贝叶斯网络（Bayesian network）：表示变量之间的依赖关系，是联合分布的一种表示

贝叶斯网络是一个有向图，其中每个节点用定量的概率信息标记。完整的描述如下。

（1）**每个节点对应一个随机变量**，它可以是离散的，也可以是连续的。

（2）有向链路或箭头连接成对的节点。如果有从节点X指向节点Y的箭头，则称X是Y的一个父节点。图中没有有向环，因此它是一个有向无环图（简称DAG）。

（3）每个节点Xi关联概率信息 ，我们使用有限的参数量化父节点对该节点的影响。

每个节点的局部概率信息采用**条件概率表（conditional probability table，CPT）**的形式表示。（条件概率表只能用于离散变量）

条件概率表的每一行包含某个条件事件 （conditioning case）下每个节点的值的条件概率。**条件事件**就是**所有父节点**的**一个可能的值的组合**，也可以把它视为一个微型的可能世界。

**13.1不确定域的知识表示**

贝叶斯网络（Bayesian network）: 表示变量之间的依赖关系

贝叶斯网络是一个有向图，其中每个节点用定量的概率信息标记。完整的描述如下。

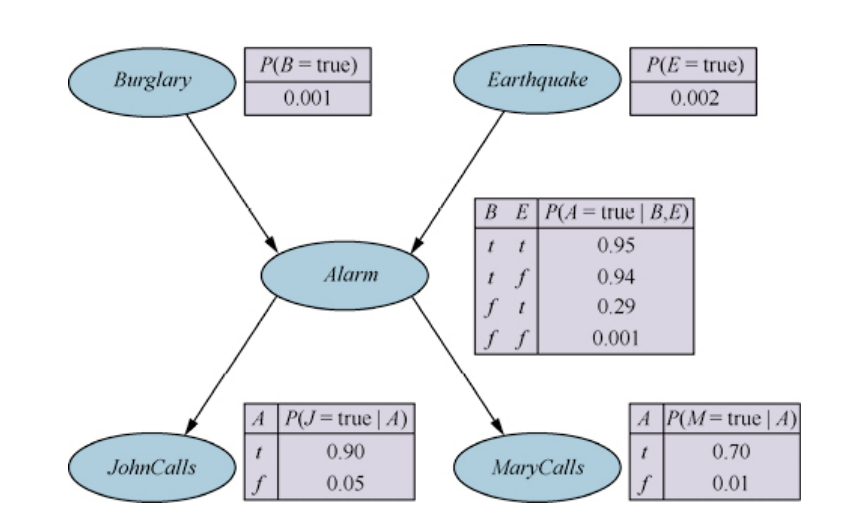
（1）每个节点对应一个随机变量，它可以是离散的，也可以是连续的。

（2）有向链路或箭头连接成对的节点。如果有从节点X指向节点Y的箭头，则称X是Y的一个父节点。图中没有有向环，因此它是一个有向**无环**图（简称DAG）。

（3）每个节点Xi关联概率信息，我们使**用有限的参数量化父节点对该节点的影响。**

**拓扑结构+局部概率信息**

每个节点的局部概率信息采用条件概率表（conditional probability table，CPT）的形式表示。



**13.2贝叶斯网络的语义**

贝叶斯网络的**语法**由一个有向无环图(DAG)和条件概率分布(CPDs)的组合组成, 来表示复杂的概率分布。

**语义**定义了语法如何对应于网络的变量的联合分布。

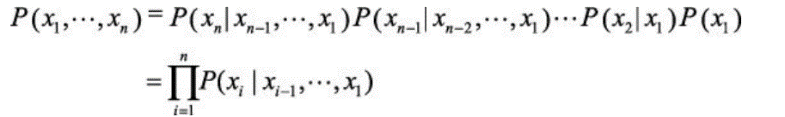
**贝叶斯网络的结构**

贝叶斯网络由两个部分组成：

1. 图形结构：一个有向无环图，节点表示随机变量，有向边表示直接依赖关系

2. 条件概率分布(CPDs)：每个节点在给定其父节点集合的条件下概率分布

贝叶斯网络的图形结构捕捉了变量之间的条件独立关系，而CPDs则提供了定量的概率信息

链式法则（chain rule）

对任何随机变量集合都成立

通过拓扑顺序（topological order，即与有向图结构一致的任何顺序）对节点进行编号

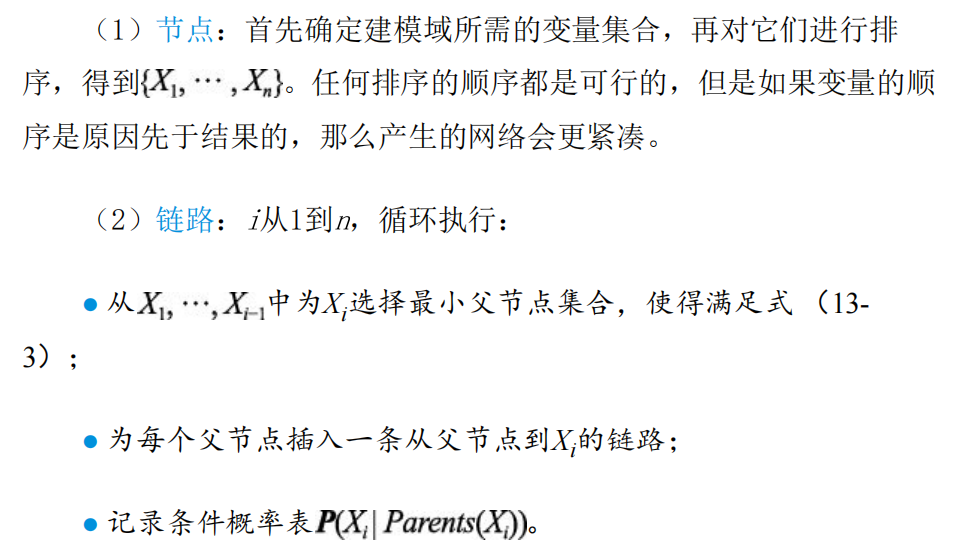
只有在给定父节点，每个节点条件独立于节点排序中的其他前驱节点时，贝叶斯网络才是域的正确表示。

贝叶斯网络的联合概率表示

贝叶斯网络提供了一种高效的联合概率分布表示方法。对于由n个节点组成的网络，其联合概率分布可以表示为各个节点条件概率的乘积：

P(X₁,X₂,…,Xₙ) = ∏ₐ P(Xₐ | Parents(Xₐ))

**1.构造贝叶斯网络的方法**



节点Xi的父节点应该包含X1, …, Xi−1中所有**直接影响**Xi的变量。（间接影响不应该被算入）

三种方式： 1. 专家知识：由领域专家根据领域知识手动构建

2. 数据学习：从数据中自动学习网络结构和参数

3. 混合方法：结合专家知识和数据学习

在实践中，通常先由专家确定大致结构，再用数据估计参数或微调结构。

**2. 紧凑性与节点顺序**

贝叶斯网络的紧凑性是局部结构化（locally structured）［也称为 稀疏（sparse）］系统一般性质的一个示例。

在局部结构化的系统中，每个子组件仅与**有限数量**的其他组件直接交互，而与组件总数无关。

复杂度上的增长速度通常是**线性**的而非指数级别的。

在使用贝叶斯网络时，在大多数域中可以合理地假设：**每个随机变量最多受到k个其他随机变量的直接影响**，其中k是某个常数。（可以使用不完全的联合分布）

**13.2.1贝叶斯网络中的条件独立性关系**

性质： 给定父节点，每个变量条件独立于它的非子孙（descendant）节点。

给定一个变量的父节点、子节点和子节点的父节点，即给定它的马尔可夫毯（Markov blanket），其条件独立于网络中所有其他节点。

**马尔可夫毯的构成与独立性基础**

马尔可夫毯 MB(Xi) 由以下三部分组成

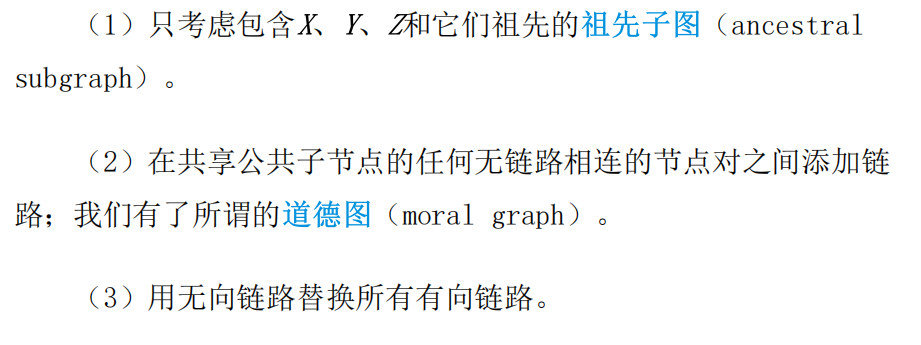
​​父节点​​（Pa(Xi)）：直接影响 Xi的变量。

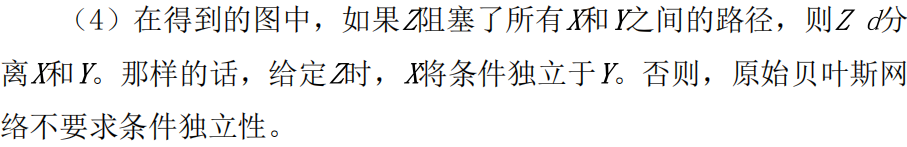
​​子节点​​（Ch(Xi)）：直接受 Xi影响的变量。

​​配偶节点​​（Sp(Xi)）：与 Xi共同影响某个子节点的其他父节点。

这一结构确保了马尔可夫毯覆盖了所有与 Xi存在直接因果或依赖关系的变量，从而形成对 Xi的“信息屏障”。

给定第三个集合Z，节点集合X是否条件独立于另一个集合Y：过检查贝叶斯网络的集合Z是否d分离（d-separate）集合X和Y：





13.2.2条件分布的高效表示

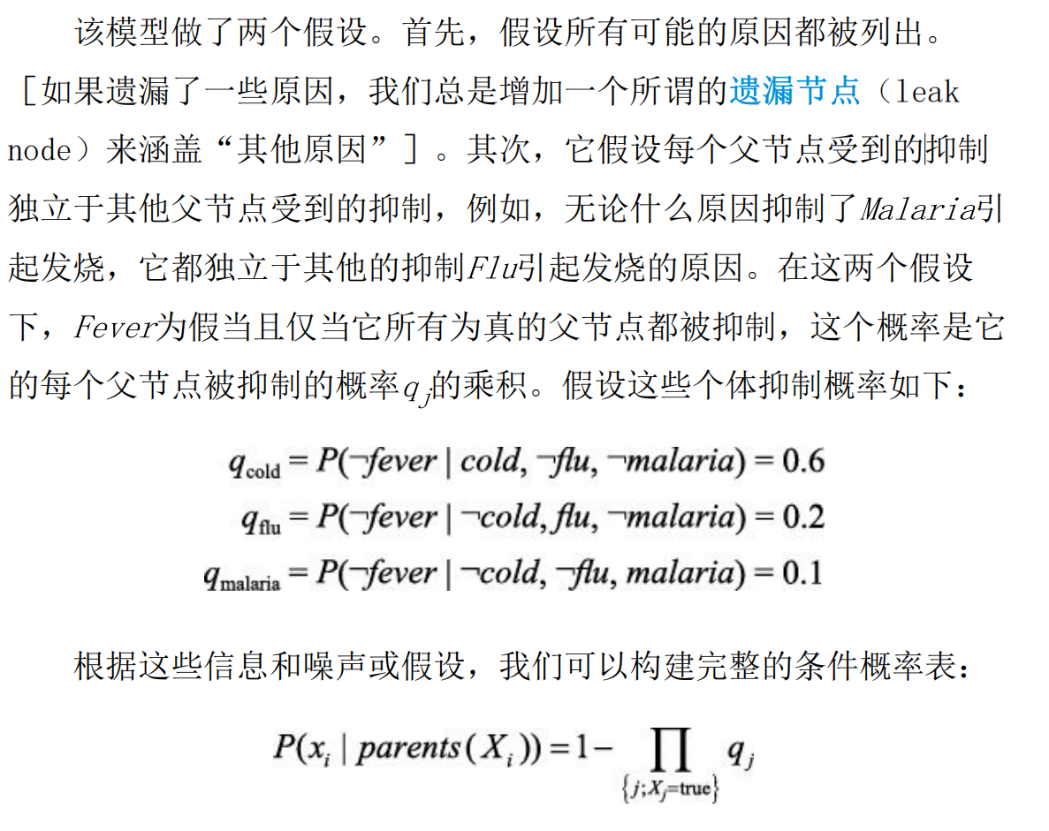
**通过符合某种标准模式的正则分布（canonical distribution）来描述父节点和子节点之间的关系。**

**确定性节点**（deterministic node）：其值完全由其父节点的值指定，没有任何不确定性。

**特定于上下文的独立性**（context-specific independence，CSI）：如果给定其他变量的某

些值，一个变量条件独立于它的一些父节点，则这个条件分布存在CSI。

**不确定关系**：通常可以利用所谓的噪声（noisy）逻辑关系来刻画。

其中一个典型的例子是噪声或（noisy-OR）关系，它是逻辑或的推广。在命题逻辑中，当且仅当Cold、Flu或Malaria为真时，我们说Fever为真。**噪声或模型为每个父节点导致子节点为真的能力引入了不确定性，父节点与子节点之间的因果关系可能会被抑制，因此病人可能感冒，但可能不发烧。**

13.2.3连续变量的贝叶斯网络

方法1：离散化

方法2：使用一族标准的概率密度函数来定义连续变量，如高斯分布N(x; μ，σ2)

方法3：非参数（nonparametric）表示，用一组实例隐式地定义条件分布，其中每个实例都包含父变量和子变量的特定值。

同时具有离散变量和连续变量的网络称为**混合贝叶斯网络**（hybrid Bayesian network）。

**13.3贝叶斯网络中的精确推断**

贝叶斯网络中的推断是指在给定一些变量的观测值(证据)的情况下，计算其他变量的概率分布。精确推断方法能够得到准确的结果，但计算复杂度可能很高。

推断的基本形式

贝叶斯网络中的推断主要有以下几种形式：

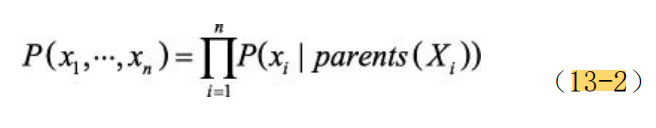
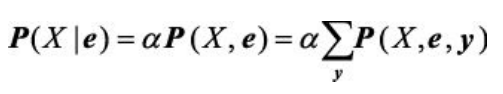
1. 后验推断：给定证据E，计算假设H的后验概率P(H|E)

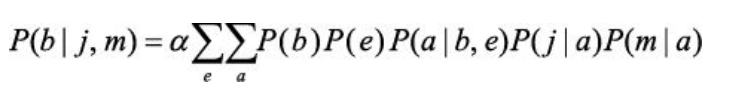
2. 预测推断：给定某些变量的值，计算其他变量的边缘概率

3. 最可能解释：找到与证据一致的最可能假设集合

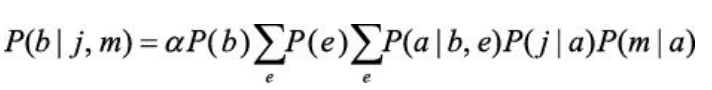
**13.3.1 通过枚举进行推断**

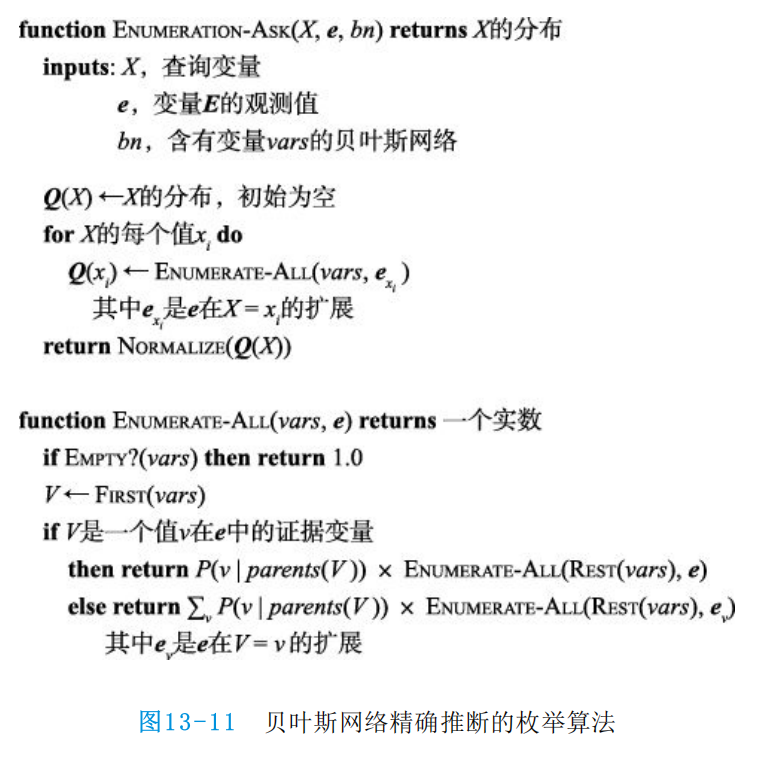
给定证据e，查询P(X|e)



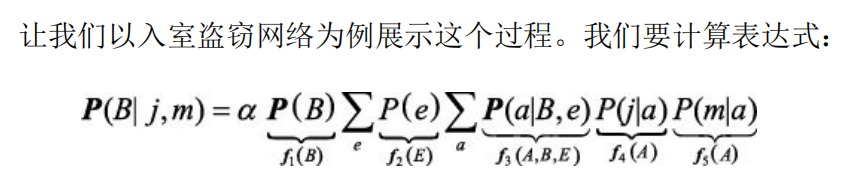
由（13-2）项*P*(X,e,y)可以写作网络中条件概率的乘积

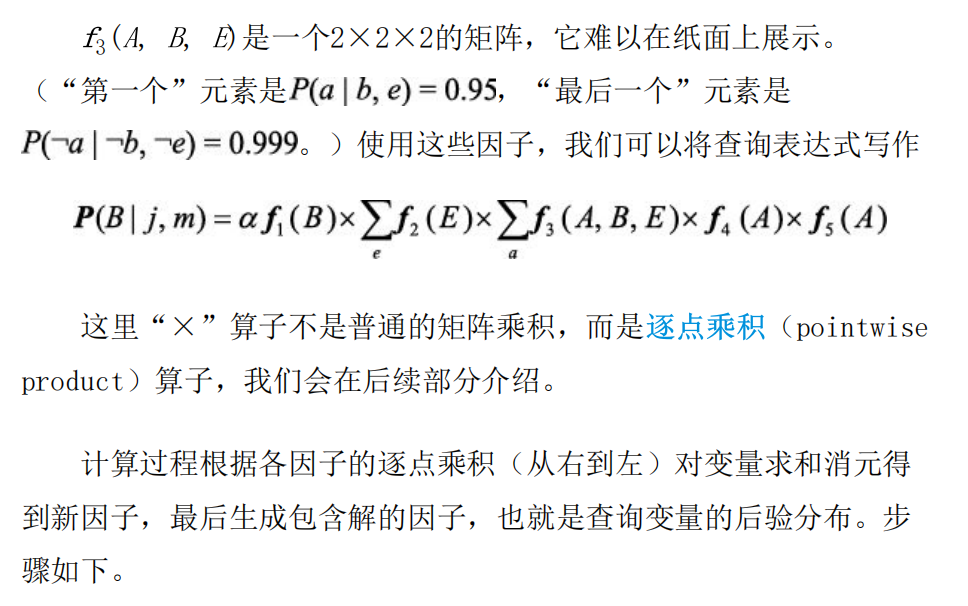
如：

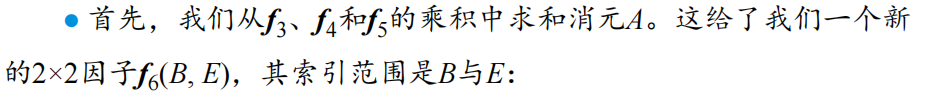


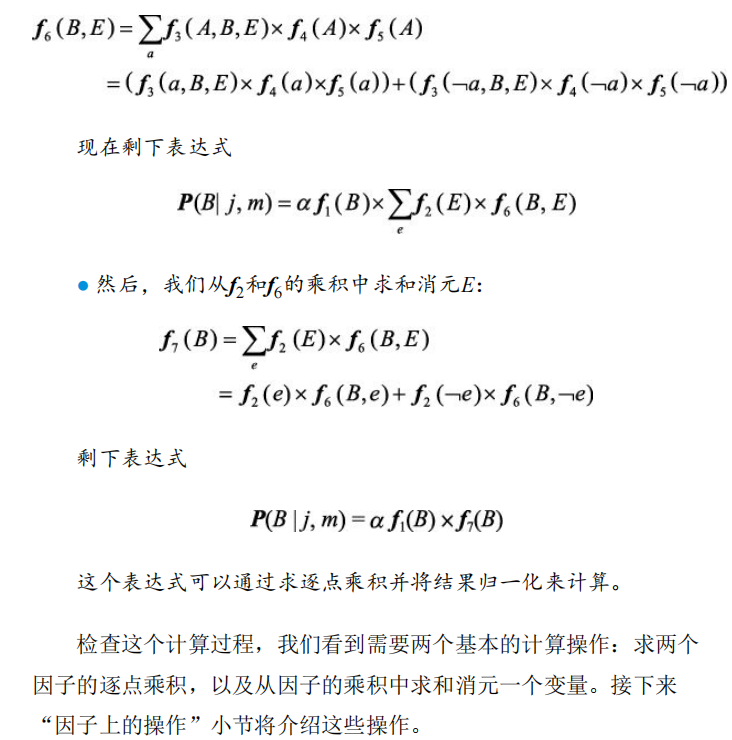


**13.3.2 变量消元算法**









**1. 因子上的操作**

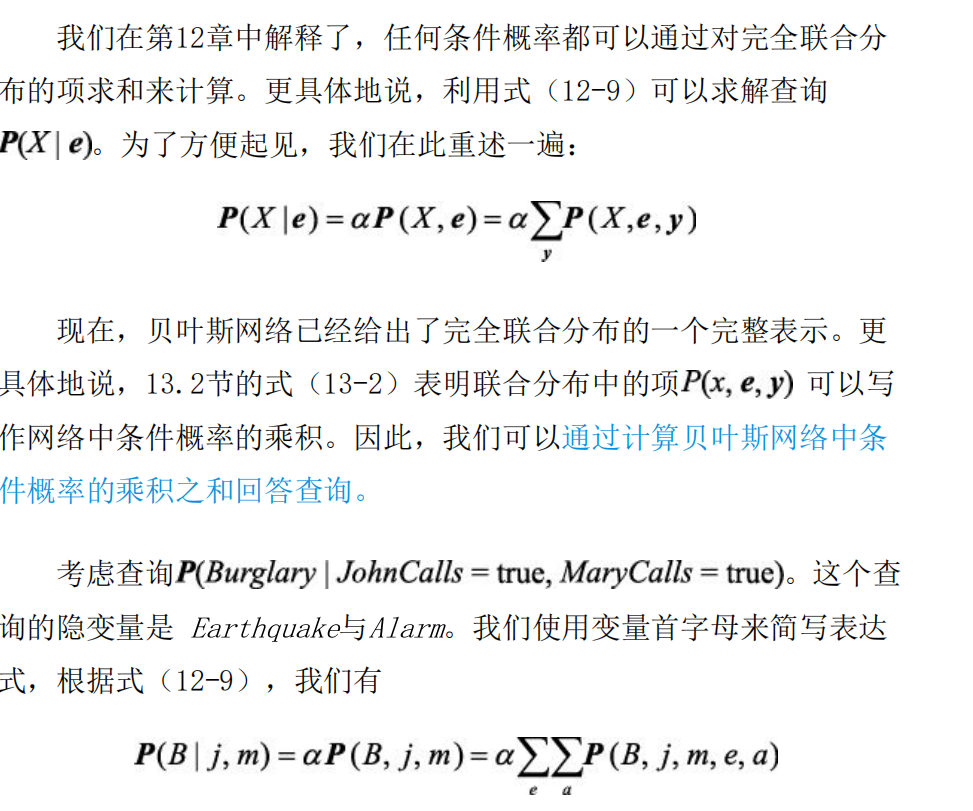
**2. 变量排序和变量相关性**

**13.4贝叶斯网络中的近似推理**

考虑一个查询变量，可以很容易地扩展到 具有多个变量的查询。

**13.3.1通过枚举进行推断**



**13.3.2变量消元算法**

**13.5因果网络**