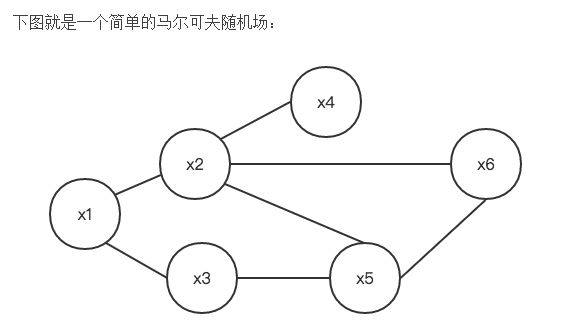
1 CRF 卡

今天主要了解了概率无向图模型，马尔可夫随机场,势函数和团,极大团,分离, 马尔可夫性, 局部马尔可夫性, 成对马尔可夫性。 条件随机场，线性链CRF等。

概率无向图模型到线性链条件随机场

### 概率无向图

****概率无向图模型:****是一个可以用无向图表示的联合概率分布[又叫: ****马尔可夫随机场****]



#### 势函数和团

****势函数：****是定义在变量子集上的非负实函数，用于定义概率分布函数。

****团:**** 图中节点的子集，其中任意两个节点之间都有边连接

****极大团:**** 一个团，其中加入任何一个其他的节点都不能再形成团.

马尔可夫随机场中： ****多个变量之间的联合概率分布可以基于团分解为多个势函数的乘积，每个势函数仅与一个团相关****

#### Hammersley-Clifford 定理

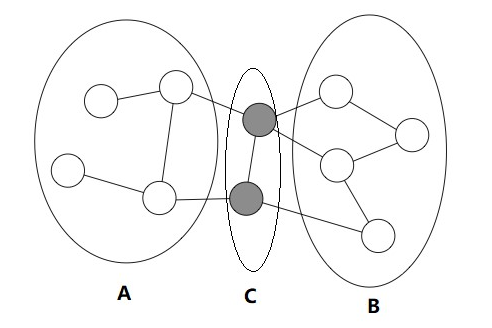
联合概率: P(X)=(1/Z\*)(∏Q∈CΨQ(XQ))

其中 ΨQ 为与团 Q 对应的势函数，用于对团 Q 中的变量关系进行建模,Z 为规范化因子.

##### ****两个概念****

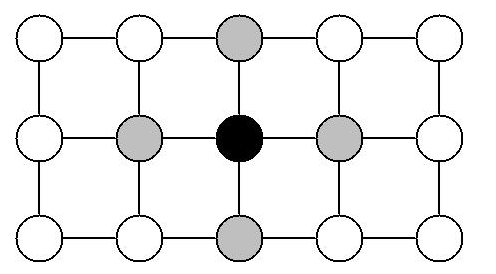
1. 分离（Separating）

AB 被C 分离, C叫分离集



1. 马尔可夫性（Markov Property）

当一个随机过程在给定当前状态及所有过去状态情况下，其未来状态的条件概率分布仅依赖于当前状态



全局马尔可夫性，又可以推导出两种性质

****局部马尔可夫性****

给定某变量的邻接变量，则该变量条件独立于其他变量。

****成对马尔可夫性****

给定所有其他变量，两个非连接变量条件独立

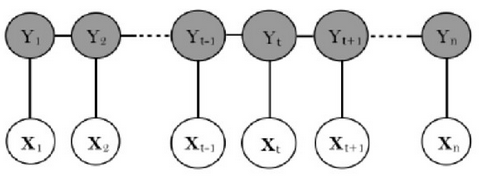
### 条件随机场

#### 无向图模型

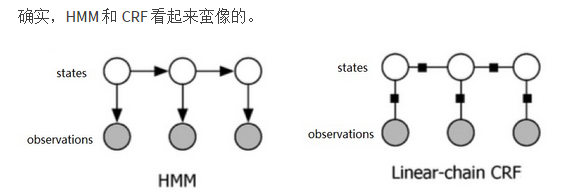
设 X 和 Y 是随机变量P(Y|X) 是给定 X 条件下 Y的条件概率分布。如果随机变量 Y 构成一个由无向图 G=<V,E> 表示的马尔可夫随机场，则称条件概率分布 P(Y|X) 为 CRF

### 线性链 CRF

X 为观测序列，Y 为状态序列。



### HMM VS 线性链 CRF



CRF——三个基本问题

### 线性链 CRF 的形式化表示

#### 一般形式

设 P(Y|X) 为线性链 CRF，在随机变量 X 取值为 x 的条件下，随机变量 Y 取值为 y 的条件概率具有如下形式：

P(y|x)=1Z(x)exp(∑i,kλktk(yi−1,yi,x,i)+∑i,lμlsl(yi,x,i))

#### 简化形式

如果我们将转移特征和状态特征及其权值用统一的符号表示，线性链 CRF 形式化的表示就会简单许多。

### CRF 的三个基本问题

##### ****1. 概率计算问题****

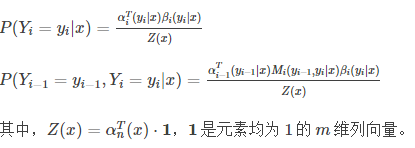
##### ****2. 预测问题****

##### ****3. 学习问题****

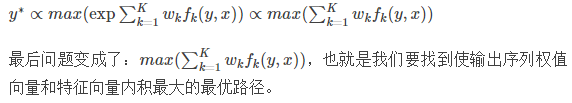
#### 对比 HMM 的三个基本问题， 两者****区别较大****的是****概率计算问题，****但对于 HMM 而言，概率计算只需要观测序列即可，无须确定的状态序列，而最终计算出的结果，则是当前观测序列出现的可能性，CRF 则需要既有已知观测序列，又有已知状态序列，这才能够去计算概率。

#### 三个基本问题的解法

##### ****概率计算问题****



##### ****预测问题****



针对这一问题，可以和应对 HMM 的预测问题一样，采用****维特比算法****。

##### ****学习问题****

线性链 CRF 模型实际上是定义在序列数据上的对数线形模型，可以通过极大化训练数据的对数似然函数来求模型参数



学习方法则有极大似然估计和正则化的极大似然估计。

具体的优化实现算法有：改进的迭代尺度法 IIS、梯度下降法以及拟牛顿法。目前应用较广的 BFGS 算法，属于拟牛顿法