**前向算法 (Forward Algorithm)**

前向算法是因该算法计算的时候使用了向前递归的方式计算最终概率而得名。

向前递归

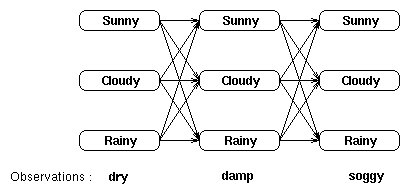
后向算法的理念与前向算法相同，只是向后递归而已

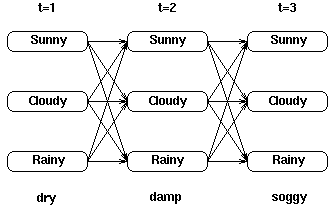
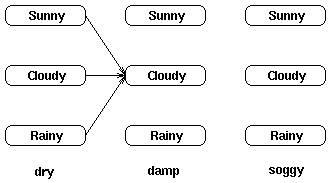
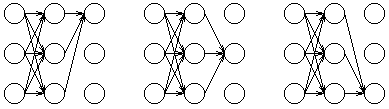
本文只介绍前向算法

该算法相比穷举时间复杂度大大降低

穷举搜索时间复杂度与观察序列的长度成指数级

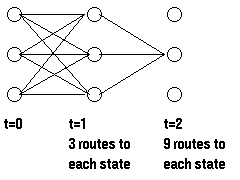
前向算法的时间复杂度与观察序列的长度成线性

 1.穷举搜索（ Exhaustive search for solution）  
　　给定隐马尔科夫模型，也就是在模型参数（pi, A, B)已知的情况下，我们想找到观察序列的概率。还是考虑天气这个例子，我们有一个用来描述天气及与它密切相关的海藻湿度状态的隐马尔科夫模型(HMM)，另外我们还有一个海藻的湿度状态观察序列。假设连续3天海藻湿度的观察结果是（干燥、湿润、湿透）——而这三天每一天都可能是晴天、多云或下雨，对于观察序列以及隐藏的状态，可以将其视为网格：  
  
　　网格中的每一列都显示了可能的的天气状态，并且每一列中的每个状态都与相邻列中的每一个状态相连。而其状态间的转移都由状态转移矩阵提供一个概率。在每一列下面都是某个时间点上的观察状态，给定任一个隐藏状态所得到的观察状态的概率由混淆矩阵提供。  
　　可以看出，一种计算观察序列概率的方法是找到每一个可能的隐藏状态，并且将这些隐藏状态下的观察序列概率相加。对于上面那个（天气）例子，将有3^3 = 27种不同的天气序列可能性，因此，观察序列的概率是：  
　　Pr(dry,damp,soggy | HMM) = Pr(dry,damp,soggy | sunny,sunny,sunny) + Pr(dry,damp,soggy | sunny,sunny ,cloudy) + Pr(dry,damp,soggy | sunny,sunny ,rainy) + . . . . Pr(dry,damp,soggy | rainy,rainy ,rainy)  
　　用这种方式计算观察序列概率极为昂贵，特别对于大的模型或较长的序列，因此我们可以利用这些概率的时间不变性来减少问题的复杂度。

2.使用递归降低问题复杂度  
　　给定一个隐马尔科夫模型（HMM），我们将考虑递归地计算一个观察序列的概率。我们首先定义局部概率（partial probability）,它是到达网格中的某个中间状态时的概率。然后，我们将介绍如何在t=1和t=n(>1)时计算这些局部概率。  
　　假设一个T-长观察序列是：  
　　　　　t-long 序列  
　　  
　2a.局部概率(alpha’s)  
　　考虑下面这个网格，它显示的是天气状态及对于观察序列干燥，湿润及湿透的一阶状态转移情况：  
　　　  
　　我们可以将计算到达网格中某个中间状态的概率作为所有到达这个状态的可能路径的概率求和问题。  
　　例如，t=2时位于“多云”状态的局部概率通过如下路径计算得出：  
　　　  
　　我们定义t时刻位于状态j的局部概率为at(j)——这个局部概率计算如下：  
　　alphat ( j )= Pr( 观察状态 | 隐藏状态j ) x Pr(t时刻所有指向j状态的路径）  
　　对于最后的观察状态，其局部概率包括了通过所有可能的路径到达这些状态的概率——例如，对于上述网格，最终的局部概率通过如下路径计算得出：  
　　　  
　　由此可见，对于这些最终局部概率求和等价于对于网格中所有可能的路径概率求和，也就求出了给定隐马尔科夫模型(HMM)后的观察序列概率。  
　　第3节给出了一个计算这些概率的动态示例

计算观察序列的概率（Finding the probability of an observed sequence）

2b.计算t=1时的局部概率alpha’s  
　　我们按如下公式计算局部概率：  
　　alphat ( j )= Pr( 观察状态 | 隐藏状态j ) x Pr(t时刻所有指向j状态的路径）  
　　特别当t=1时，没有任何指向当前状态的路径。故t=1时位于当前状态的概率是初始概率，即Pr(state|t=1)=P(state)，因此，t=1时的局部概率等于当前状态的初始概率乘以相关的观察概率：  
　　　　　　　　　5.2_2  
　　所以初始时刻状态j的局部概率依赖于此状态的初始概率及相应时刻我们所见的观察概率。

2c.计算t>1时的局部概率alpha’s  
　　我们再次回顾局部概率的计算公式如下：  
　　alphat ( j )= Pr( 观察状态 | 隐藏状态j ) x Pr(t时刻所有指向j状态的路径）  
　　我们可以假设（递归地），乘号左边项“Pr( 观察状态 | 隐藏状态j )”已经有了，现在考虑其右边项“Pr(t时刻所有指向j状态的路径）”。  
　　为了计算到达某个状态的所有路径的概率，我们可以计算到达此状态的每条路径的概率并对它们求和，例如：  
　　　　　　  
　　计算alpha所需要的路径数目随着观察序列的增加而指数级递增，但是t-1时刻alpha’s给出了所有到达此状态的前一路径概率，因此，我们可以通过t-1时刻的局部概率定义t时刻的alpha’s，即：  
　　　　　5.1.2.3_1  
　　故我们所计算的这个概率等于相应的观察概率（亦即，t+1时在状态j所观察到的符号的概率）与该时刻到达此状态的概率总和——这来自于上一步每一个局部概率的计算结果与相应的状态转移概率乘积后再相加——的乘积。  
　　注意我们已经有了一个仅利用t时刻局部概率计算t+1时刻局部概率的表达式。  
　　现在我们就可以递归地计算给定隐马尔科夫模型(HMM)后一个观察序列的概率了——即通过t=1时刻的局部概率alpha’s计算t=2时刻的alpha’s，通过t=2时刻的alpha’s计算t=3时刻的alpha’s等等直到t=T。给定隐马尔科夫模型(HMM)的观察序列的概率就等于t=T时刻的局部概率之和。

2d.降低计算复杂度  
　　我们可以比较通过穷举搜索（评估）和通过递归前向算法计算观察序列概率的时间复杂度。  
　　我们有一个长度为T的观察序列O以及一个含有n个隐藏状态的隐马尔科夫模型l=(pi,A,B)。  
　　穷举搜索将包括计算所有可能的序列：  
　　　5.1.2.4_1  
　　公式  
　　　　5.1.2.4_2  
　　对我们所观察到的概率求和——注意其复杂度与T成指数级关系。相反的，使用前向算法我们可以利用上一步计算的信息，相应地，其时间复杂度与T成线性关系。  
**注：穷举搜索的时间复杂度是2TN^T，前向算法的时间复杂度是N^2T，其中T指的是观察序列长度，N指的是隐藏状态数目。**

3.总结  
　　我们的目标是计算给定隐马尔科夫模型HMM下的观察序列的概率——Pr(observations |lamda)。  
　　我们首先通过计算局部概率（alpha’s）降低计算整个概率的复杂度，局部概率表示的是t时刻到达某个状态s的概率。  
　　t=1时，可以利用初始概率(来自于P向量）和观察概率Pr(observation|state)（来自于混淆矩阵）计算局部概率；而t>1时的局部概率可以利用t-时的局部概率计算。  
　　因此，这个问题是递归定义的，观察序列的概率就是通过依次计算t=1,2,…,T时的局部概率，并且对于t=T时所有局部概率alpha’s相加得到的。  
　　注意，用这种方式计算观察序列概率的时间复杂度远远小于计算所有序列的概率并对其相加（穷举搜索）的时间复杂度。