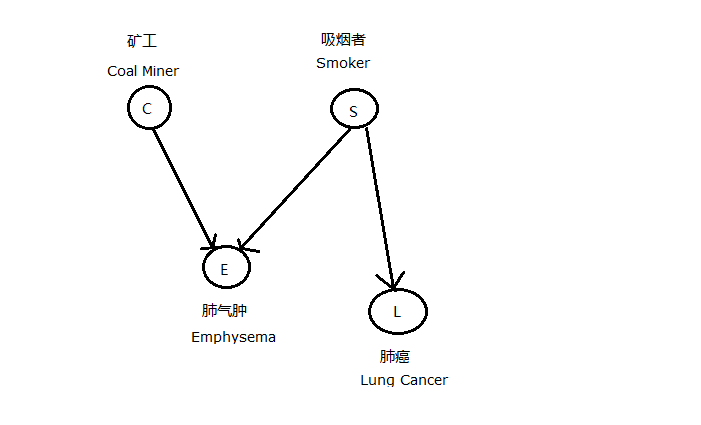
贝叶斯网络  
　　贝叶斯网络是一系列变量的联合概率分布的图形表示。  
  
　　　　一般包含两个部分，一个就是贝叶斯网络结构图，这是一个有向无环图（DAG），其中图中的每个节点代表相应的变量，节点之间的连接关系代表了贝叶斯网络的条件独立语义。另一部分，就是节点和节点之间的**条件概率**表（CPT），也就是一系列的概率值。如果一个贝叶斯网络提供了足够的**条件概率**值，足以计算任何给定的联合概率，我们就称，它是可计算的，即可推理的。  
3.5.1 贝叶斯网络基础  
　　首先从一个具体的实例（医疗诊断的例子）来说明贝叶斯网络的构造。  
　　假设：  
　　命题S(moker)：该患者是一个吸烟者  
　　命题C(oal Miner)：该患者是一个煤矿矿井工人  
　　命题L(ung Cancer)：他患了肺癌  
　　命题E(mphysema)：他患了肺气肿  
　　命题S对命题L和命题E有因果影响，而C对E也有因果影响。  
　　命题之间的关系可以描绘成如右图所示的因果关系网。  
　　因此，贝叶斯网有时也叫因果网，因为可以将连接结点的弧认为是表达了直接的因果关系。  
  
图3-5 贝叶斯网络的实例  
  
图中表达了贝叶斯网的两个要素：其一为贝叶斯网的结构，也就是各节点的继承关系，其二就是**条件概率**表CPT。若一个贝叶斯网可计算，则这两个条件缺一不可。  
　　贝叶斯网由一个有向无环图（DAG）及描述顶点之间的概率表组成。其中每个顶点对应一个随机变量。这个图表达了分布的一系列有条件独立属性：在给定了父亲节点的状态后，每个变量与它在图中的非继承节点在概率上是独立的。该图抓住了概率分布的定性结构，并被开发来做高效推理和决策。  
　　贝叶斯网络能表示任意概率分布的同时，它们为这些能用简单结构表示的分布提供了可计算优势。   
　　假设对于顶点xi，其双亲节点集为Pai，每个变量xi的**条件概率**P(xi|Pai)。 则顶点集合X={x1,x2,…,xn}的联合概率分布可如下计算：  
。  
  
　　双亲结点。该结点得上一代结点。  
　　该等式暗示了早先给定的图结构有条件独立语义。它说明贝叶斯网络所表示的联合分布作为一些单独的局部交互作用模型的结果具有因式分解的表示形式。  
从贝叶斯网的实例图中，我们不仅看到一个表示因果关系的结点图，还看到了贝叶斯网中的每个变量的**条件概率**表（CPT）。因此一个完整的随机变量集合的概率的完整说明不仅包含这些变量的贝叶斯网，还包含网中变量的**条件概率**表。  
  
　　图例中的联合概率密度：  
　　P(S,C,L,E)=P(E|S,C)\*P(L|S)\*P(C)\*P(S)  
　　推导过程：P(S,C,L,E)=P(E|S,C,L)\*P(L|S,C)\*P(C|S)\*P(S)（贝叶斯定理）  
　　＝P(E|S,C)\*P(L|S)\*P(C)\*P(S)  
　　即：P(E|S,C,L) ＝ P(E|S,C), E与L无关  
　　　　P(L|S,C)= P(L|S)　　　　L与C无关  
　　　　P(C|S)=P(C) 　　　　　　C与S无关  
　　以上三条等式的正确性，可以从贝叶斯网的条件独立属性推出：每个变量与它在图中的非继承节点在概率上是独立的。  
　　相比原始的数学公式：  
　　P(S,C,L,E)=P(E|S,C,L)\*P(L|S,C)\*P(C|S)\*P(S)  
　　推导过程：  
　　由贝叶斯定理， P(S,C,L,E)=P(E|S,C,L)\*P(S,C,L)  
　　再由贝叶斯定理 P(S,C,L)= P(L|S,C)\* P(S,C)  
　　同样， P(S,C)=P(C|S)\*P(S)  
　　以上几个等式相乘即得原式。   
　　显然，简化后的公式更加简单明了，计算复杂度低很多。如果原贝叶斯网中的条件独立语义数量较多，这种减少更加明显。  
　　贝叶斯网络是一系列变量的联合概率分布的图形表示。这种表示法最早被用来对专家的不确定知识编码，今天它们在现代专家系统、诊断引擎和决策支持系统中发挥了关键作用。贝叶斯网络的一个被经常提起的优点是它们具有形式的概率语义并且能作为存在于人类头脑中的知识结构的自然映像。这有助于知识在概率分布方面的编码和解释，使基于概率的推理和最佳决策成为可能。  
3.5.2 贝叶斯网的推理模式  
　　在贝叶斯网中有三种重要的推理模式，因果推理（由上向下推理），诊断推理（自底向上推理）和辩解。  
3．5．2．1 因果推理  
　　让我们通过概述的实例来说明因果推理得过程。给定患者是一个吸烟者（S），计算他患肺气肿（E）的概率P(E|S)。S称作推理的证据，E叫询问结点。  
  
　　首先，我们寻找E的另一个父结点（C），并进行概率扩展  
　　　P(E|S)=P(E,C|S)+P(E,~C|S);  
　　即，吸烟的人得肺气肿的概率为吸烟得肺气肿又是矿工的人的概率与吸烟得肺气肿不是矿工的人的概率之和，也就是全概率公式。  
　　然后利用Bayes定理：  
　　　P(E|S)=P(E|C,S)\*P(C|S)+P(E|~C,S)\*P(~C|S);   
　　公式解释：P(E,C|S)＝P(E,C,S)/P(S)  
　　　　　　　　　　　＝P(E|C,S)\*P(C,S)/P(S)(贝叶斯定理)  
　　　　　　　　　　　＝P(E|C,S)\*P(C|S)(反向利用贝叶斯定理)  
　　同理可以得出P(E,~C|S)的推导过程。  
　　需要寻找该表达式的双亲结点的**条件概率**，重新表达联合概率（指P(E,C|S)，P(E,~C|S)）。  
　　在图中，C和S并没有双亲关系，符合条件独立条件：  
　　　P(C|S)=P(C),   
　　　P(~C|S) = P(~C),  
　　由此可得：  
　　　P(E|S) = P(E|S,C)\*P(C)+P(E|~C,S)\*P(~C)  
　　如果采用概述中的例题数据，则有P(E|S)＝0.9\*0.3+0.3\*(1-0.3)=0.48  
　　从这个例子中，不难得出这种推理的主要操作：  
　　1） 按照给定证据的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**。  
　　2） 回到以所有双亲为条件的概率，重新表达这个联合概率。  
　　3） 直到所有的概率值可从CPT表中得到，推理完成。  
3．5．2．2 诊断推理  
　　同样以概述中的例题为例，我们计算"不得肺气肿的不是矿工"的概率P(~C|~E),即在贝叶斯网中，从一个子结点计算父结点的**条件概率**。也即从结果推测一个起因，这类推理叫做诊断推理。使用Bayes公式就可以把这种推理转换成因果推理。  
  
　　　P(~C|~E)＝P(~E|~C)\*P(~C)/P(~E)，  
　　从因果推理可知  
　　　P(~E|~C) = P(~E,S|~C)+P(~E,~S|~C)  
　　　　　 　　= P(~E|S,~C)\*P(S)+P(~E|~S,~C)\*P(~S)  
　　　　　 　　= (1-0.3)\*0.4+(1-0.10)\*(1-0.4)=0.82;  
　　由此得：  
　　　P(~C|~E)＝P(~E|~C)\*P(~C)/ P(~E)(贝叶斯公式)  
　　　　　　　＝0.82\*(1-0.3)/ P(~E)  
　　　　　　　＝0.574/ P(~E)  
　　同样的，  
　　　P(C|~E) ＝P(~E|C)\* P(C)/ P(~E)  
　　　　　　　＝0.34\*0.3/ P(~E)  
　　　　　　　＝0.102 /P(~E)  
　　由于全概率公式：  
　　　P(~C|~E)+P(C|~E)＝1  
　　代入可得  
　　　P(~E)=0.676  
　　所以， P(~C|~E)＝0.849  
　　这种推理方式主要利用Bayes规则转换成因果推理。  
3．5．2．3 辩解  
　　如果我们的证据仅仅是～E（不是肺气肿），象上述那样，我们可以计算～C患者不是煤矿工人的概率。但是如果也给定～S（患者不是吸烟者），那么～C也应该变得不确定。这种情况下，我们说～S解释～E，使～C变得不确定。这类推理使用嵌入在一个诊断推理中的因果推理。  
  
  
作为思考题，读者可以沿着这个思路计算上式。在这个过程中，贝叶斯规则的使用，是辩解过程中一个重要的步骤。   
3.5.3 D分离  
　　在本节最开始的贝叶斯网图中，有三个这样的结点：S，L，E。从直观来说，L的知识（结果）会影响S的知识（起因），S会影响E的知识（另一个结果）。因此，在计算推理时必须考虑的相关因素非常多，大大影响了算法的计算复杂度，甚至可能影响算法的可实现性。但是如果给定原因S，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，那么在计算S和L的关系时就不用过多地考虑E，将会大大减少计算复杂度。这种情况下，我们称S能D分离L和E。D分离是一种寻找条件独立的有效方法。  
  
　　如下图，对于给定的结点集ε，如果对贝叶斯网中的结点Vi和Vj之间的每个无向路径，在路径上有某个结点Vb，如果有属性：   
　　1）Vb在ε中，且路径上的两条弧都以Vb为尾（即弧在Vb处开始（出发））  
　　2）Vb在ε中，路径上的一条弧以Vb为头，一条以Vb为尾  
　　3）Vb和它的任何后继都不在ε中，路径上的两条弧都以Vb为头（即弧在Vb处结束）  
　　则称Vi和Vj 被Vb结点阻塞。  
　　结论：如果Vi和Vj被证据集合ε中的任意结点阻塞，则称Vi和Vj是被ε集合D分离，结点Vi和Vj条件独立于给定的证据集合ε，即  
　　　　P(Vi|Vj,ε) ＝P(Vi|ε)  
　　　　P(Vj|Vi,ε) ＝P(Vj|ε)   
　　表示为：I(Vi,Vj|ε) 或I(Vj,Vi|ε)   
　　无向路径：DAG图是有向图，所以其中的路径也应该是有向路径，这里所指的无向路径是不考虑DAG图中的方向性时的路径。  
　　条件独立：如具有以上三个属性之一，就说结点Vi和Vj条件独立于给定的结点集ε。  
　　阻塞：给定证据集合ε，当上述条件中的任何一个满足时，就说Vb阻塞相应的那条路径。  
　　D分离：如果Vi和Vj之间所有的路径被阻塞，就叫证据集合ε可以D分离Vi和Vj  
　　注意：在论及路径时，是不考虑方向的；在论及"头"和"尾"时，则必须考虑弧的方向。"头"的含义是箭头方向（有向弧）的终止点，"尾"的含义是箭头方向（有向弧）的起始点。  
回到最开始的医疗诊断实例：为简单起见，选择证据集合ε为单个结点集合。  
　　对于给定的结点S，结点E阻塞了结点C和结点L之间的路径，因此C和L是条件独立的，有I（C，L|S）成立。  
　　而对于给定结点E，S和L之间找不到阻塞结点。因此，S和L不是条件独立的。  
  
　 即使使用了D分离，一般地讲，在贝叶斯网中，概率推理仍是NP难题。然而，有些简化能在一个叫Polytree的重要网络分类中使用。一个Polytree网是一个DAG，在该DAG的任意两个结点间，顺着弧的每一个方向只有一条路径。如图就是一个典型的Polytree。  
图3-7 Polytree   
  
  
　　D分离的实质就是寻找贝叶斯网中的条件独立语义，以简化推理计算。   
　　总结  
　　本节就Bayes网络的基本问题进行了阐述，着重点在推理计算上。其本质就是通过各种方法寻找网络中的条件独立性，达到减少计算量和复杂性的目的。这些都只是粗浅的描述，进一步的学习，请参考相应的参考书的"http://bbs.matwav.com/images/smilies/default/titter.gifolytree的概率推理"和"Bayes网的学习和动作"等章节，其中有很详细的阐述。