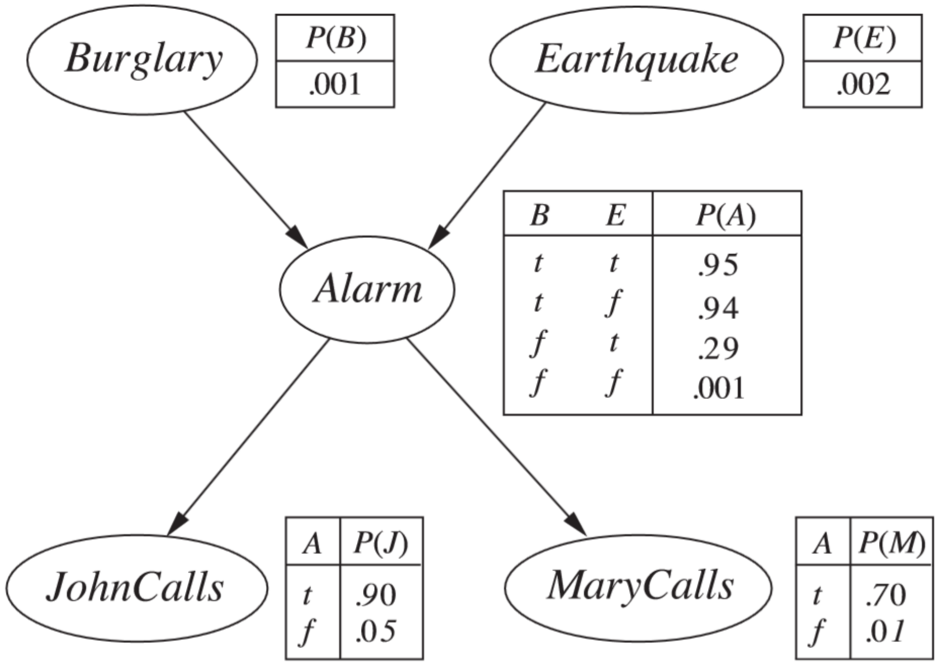
  - 例如:

$P(j, m, a, \neg b, \neg e)$

$= P(j|a)\, P(m|a)\, P(a|\neg b,\neg e)\, P(\neg b)\, P(\neg e)$

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

≈ 0.00063



构造贝叶斯网络的方法

- 1. 结点：选择一组排好序的随机变量$X_1,\ldots,X_n$
- 2. 边：i从1 到 n，执行：

  2.1 从$X_1,\ldots,X_{i-1}$中选择X_i的父结点最小的集合，使得

  $$\mathbf{P}(X_i|Parents(X_i)) = \mathbf{P}(X_i \mid X_1, \ldots, X_{i-1})$$

  2.2 在每一个父结点与X_i之间插入一条边

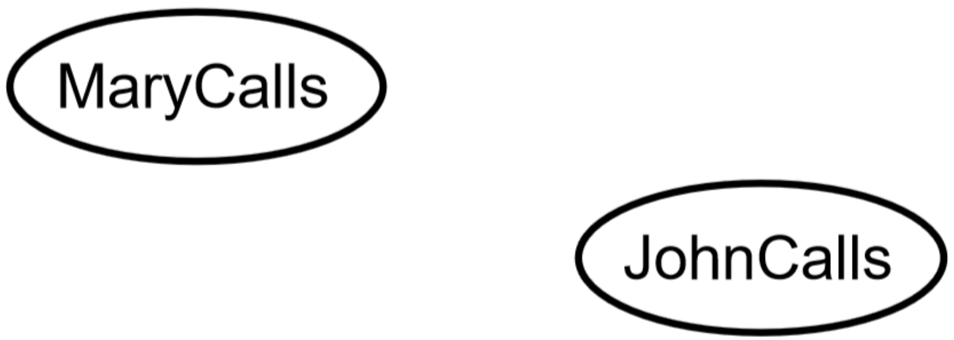
  2.3 条件概率表（CPTs）：写出条件概率表$\mathbf{P}(X_i|Parents(X_i))$

构造贝叶斯网络的方法

示例： Alarm Network

假设选定随机变量的次序：M, J, A, B, E

$$P(J|M) = P(J)?$$



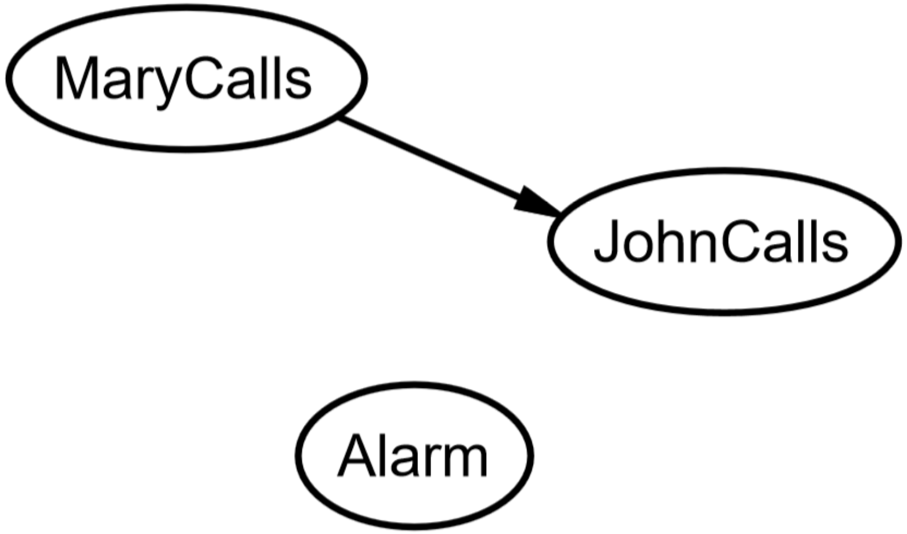
构造贝叶斯网络的方法

示例： Alarm Network

假设选定随机变量的次序：M, J, A, B, E

$P(J|M) = P(J)$? No

$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$?



构造贝叶斯网络的方法

示例： Alarm Network

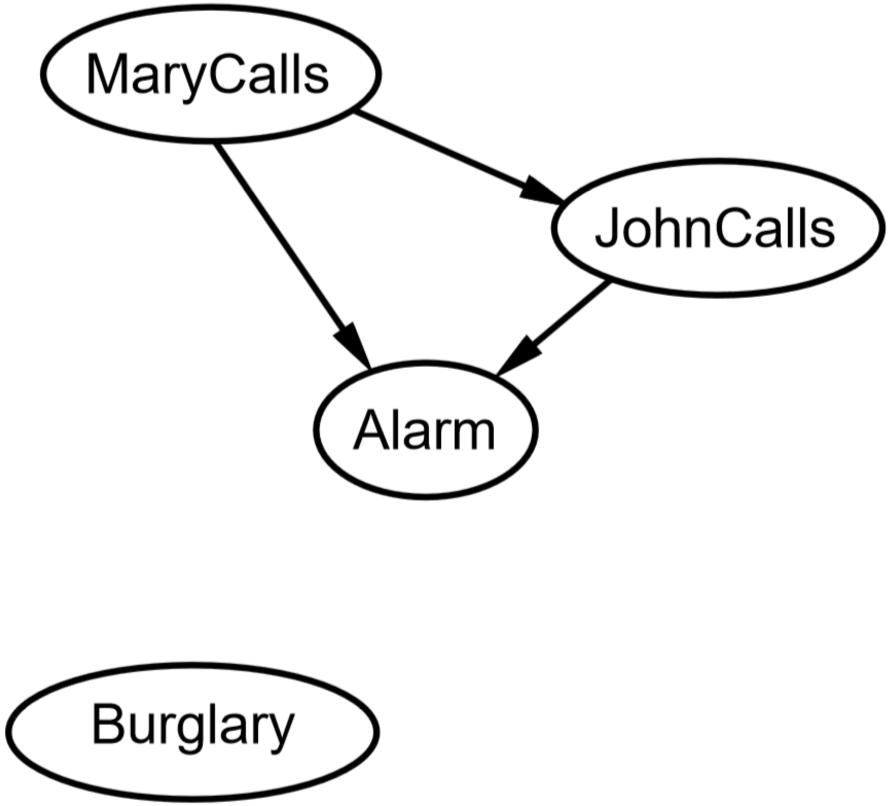
假设选定随机变量的次序：M, J, A, B, E

$P(J|M) = P(J)$? No

$P(A|J,M) = P(A|J)$? $P(A|J,M) = P(A)$? No

$P(B|A,J,M) = P(B|A)$?

$P(B|A,J,M) = P(B)$?



构造贝叶斯网络的方法

示例： Alarm Network

假设选定随机变量的次序：M, J, A, B, E

$P(J|M) = P(J)?$ No

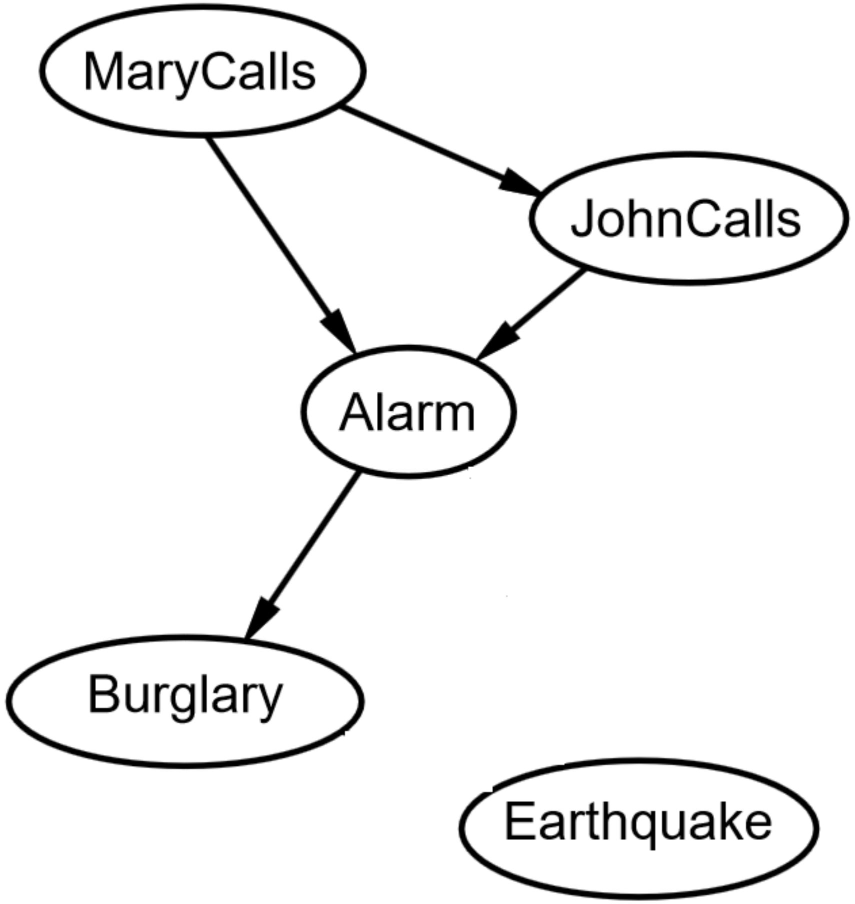
$P(A|J,M) = P(A|J)?$ $P(A|J,M) = P(A)?$ No

$P(B|A,J,M) = P(B|A)?$ Yes

$P(B|A,J,M) = P(B)?$ No

$P(E|B,A,J,M) = P(E|A)?$

$P(E|B,A,J,M) = P(E|A,B)?$



构造贝叶斯网络的方法

示例： Alarm Network

假设选定随机变量的次序：M, J, A, B, E

$P(J|M) = P(J)?$ No

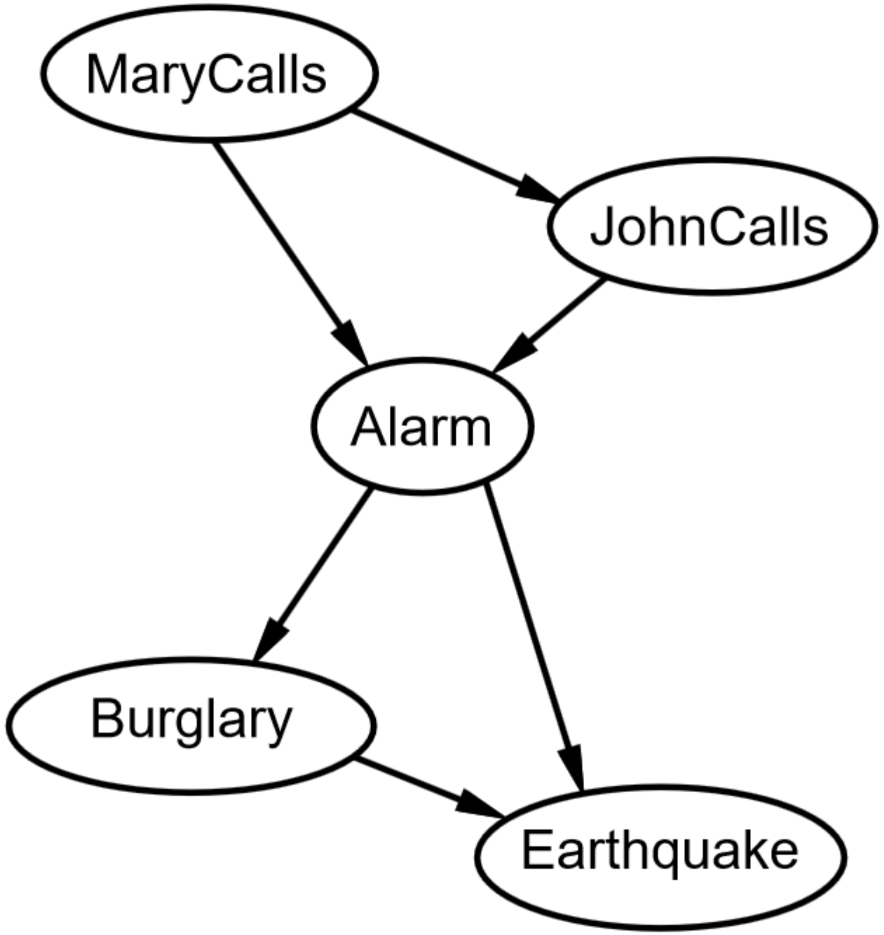
$P(A|J,M) = P(A|J)?\ P(A|J,M) = P(A)?$ No

$P(B|A,J,M) = P(B|A)?$ Yes

$P(B|A,J,M) = P(B)?$ No

$P(E|B,A,J,M) = P(E|A)?$ No

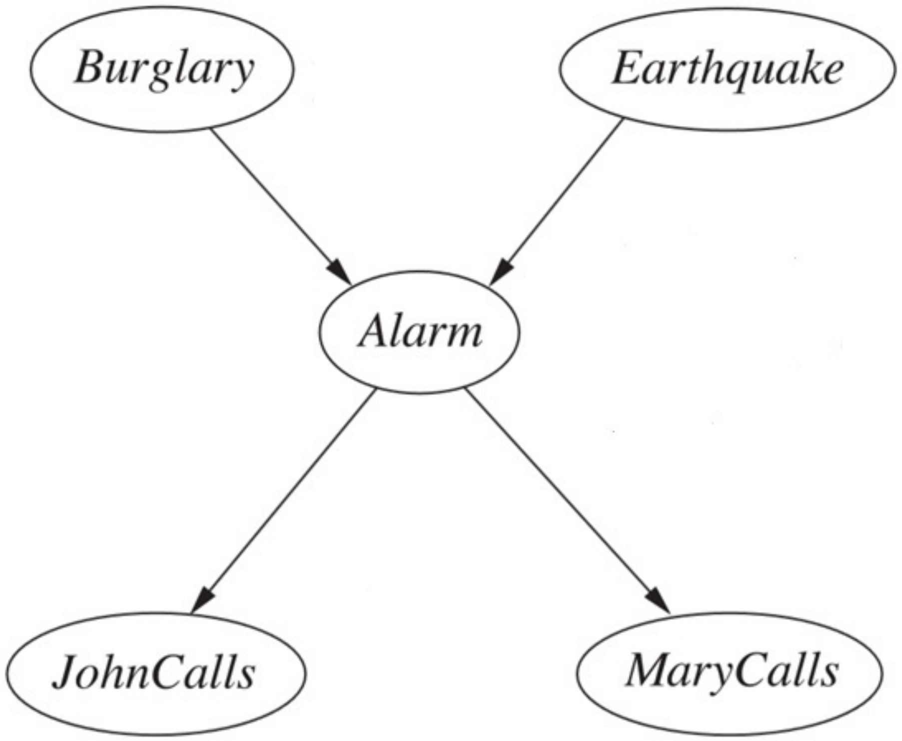
$P(E|B,A,J,M) = P(E|A,B)?$ Yes

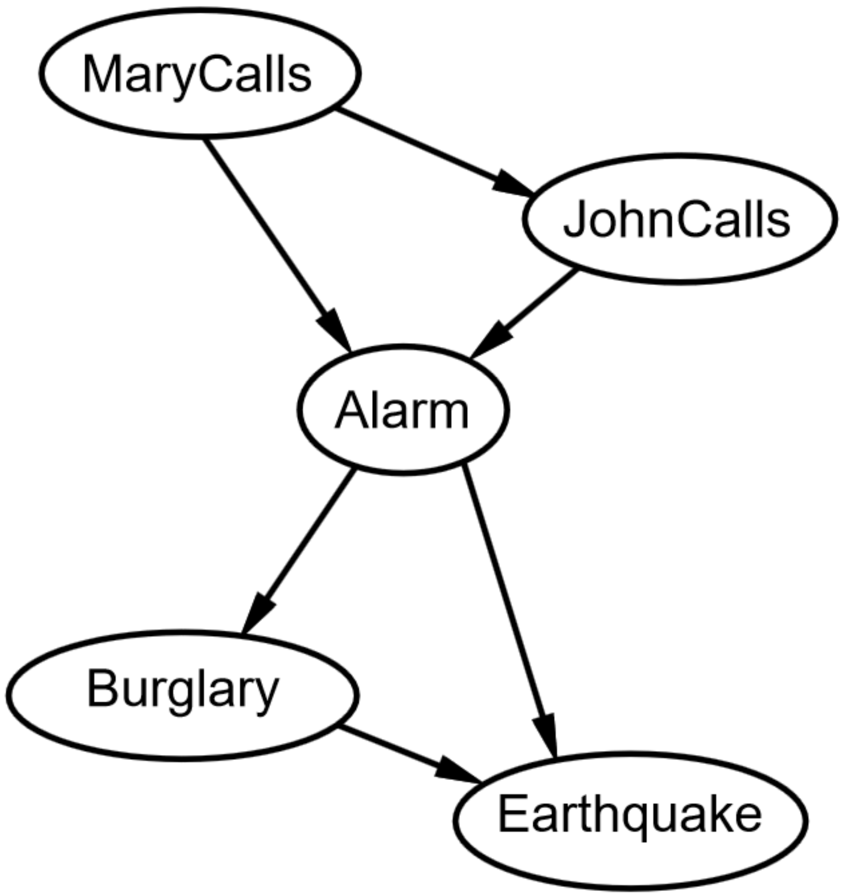


构造贝叶斯网络的方法

示例： Alarm Network

假设选定随机变量的次序：M, J, A, B, E





贝叶斯枚举推理

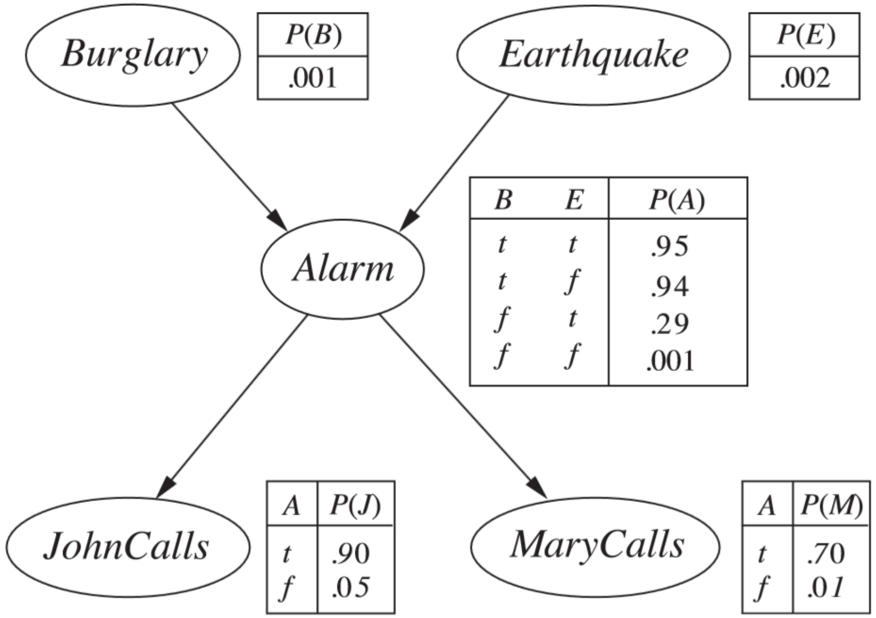
- 基本任务:
  - 证据变量: $E_1 \ldots E_k = e_1 \ldots e_k$
  - 查询变量: Q
  - 隐藏变量: $H_1 \ldots H_r$

$\left.\begin{matrix} E_1 \ldots E_k = e_1 \ldots e_k \\ Q \\ H_1 \ldots H_r \end{matrix}\right\} X_1, X_2, \ldots X_n$ *All variables*

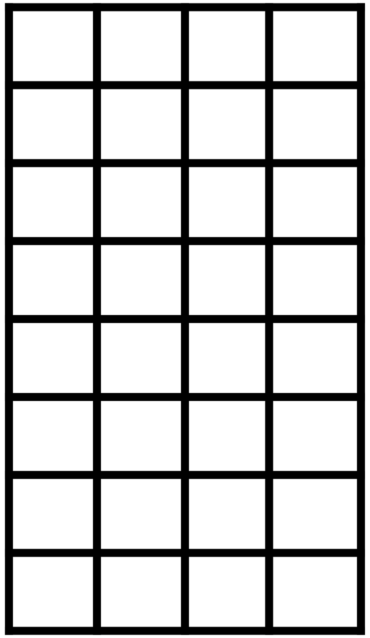
- 典型的查询是询问后验概率:

$$P(Q|e_1 \ldots e_k)$$

贝叶斯网络



联合分布



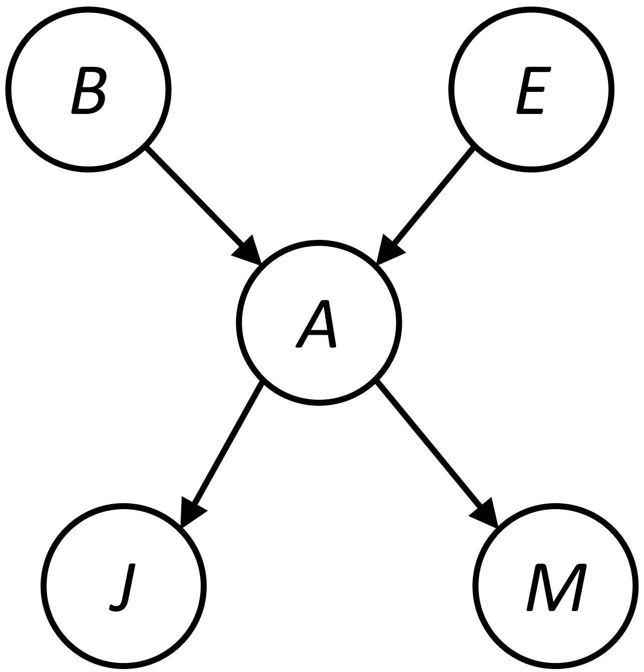
查询

$$P(a \mid e)$$

贝叶斯枚举推理

示例： Alarm Network

- 已知John和Mary都打来了电话，问出现盗贼的概率是多少？
- 分析：
  - 证据变量: $JohnCalls$ =true， $MaryCalls$ =true
  - 查询变量: Burglary
  - 隐藏变量: $Earthquake$, $Alarm$
- 问题：$P(Burglary|JohnCalls = \text{true}, MaryCalls = \text{true}) = ?$
- 简写为：$P(B|j,m)$ ？



贝叶斯枚举推理

- 直接求和联合概率分布中的变量，而不进行实际概率的显式计算

根据归一化方法

$$\mathbf{P}(B|j,m)$$
$$=\alpha\ \mathbf{P}(B,j,m)$$
$$=\alpha\sum_{e}\sum_{a}\mathbf{P}(B,e,a,j,m)$$

贝叶斯枚举推理

- 直接求和联合概率分布中的变量，而不进行实际概率的显式计算

根据归一化方法

$$\mathbf{P}(B|j,m)$$
$$=\alpha\ \mathbf{P}(B,j,m)$$
$$=\alpha\sum_{e}\sum_{a}\mathbf{P}(B,e,a,j,m)$$

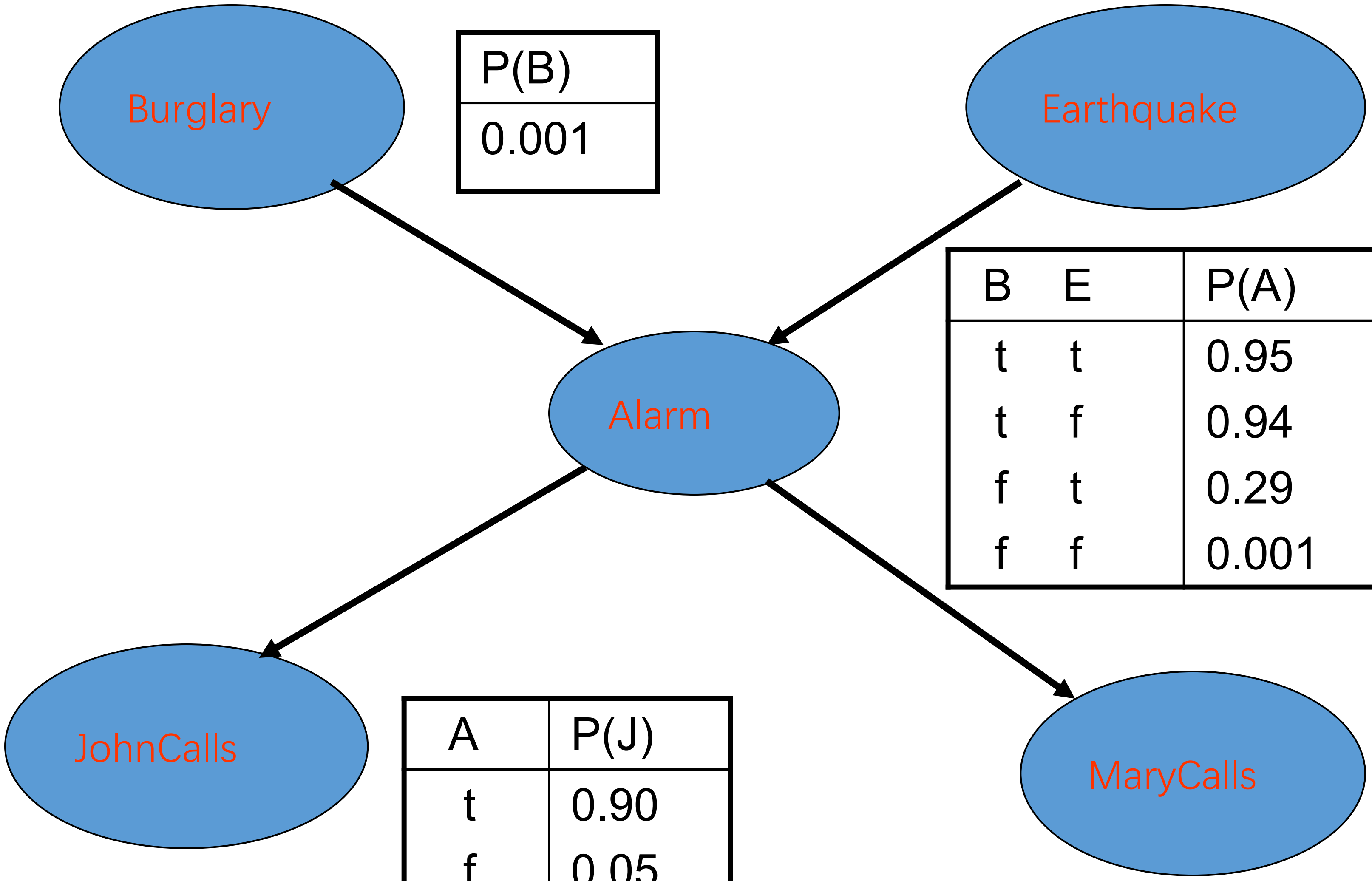
根据贝叶斯网络的语义

$$=\alpha\sum_{e}\sum_{a}\mathbf{P}(B)\ P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a)$$

简化计算

$$=\alpha\mathbf{P}(B)\sum_{e}P(e)\sum_{a}\mathbf{P}(a|B,e)P(j|a)\ P(m|a)$$

贝叶斯枚举推理



P(B)
0.001

P(E)
0.002

B	E	P(A)
t	t	0.95
t	f	0.94
f	t	0.29
f	f	0.001

A	P(J)
t	0.90
f	0.05

A	P(M)
t	0.70
f	0.01

贝叶斯枚举推理

- 已知，一个事件e = {JohnCalls = true, and MaryCalls = true}，试问出现盗贼的概率是多少？
- 解：$P(X|e) = \alpha P(X,e) = \alpha \sum_y P(X,e,y)$

  而P(X,e,y)可写成条件概率乘积的形式。

  因此，在贝叶斯网络中可通过计算条件概率的乘积并求和来回答查询。

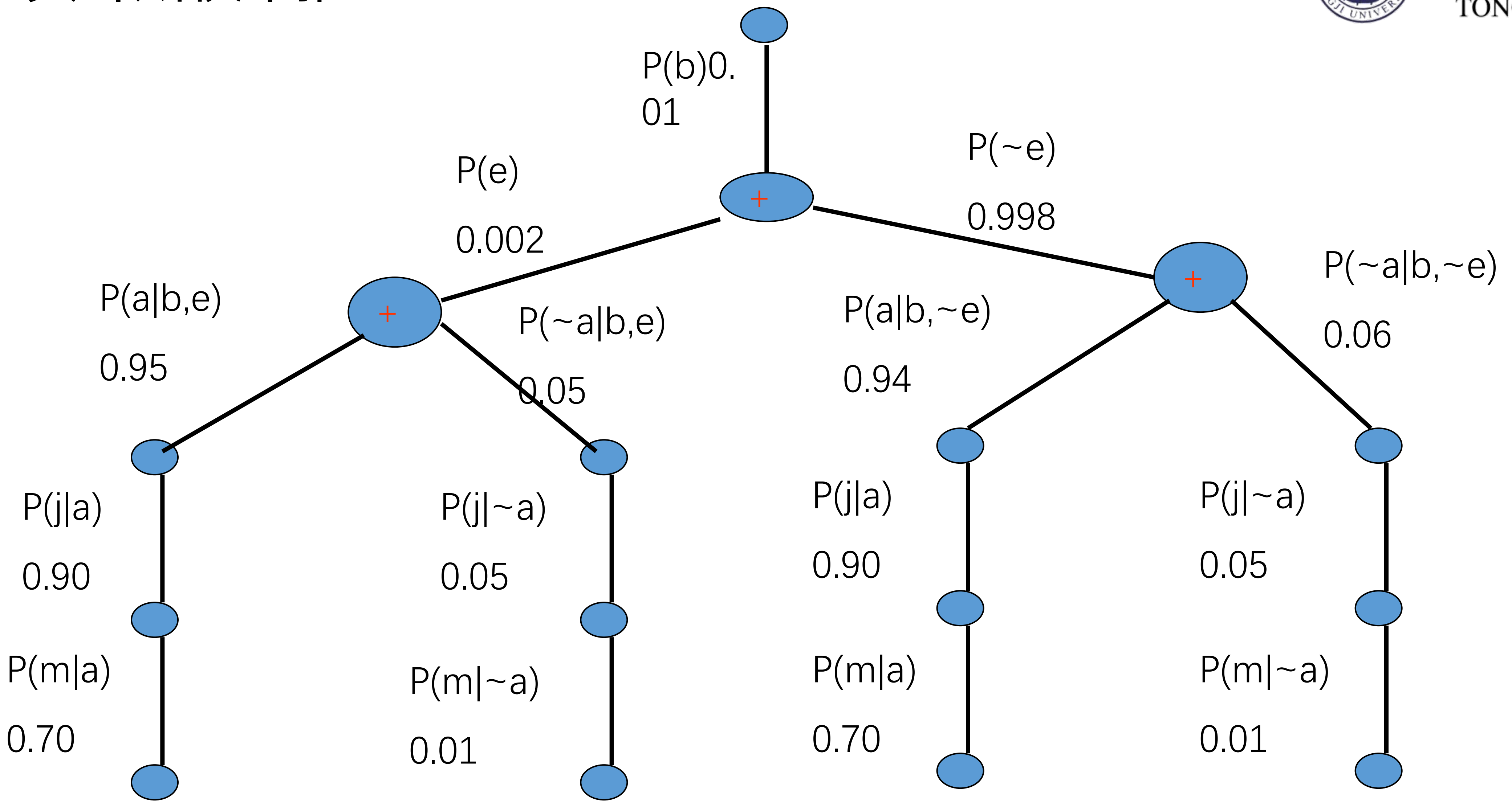
P(Burgary | JohnCalls = true, MaryCalls = true)简写为：

$$P(B \mid j, m) = \alpha P(B, j, m) = \alpha \sum_e \sum_a P(B, e, a, j, m)$$

$$= \alpha \sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a)$$

$$= \alpha P(b) \sum_e P(e) \sum_a P(a|b,e)P(j|a)P(m|a)$$

贝叶斯枚举推理



P(b | j, m)的自顶向下的计算过程

贝叶斯枚举推理

同濟大學 TONGJI UNIVERSITY

- $P(B \mid j, m) = \alpha P(B, j, m) = \alpha \sum_e \sum_a P(B, e, a, j, m)$

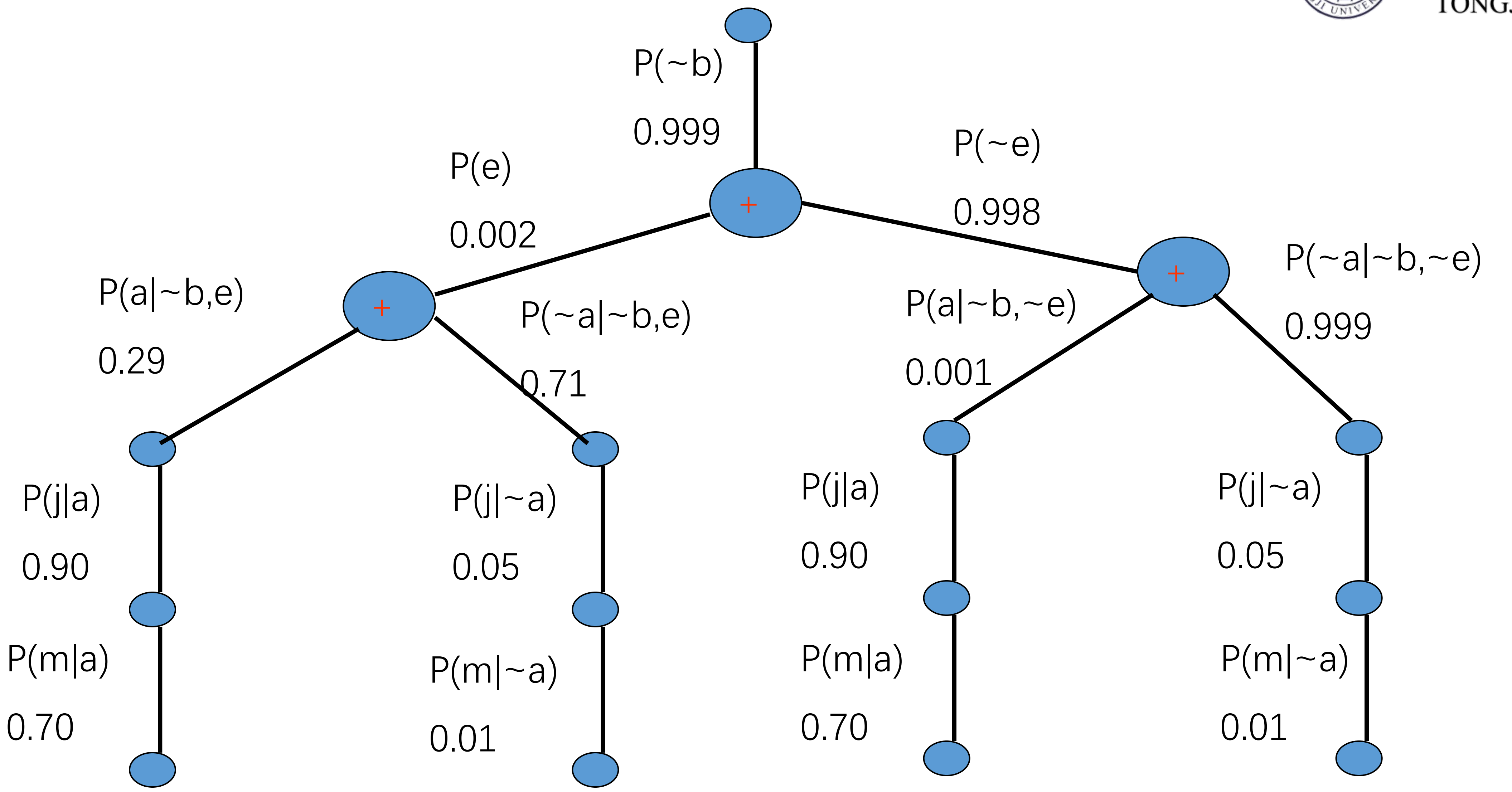
  $= \alpha \sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a)$

  $= \alpha P(b) \sum_e P(e) \sum_a P(a|b,e)P(j|a)P(m|a)$

  $= \alpha \times 0.001 \times \{[0.002 \times (0.95 \times 0.9 \times 0.7 + 0.05 \times 0.05 \times 0.01)] + [0.998 \times (0.94 \times 0.9 \times 0.7 + 0.06 \times 0.05 \times 0.01)]\}$

  $= \alpha \times 0.00059224$

贝叶斯枚举推理



P(~ b | j, m)的自顶向下的计算过程

贝叶斯枚举推理

- $P(\sim B \mid j, m) = \alpha P(\sim B, j, m) = \alpha \sum_e \sum_a P(\sim B, e, a, j, m)$

$= \alpha \sum_e \sum_a P(\sim b)P(e)P(a|\sim b,e)P(j|a)P(m|a)$

$= \alpha P(\sim b) \sum_e P(e) \sum_a P(a|\sim b,e)P(j|a)P(m|a)$

$= \alpha \times 0.999 \times \{[0.002 \times (0.29 \times 0.9 \times 0.7 + 0.71 \times 0.05 \times 0.01)] + [0.998 \times (0.001 \times 0.9 \times 0.7 + 0.999 \times 0.05 \times 0.01)]\}$

$= \alpha \times 0.0014919$

因此，$P(B|j, m) = \alpha \langle 0.00059224, 0.0014919 \rangle$

$\approx \langle 0.284, 0.716 \rangle$

即在John和Mary都打电话的条件下，出现盗贼的概率约为28%。

概率推理知识点

01

贝叶斯规则及应用

02

贝叶斯网络推理

03

贝叶斯网络推理案例

04

案例：学生健康风险预测模型

背景

已知：一个事件e = {学校政策U = true, and 工作压力大 = true}，

请根据上述枚举法计算学生出现健康问题的概率。 目标是基于贝叶斯网络的推理来计算该学生出现健康问题的概率。

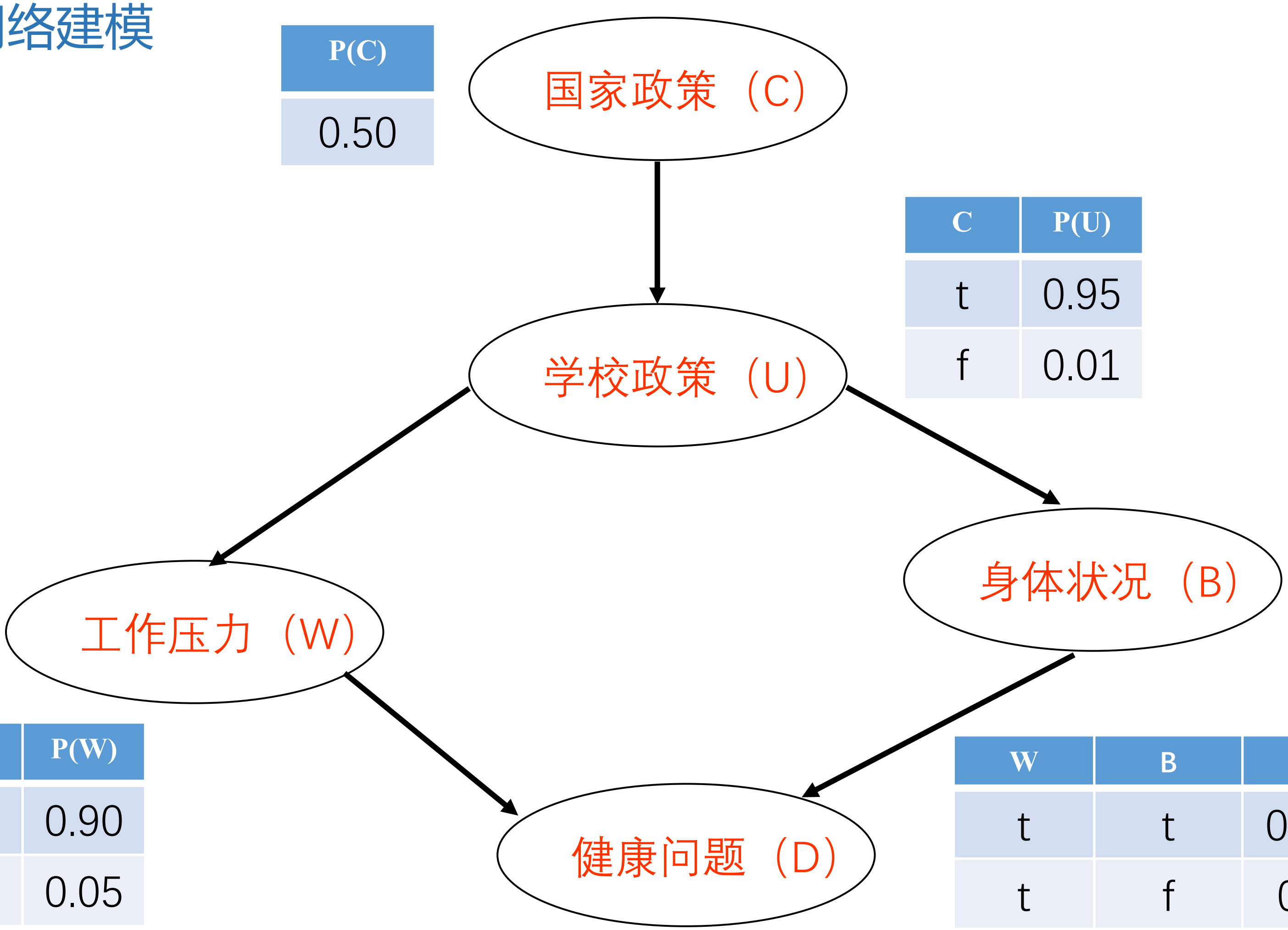
贝叶斯网络中涉及到的节点和条件概率如下所示：

1. 国家政策（C）: 可能的值为True和False，且$P(C=True) = 0.50$。
2. 学校政策（U）: 根据国家政策C的不同，U的概率也不同。当C=True时，$P(U=True) = 0.95$；当C=False时，$P(U=True) = 0.01$。
3. 工作压力大（W）: 该值依赖于学校政策U，$P(W=True \mid U=True) = 0.90$，$P(W=True \mid U=False) = 0.05$。
4. 身体状况差（B）: 该值依赖于学校政策U，$P(B=True \mid U=True) = 0.30$，$P(B=True \mid U=False) = 0.01$。
5. 健康问题（D）: 该值依赖于工作压力W和身体状况B的值。健康问题的概率如下：

- $P(D=True \mid W=True, B=True) = 0.335$
- $P(D=True \mid W=True, B=False) = 0.30$
- $P(D=True \mid W=False, B=True) = 0.05$
- $P(D=True \mid W=False, B=False) = 0.00$

案例：学生健康风险预测模型

贝叶斯网络建模



P(C)
0.50

C	P(U)
t	0.95
f	0.01

U	P(W)
t	0.90
f	0.05

U	P(B)
t	0.30
f	0.01

W	B	P(D)
t	t	0.335
t	f	0.30
f	t	0.05
f	f	0.00

案例：学生健康风险预测模型

推理过程：

- 计算P(D=True | U=True, W=True)

解：步骤 1：根据贝叶斯网络公式，需要计算的条件概率为：

$$P(D = \text{True} \mid U = \text{True}, W = \text{True}) = \textstyle\sum_B P\,(D = \text{True} \mid W = \text{True}, B)\ P(B \mid U = \text{True})$$

步骤 2：根据给定的条件概率计算

我们有以下条件概率表：

$P(B = \text{True} \mid U = \text{True}) = 0.30$

$P(B = \text{False} \mid U = \text{True}) = 0.70$

$P(D = \text{True} \mid W = \text{True}, B = \text{True}) = 0.335$

$P(D = \text{True} \mid W = \text{True}, B = \text{False}) = 0.30$

推理过程：

步骤 3 : 计算 $P(D = True \mid W = True, U = True)$

$$P(D = True \mid W = True, U = True) = P(D = True \mid W = True, B = True)\, P(B = True \mid U = True) + P(D = True \mid W = True, B = False)\, P(B = False \mid U = True)$$

代入已知的值：

$$\begin{aligned} P(D = True \mid W = True, U = True) &= (0.335 \times 0.30) + (0.30 \times 0.70) \\ &= 0.1005 + 0.21 \\ &= 0.3105 \end{aligned}$$

结论： 学生出现健康问题的概率约为 **31.05%**。

案例：学生健康风险预测模型

代码：

```
# 给定条件概率
P_B_true_given_U_true = 0.30
P_B_false_given_U_true = 0.70
P_D_true_given_W_true_B_true = 0.335
P_D_true_given_W_true_B_false = 0.30

# 计算 P(D=True | U=True, W=True)
P_D_true_given_U_true_W_true = (P_D_true_given_W_true_B_true * P_B_true_given_U_true) +
(P_D_true_given_W_true_B_false * P_B_false_given_U_true)

print(f"P(D = True | U = True, W = True) = {P_D_true_given_U_true_W_true:.4f}")
```

```
P(D = True | U = True, W = True) = 0.3105
```