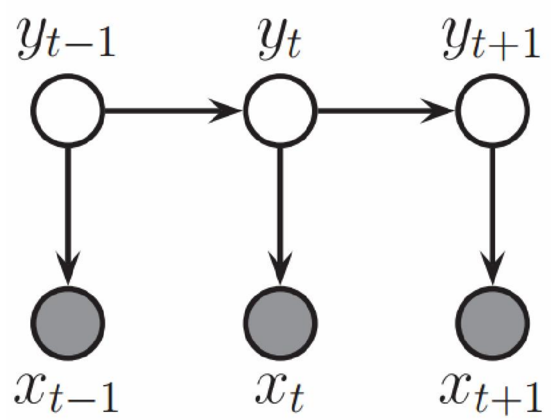
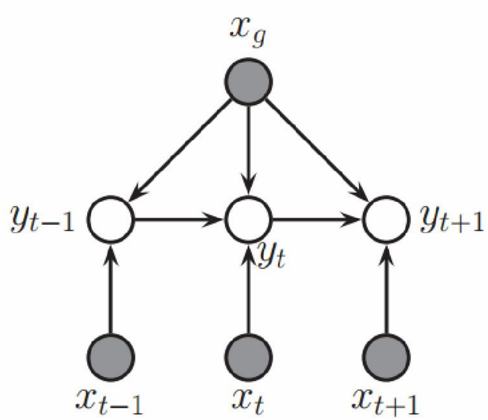
条件随机场

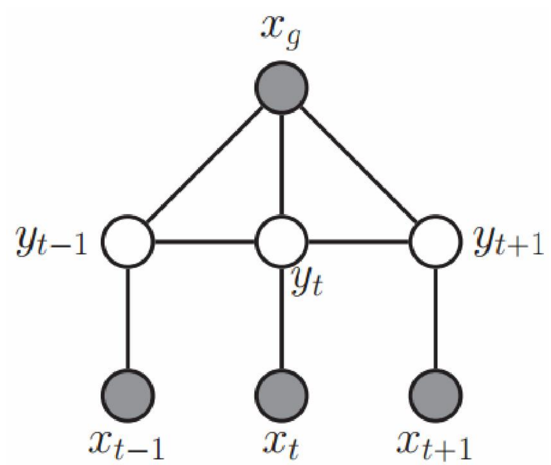
七月算法 邹博

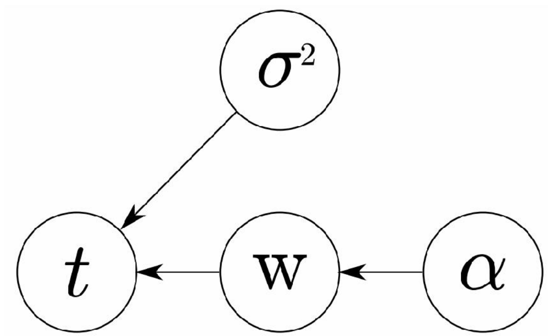
2015年12月13日

网络模型比较HMM/MEMM/CRF/RVM



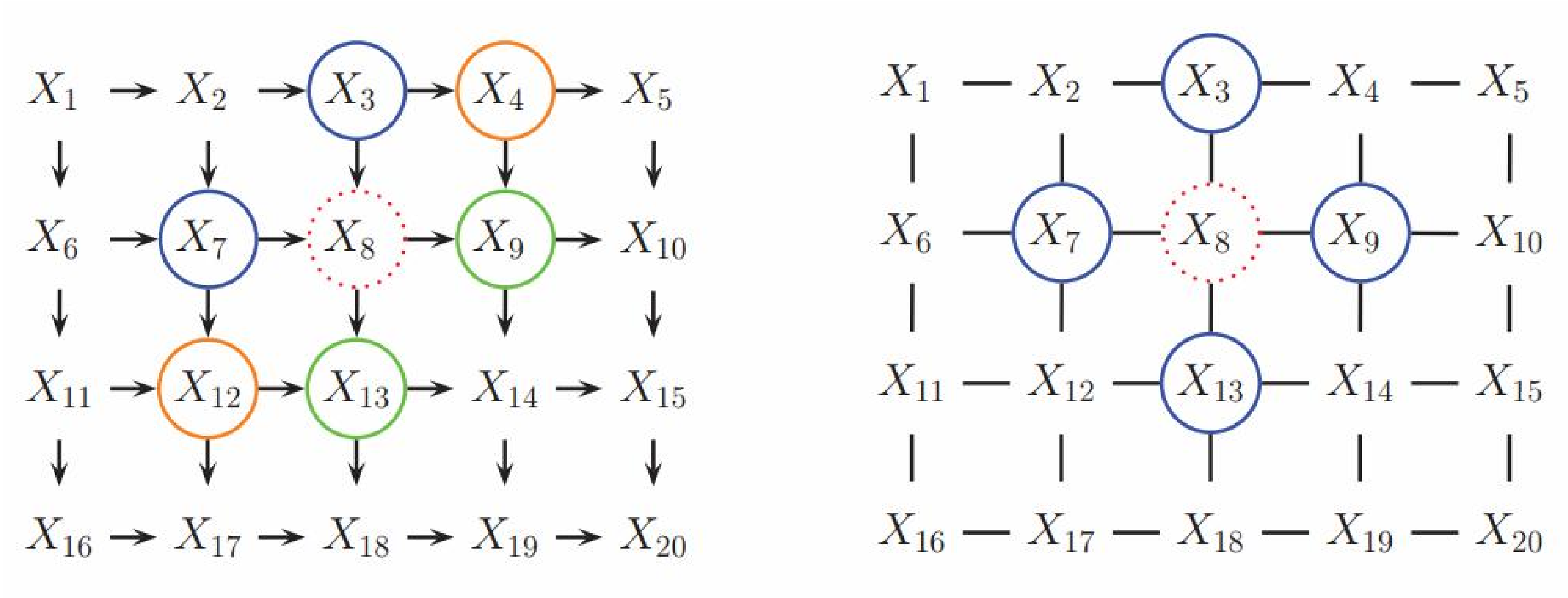






分析对图像像素间建立贝叶斯网络

- 考察X8的马尔科夫毯(Markov blanket)



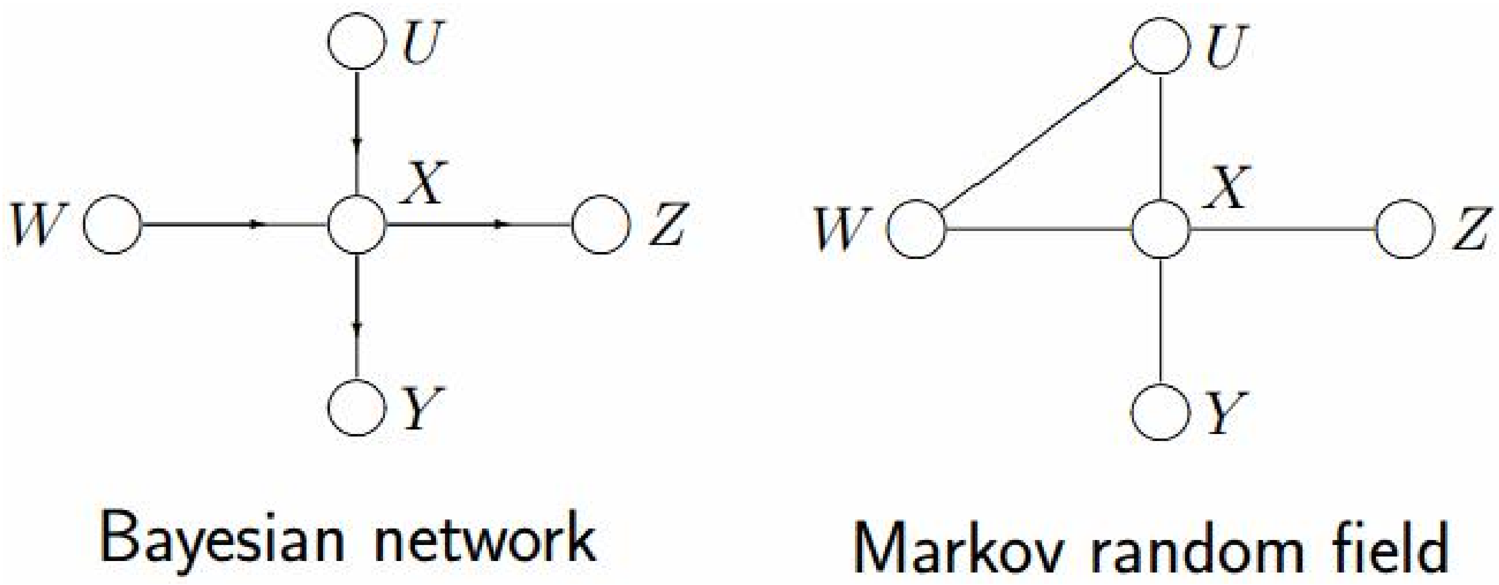
无向图模型

- 有向图模型，又称作贝叶斯网络(Directed Graphical Models, DGM, Bayesian Network)
  - 事实上，在有些情况下，强制对某些结点之间的边增加方向是不合适的。
- 使用没有方向的无向边，形成了无向图模型(Undirected Graphical Model,UGM), 又称马尔科夫随机场或马尔科夫网络(Markov Random Field, MRF or Markov network)
  - 注：概率有向图模型/概率无向图模型

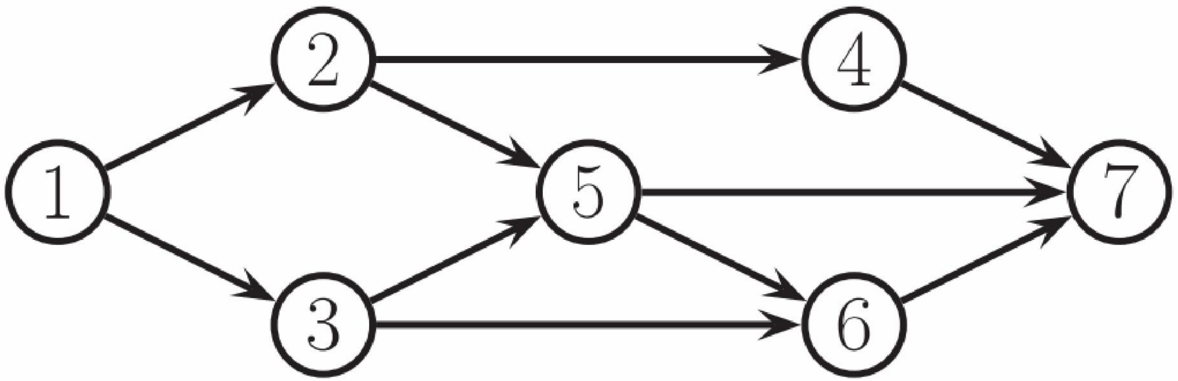
条件随机场

- 设$X=(X_1,X_2...X_n)$和$Y=(Y_1,Y_2...Y_m)$都是联合随机变量，若随机变量Y构成一个无向图G=(V,E)表示的马尔科夫随机场(MRF)，则条件概率分布P(Y|X)称为条件随机场(Conditional Random Field, CRF)
  - X称为输入变量、观测序列
  - Y称为输出序列、标记序列、状态序列
  - 大量文献将MRF和CRF混用，包括经典著作。
  - 一般而言，MRF是关于隐变量(状态变量、标记变量)的图模型，而给定观测变量后考察隐变量的条件概率，即为CRF。
  - 但这种混用，类似较真总理和周恩来的区别。
    - 有时候没必要区分的那么严格
    - 混用的原因：在计算P(Y|X)时需要将X也纳入MRF中一起考虑

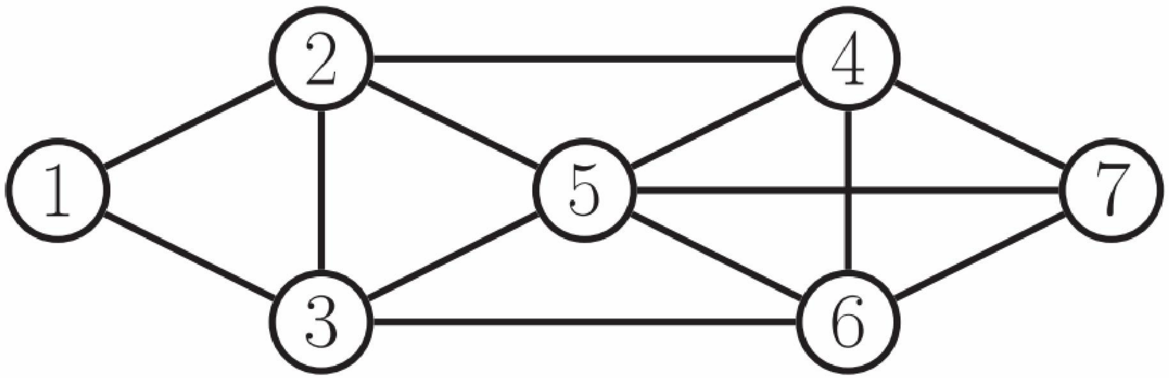
DGM转换成UGM



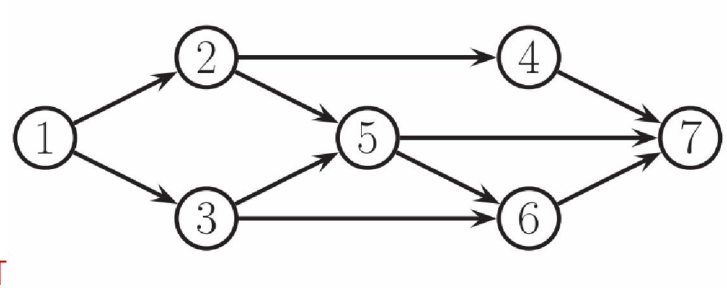
DGM转换成UGM



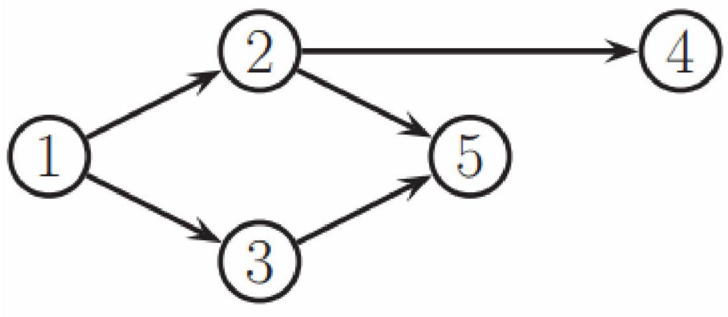
$$4 \perp 5 | 2$$

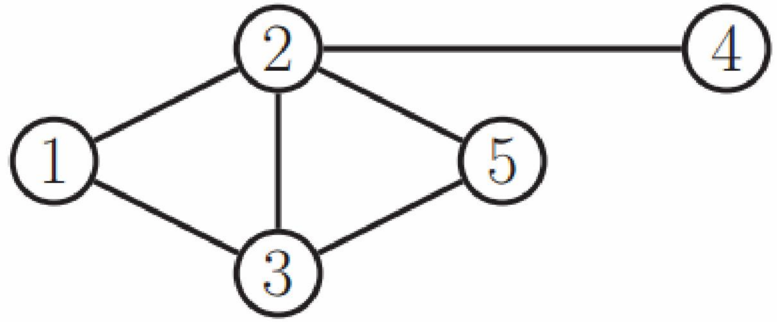


条件独立的破坏



- 靠考察是否有$A \perp B|C$，则计算U的祖先图(ancestral graph)：$U = A \cup B \cup C$

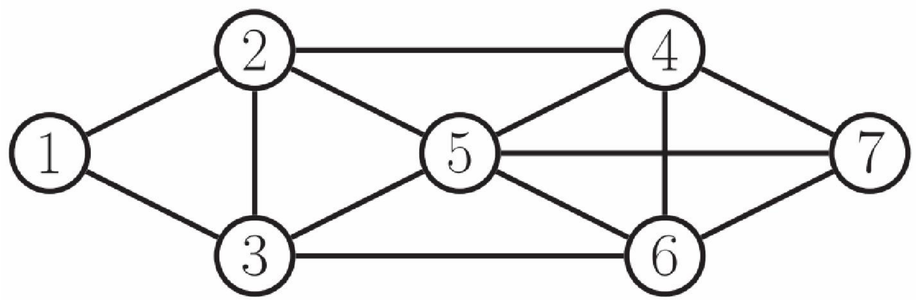




$$4 \perp 5|2$$

MRF的性质

- □ 成对马尔科夫性
  - ■ parewise Markov property
- □ 局部马尔科夫性
  - ■ local Markov property
- □ 全局马尔科夫性
  - ■ global Markov property



Pairwise $1 \perp 7|\text{rest}$

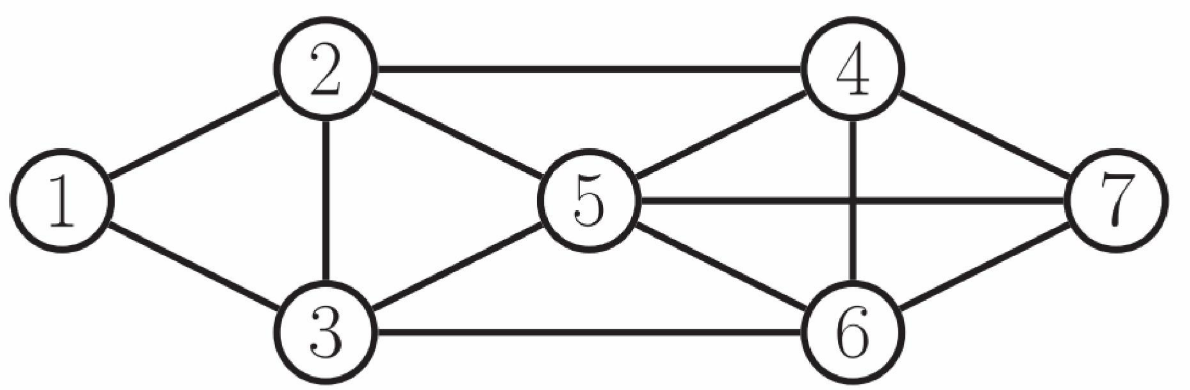
Local $1 \perp \text{rest}|2,3$

Global $1,2 \perp 6,7|3,4,5$

- □ 记号：随机变量$Y=(Y_1,Y_2\dots Y_m)$构成无向图$G=(V,E)$，结点(集)v对应的(联合)随机变量是Y_v。

成对马尔科夫性

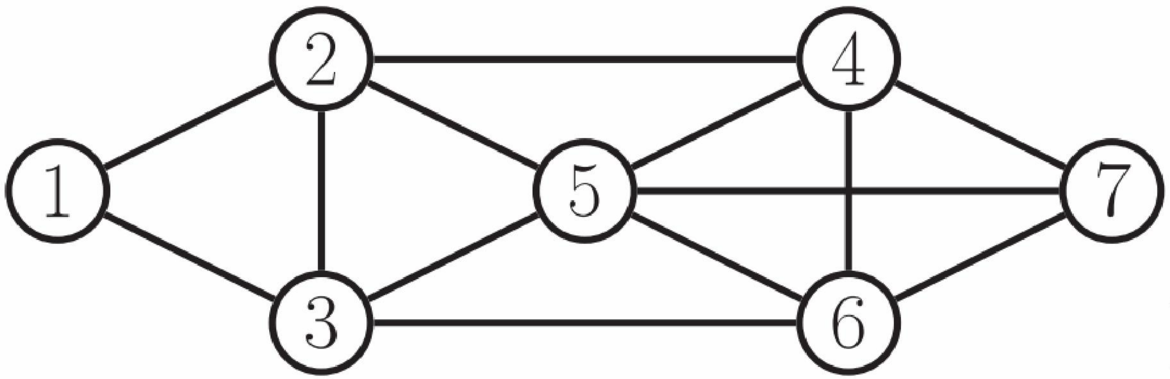
- 设u和v是无向图G中任意两个没有边直接连接的结点，G中其他结点的集合记做O；则在给定随机变量Yo的条件下，随机变量Yu和Yv条件独立。
- 即：$P(Y_u,Y_v|Y_o)= P(Y_u|Y_o)*P(Y_v|Y_o)$



$$1 \perp 7 | \text{rest}$$

局部马尔科夫性

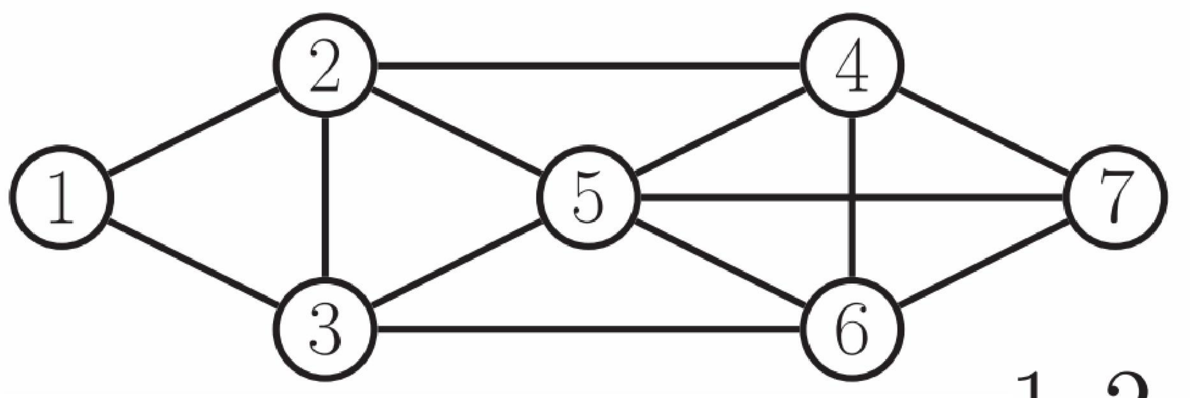
- 设v是无向图G中任意一个结点，W是与v有边相连的所有结点，G中其他结点记做O；则在给定随机变量Yw的条件下，随机变量Yv和Yo条件独立。
- 即：P(Yv,Yo|Yw)= P(Yv|Yw)* P(Yo|Yw)



$$1 \perp \text{rest} | 2, 3$$

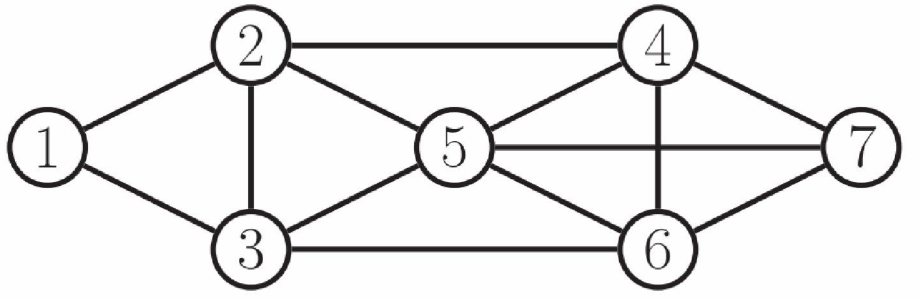
全局马尔科夫性

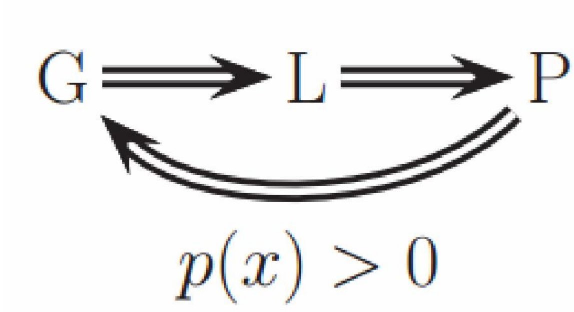
- 设结点集合A，B是在无向图G中被结点集合C分开的任意结点集合，则在给定随机变量Y_C的条件下，随机变量Y_A和Y_B条件独立。
- 即：$P(Y_A, Y_B | Y_C) = P(Y_A | Y_C) * P(Y_B | Y_C)$



$$1,2 \perp 6,7 | 3,4,5$$

三个性质的等价性





- 根据全局马尔科夫性，能够得到局部马尔科夫性；
- 根据局部马尔科夫性，能够得到成对马尔科夫性；
- 根据成对马尔科夫性，能够得到全局马尔科夫性；

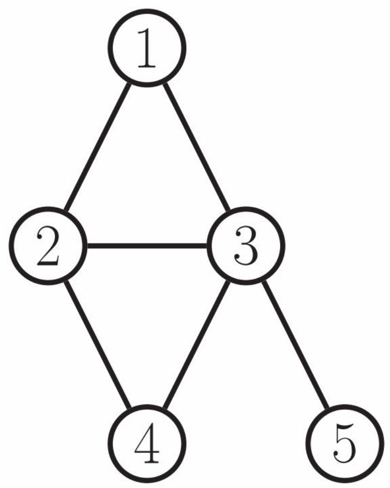
- 事实上，这个性质对MRF具有**决定性**作用：
  - 满足这三个性质(或其一)的无向图，称为MRF。

贝叶斯网络与条件随机场

- 在研究贝叶斯网络的过程中，重点考察了LDA模型、HMM模型，使用Markov模型进行了MCMC；
- 在条件随机场的研究学习中，将重点考察线性链MRF(LC-CRF,Linear Chain Conditional Random Field)

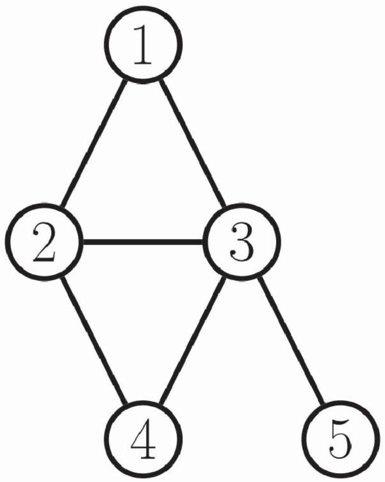
团和最大团

- 无向图G中的某个子图S，若S中任何两个结点均有边，则S称作G的团(Clique)。
  - 若C是G的一个团，并且不能再加入任何一个G的结点使其称为团，则C称作G的最大团(Maximal Clique)。
- 团：{1,2}, {1,3}, {2,3}, {2,4}, {3,4}, {3,5},{1,2,3}, {2,3,4}
- 最大团：{1,2,3}, {2,3,4}, {3,5}



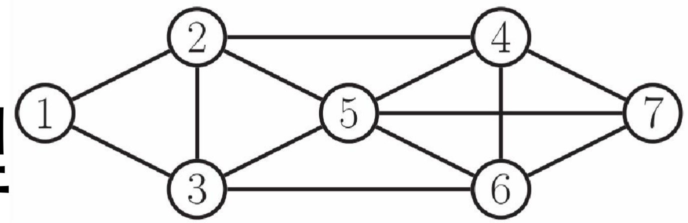
Hammersley-Clifford定理

- UGM的联合分布可以表示成最大团上的随机变量的函数的乘积的形式；这个操作叫做UGM的因子分解(Factorization)。



$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{1}{Z(\boldsymbol{\theta})}\psi_{123}(y_1, y_2, y_3)\psi_{234}(y_2, y_3, y_4)\psi_{35}(y_3, y_5)$$

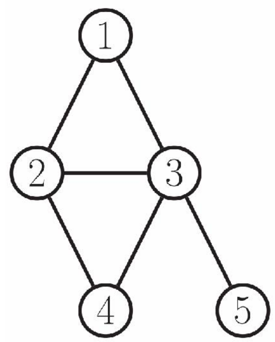
Hammersley-Clifford定理



- UGM的联合概率分布P(Y)可以表示成如下形式：

$$P(Y)=\frac{1}{Z}\prod_{c}\psi_{c}(Y_{c})$$

$$Z=\sum_{Y}\prod_{c}\psi_{c}(Y_{c})$$



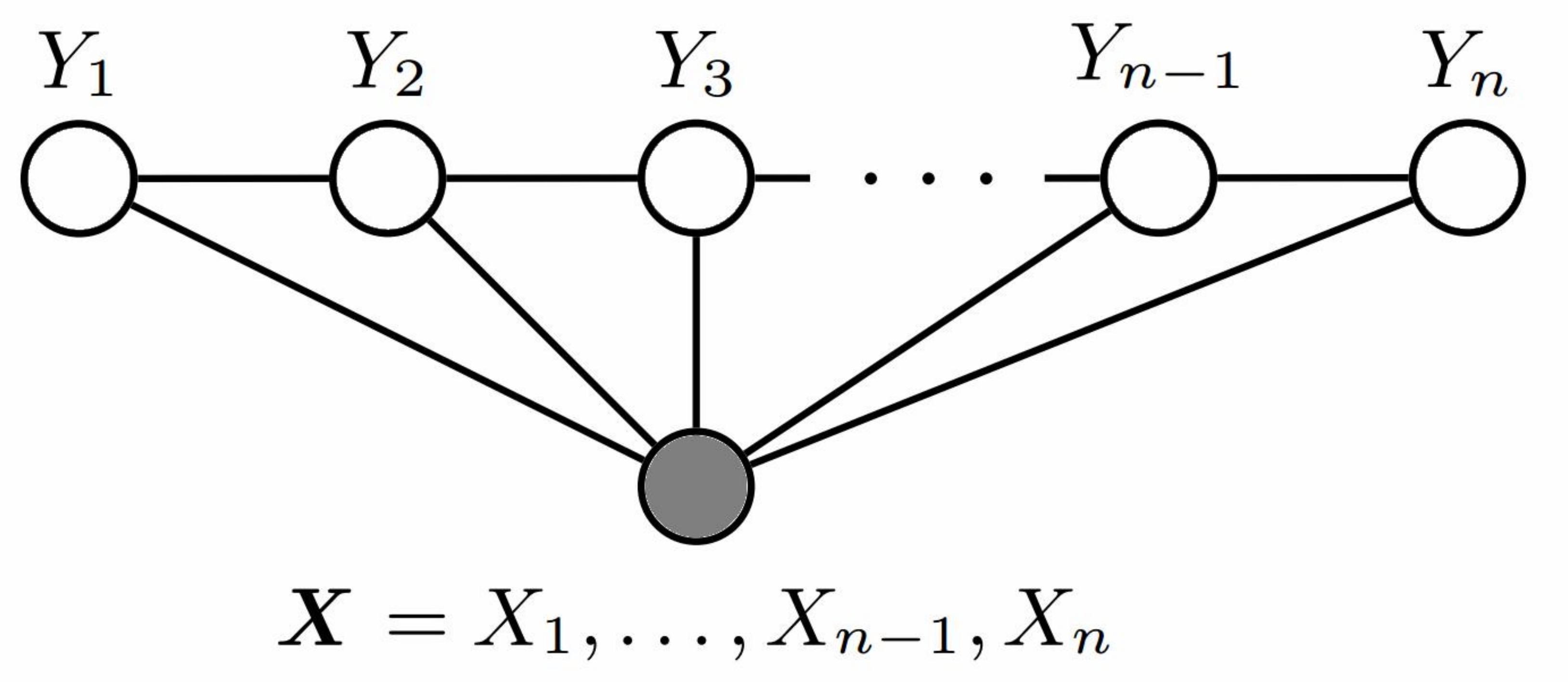
- 其中，C是G的最大团，$\psi_{c}(Y_{c})$ 是C上定义的严格正函数，被称作势函数(Potential Function)。因子分解是在UGM所有的最大团上进行的。

线性链条件随机场

- 设$X=(X_1,X_2\dots X_n)$和$Y=(Y_1,Y_2\dots Y_m)$都是联合随机变量，若随机变量Y构成一个无向图G=(V,E)表示的马尔科夫随机场(MRF)，则条件概率分布P(Y|X)称为条件随机场(Conditional Random Field, CRF)
- 即：　$P(Y_v|X, Y_w, w \neq v) = P(Y_v|X, Y_w, w \cong v)$
- 其中，$w \cong v$表示与结点v相连的所有结点w
- 一种重要而特殊的CRF是线性链条件随机场(Linear Chain Conditional Random Field)，可用于标注等问题。这时，条件概率P(Y|X)中，Y表示标记序列(或称状态序列)，X是需要标注的观测序列。

线性链条件随机场

- 线性链条件随机场的无向图模型



线性链条件随机场的定义

- 设 $X=(X_1,X_2\ldots X_n)$ 和 $Y=(Y_1,Y_2\ldots Y_n)$ 均为线性链表示的随机变量序列，若在给定随机变量序列X的条件下，随机变量序列Y的条件概率分布P(Y|X)构成条件随机场，即满足马尔科夫性

$$P(Y_i|X,Y_1,Y_2\cdots Y_n)=P(Y_i|X,Y_{i-1},Y_{i+1})$$

- 则称P(Y|X)为**线性链条件随机场**。在标注问题中，X表示输入序列或称观测序列，Y表述对应的输出标记序列或称状态序列。

线性链条件随机场的参数化形式

- 设P(Y|X)为线性链条件随机场，则在随机变量X取值为x的条件下，随机变量Y取值为y的条件概率有以下形式：

$$P(y|x)=\frac{1}{Z(x)}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1},y_i,x,i)+\sum_{i,l}\mu_l s_l(y_i,x,i)\right)$$

- 其中，$Z(x)=\sum_{y}\exp\left(\sum_{i,k}\lambda_k t_k(y_{i-1},y_i,x,i)+\sum_{i,l}\mu_l s_l(y_i,x,i)\right)$
  - 特征函数：$t_k(y_{i-1},y_i,x,i)$、$s_l(y_i,x,i)$
  - 特征函数对应的权值：λ_k、μ_l
  - Z(x)为规范化因子，保证P(Y|X)为概率分布。

参数说明

- t_k是定义在边上的特征函数，称为转移特征，依赖于当前和前一个位置；
- s_l是定义在结点上的特征函数，称为状态特征，依赖于当前位置；
- t_k 和s_l都依赖于位置，是局部特征函数；
- 通常， t_k 和s_l取值为1或者0；满足特征条件时取1，否则取0；
- CRF完全由特征函数t_k、s_l和对应的权值λ_k μ_l确定。
- 线性链条件随机场模型属于对数线性模型。

条件随机场举例

[PRP He] [VBZ reckons] [DT the] [JJ current] [NN account] [NN deficit] [MD will] [VB narrow] [TO to] [RB only] [# #] [CD 1.8] [CD billion] [IN in] [NNP September] [. .]

- NN、NNS、NNP、NNPS、PRP、DT、JJ分别代表普通名词单数形式、普通名词复数形式、专有名词单数形式、专有名词复数形式、代词、限定词、形容词

$$b(\boldsymbol{x}, i) = \begin{cases} 1 & \text{if the observation at position } i \text{ is the word "September"} \\ 0 & \text{otherwise.} \end{cases}$$

$$t_j(y_{i-1}, y_i, \boldsymbol{x}, i) = \begin{cases} b(\boldsymbol{x}, i) & \text{if } y_{i-1} = \text{IN and } y_i = \text{NNP} \\ 0 & \text{otherwise.} \end{cases}$$

条件随机场举例

- 设有一标注问题：输入观测序列为X=(X_1,X_2,X_3)，输出标记序列为Y=(Y_1,Y_2,Y_3)，Y_1,Y_2,Y_3的取值范围为{1,2}。
- 对于第一条连接边，假设特征和权值如下：

$$t_1 = t_1\left(y_{i-1}=1, y_i=2, x, i\right), \quad i=2,3, \quad \lambda_1 = 1$$

- 其含义是对于给出条件的式子，值为1；否则，值为0，即：

$$t_1\left(y_{i-1}, y_i, x, i\right) = \begin{cases} 1 & \text{当} y_{i-1}=1, y_i=2, i=2,3 \\ 0 & \text{其他} \end{cases}$$

条件随机场举例

- $t1=t1(y1=1,y2=2,x,2)$ $\lambda_1=1$
- $t1=t1(y2=1,y3=1,x,3)$ $\lambda_1=1$
- $t2=t2(y1=1,y2=1,x,2)$ $\lambda_2=0.5$
- $t3=t3(y2=2,y3=1,x,3)$ $\lambda_3=1$
- $t4=t4(y1=2,y2=1,x,2)$ $\lambda_4=1$
- $t5=t5(y2=2,y3=2,x,3)$ $\lambda_5=0.2$
- $s1=s1(y1=1,x,1)$ $\mu_1=1$
- $s2=s2(y2=2,x,i)$ $\mu\ 2=0.5$
- $s3=s3(y1=1,x,i)$ $\mu_3=0.8$
- $s4=s4(y3=2,x,i)$ $\mu_4=0.5$

条件随机场举例

- 则标记序列为y=(1,2,2)的非规范化概率为：

$$
\begin{aligned}
P(y|x) &= \frac{1}{Z(x)} \exp\left(\sum_{i,k} \lambda_k t_k (y_{i-1}, y_i, x, i) + \sum_{i,l} \mu_l s_l (y_i, x, i) \right) \\
&\propto \exp\left(\sum_{k=1}^{5} \lambda_k \sum_{i=2}^{3} t_k (y_{i-1}, y_i, x, i) + \sum_{l=1}^{4} \mu_l \sum_{i=1}^{3} s_l (y_i, x, i) \right) \\
&= \exp(3.2)
\end{aligned}
$$

使用统一的函数记号表达特征

- 为方便起见，将转移特征和状态特征及其权值同统一的符号表示。
- 设有K_1个转移特征，K_2个状态特征，$K=K_1+K_2$，则

$$f_k(y_{i-1},y_i,x,i)=\begin{cases} t_k(y_{i-1},y_i,x,i) & k=1,2,\cdots K_1 \\ s_l(y_i,x,i) & k=K_1+l;l=1,2,\cdots K_2 \end{cases}$$

- 在各个位置求和：

$$f_k(y,x)=\sum_{i=1}^{n} f_k(y_{i-1},y_i,x,i),\quad k=1,2,\cdots K$$

CRF的简化形式

- 用w表示统一的权值：

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{pmatrix}$$

- 则CRF可表示成：

$$P(y|x) = \frac{1}{Z(x)} \exp\left(\sum_{k=1}^{K} w_k f_k(y, x) \right)$$

$$Z(x) = \sum_{y} \exp\left(\sum_{k=1}^{K} w_k f_k(y, x) \right)$$

CRF的简化形式

- 以F(y,x)表示全局特征向量：

$$F(y,x)=\begin{pmatrix} f_1(y,x) \\ f_2(y,x) \\ \vdots \\ f_K(y,x) \end{pmatrix}$$

- 则CRF可以写成向量内积的形式：

$$P(y|x)=\frac{\exp\left(w^T F(y,x)\right)}{Z(x)}$$

$$Z_w(x)=\sum_y \exp\left(w^T F(y,x)\right)$$

CRF的矩阵形式

- 引入特殊的起点、终点标记：
  - y_0=start，y_{n+1}=stop
  - start,stop也是属于标记空间的某两个值
- 定义m阶矩阵(m是标记y_i取值的个数) $M_i(x)=\left[M_i\left(y_{i-1},y_i\middle|x\right)\right]$
- $M_i(x)$的元素是状态序列的转移概率，它的第r行第s列表示当前状态是r时，转移到状态s的概率。
  - 事实上，这里的状态r即为y_{i-1}，下一个状态即y_i
- 根据CRF的简化记号，矩阵Mi的元素为：

$$M_i\left(y_{i-1},y_i\middle|x\right)=\exp\left(W_i\left(y_{i-1},y_i\middle|x\right)\right)$$

- 其中，

$$W_i\left(y_{i-1},y_i\middle|x\right)=\sum_{i=1}^{K}w_k f_k\left(y_{i-1},y_i,x,i\right)$$

CRF的矩阵乘积和条件概率

- 条件概率P(y|x)可以写成如上这些n+1个状态转移矩阵中相应状态$y_1y_2\ldots y_n$对应的元素的乘积：

$$P_w(y|x)=\frac{1}{Z_w(x)}\prod_{i=1}^{n+1}M_i(y_{i-1},y_i|x)$$

- 其中，Z(x)是归一化因子，对上式两边求关于y的积分(加和)，得：

$$Z_w(x)=\left(M_1(x)M_2(x)\cdots M_{n+1}(x)\right)_{start,stop}$$

  - y_0=start，y_{n+1}=stop表示开始和终止状态，是以start为起点stop为终点通过状态$y_1y_2\ldots y_n$的所有路径的非规范化概率之和，所以，恰好是概率转移矩阵连乘n+1次后得到的新矩阵(start,end)位置的元素值。

CRF的三个问题

- CRF的概率计算问题
  - 前向后向算法
- CRF的参数学习问题
  - IIS：改进的迭代尺度算法
- CRF的预测算法
  - Viterbi算法

CRF的概率计算问题

- 给定条件随机场P(Y|X)，输入序列x和输出序列y，计算：
  - 条件概率$P(Y_i=y_i|x)$
  - 条件联合概率$P(Y_{i-1}=y_{i-1}, Y_i=y_i|x)$
  - 手段：前向后向算法

前向向量

- $\alpha_i(y_i|x)$：在位置i的标记是y_i并且到位置i的前部分标记序列的非规范化概率。
  - yi可取的值有m个：$\alpha_i(y_i|x)$是m维列向量

$$\alpha_0(y|x)=\begin{cases}1, & y=start\\0, & otherwise\end{cases}$$

$$\alpha_i^T(y_i|x)=\alpha_{i-1}^T(y_{i-1}|x)M_i(y_{i-1},y_i|x),\quad i=1,2,\cdots,n+1$$

$$\alpha_i^T(x)=\alpha_{i-1}^T(x)M_i(x)$$

后向向量

- β$_i$(y$_i$|x)：在位置i的标记是y$_i$并且从i+1到n的后部分标记序列的非规范化概率。
  - yi可取的值有m个：β$_i$(y$_i$|x)是m维列向量

$$\beta_{n+1}\left(y_{n+1}\middle|x\right)=\begin{cases}1 & y_{n+1}=stop\\ 0 & y_{n+1}\neq stop\end{cases}$$

$$\beta_i\left(y_i\middle|x\right)=M_i\left(y_i,y_{i+1}\middle|x\right)\beta_{i+1}\left(y_{i+1}\middle|x\right),\quad i=1,2,\cdots n$$

$$\beta_i\left(x\right)=M_{i+1}\left(x\right)\beta_{i+1}\left(x\right)$$

归一化因子

- 归一化因子Z(x)：

$$Z(x)=\alpha_n^T(x)\cdot 1=1^T\cdot\beta_1(x)$$

概率计算

- 根据非规范化的前向、后向概率，条件概率 $P(Y_i=y_i|x)$，联合条件概率 $P(Y_{i-1}=y_{i-1}, Y_i=y_i|x)$ 分别为：

$$P(Y_i = y_i|x) = \frac{\alpha_i^T(y_i|x)\beta_i(y_i|x)}{Z(x)}$$

$$P(Y_{i-1} = y_{i-1}, Y_i = y_i|x) = \frac{\alpha_{i-1}^T(y_{i-1}|x)M_i(y_{i-1}, y_i|x)\beta_i(y_i|x)}{Z(x)}$$

CRF的参数学习问题

- 根据训练数据集合(x,y)，可以求出经验概率分布P(x,y)
  - 大数定理：用频率估计概率
- 使用极大似然估计给出目标函数：

$$L(w)=\log\prod_{x,y}P_w(y\mid x)^{\widetilde{P}(x,y)}=\sum_{x,y}\widetilde{P}(x,y)\log P_w(y\mid x)$$

$$=\sum_{x,y}\left[\widetilde{P}(x,y)\left(\sum_{k=1}^{K}w_k f_k(x,y)\right)-\widetilde{P}(x,y)\log Z_w(x)\right]$$

$$=\left(\sum_{j=1}^{N}\sum_{k=1}^{K}w_k f_k(x_j,y_j)\right)-\sum_{j=1}^{N}\log Z_w(x_j)$$

改进的迭代尺度算法

- 算法通过计算似然函数该变量的下界，从而给出下界的最优值，迭代进行，不断更新参数W。

$$w = \left(w_1, w_2, \cdots, w_K\right)^T$$

$$\delta = \left(\delta_1, \delta_2, \cdots, \delta_K\right)^T$$

$$w + \delta = \left(w_1 + \delta_1, w_2 + \delta_2, \cdots, w_K + \delta_K\right)^T$$

变化率δ的函数

- 经过计算，关于转移特征t_k的更新方程为：

$$E_{\tilde{P}}[t_k]=\sum_{x,y}\tilde{P}(x)P(y\mid x)\left(\sum_{i=1}^{n+1}t_k(y_{i-1},y_i,x,i)\exp(\delta_k T(x,y))\right),\quad k=1,2,\cdots,K_1$$

- 关于状态特征s_l的更新方程为：

$$E_{\tilde{P}}[s_l]=\sum_{x,y}\tilde{P}(x)P(y\mid x)\left(\sum_{i=1}^{n}s_l(y_i,x,i)\exp\left(\delta_{K_1+l}T(x,y)\right)\right),\quad l=1,2,\cdots,K_2$$

- 其中T(x,y)为在数据(x,y)中出现的所有特征数之和

$$T(x,y)=\sum_{k=1}^{K}\sum_{i=1}^{n+1}f_k(y_{i-1},y_i,x,i)$$

参数学习：改进的迭代尺度算法IIS

- 在当前转移特征t_k和状态特征s_l给定的条件下，求关于未知向量δ的式子：

$$E_{\widetilde{P}}[t_k]=\sum_{x,y}\widetilde{P}(x)P(y\mid x)\left(\sum_{i=1}^{n+1}t_k(y_{i-1},y_i,x,i)\exp(\delta_k T(x,y))\right),\quad k=1,2,\cdots,K_1$$

- 取所有T(x,y)的最大值S，并用S代替其他T(x,y)，则上式可直接给出解析式：

$$\delta_k=\frac{1}{S}\log\frac{E_{\widetilde{P}}[t_k]}{E_P[t_k]}$$

- 其中，

$$E_P[t_k]=\sum_x\widetilde{P}(x)\left(\sum_{i=1}^{n+1}\sum_{y_{i-1},y_i}t_k(y_{i-1},y_i,x,i)\frac{\alpha_{i-1}^T(y_{i-1}\mid x)M_i(y_{i-1},y_i\mid x)\beta_i(y_i\mid x)}{Z(x)}\right)$$

- 否则，需要使用拟牛顿或L-BFGS等方式计算δ。

CRF的预测算法

- CRF的预测问题，是给定条件随机场P(Y|X)和输入序列(观测序列)x，求条件概率最大的输出序列(标记序列)y*，即对观测序列进行标注。
  - Viterbi算法

最优路径的目标函数

- 根据推导：

$$\begin{aligned} y^* &= \arg\max_y P_w(y \mid x) \\ &= \arg\max_y \frac{\exp(wF(y,x))}{Z_w(x)} \\ &= \arg\max_y \exp(wF(y,x)) \\ &= \arg\max_y wF(y,x) \end{aligned}$$

- CFR的预测问题即非规范化概率最大的最优路径问题。

最优路径的目标函数

- 其中，

$$w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{pmatrix} \qquad F(y,x) = \begin{pmatrix} f_1(y,x) \\ f_2(y,x) \\ \vdots \\ f_K(y,x) \end{pmatrix}$$

$$f_k(y,x) = \sum_{i=1}^{n} f_k(y_{i-1}, y_i, x, i)$$

最优路径的目标函数

- 目标函数：

$$\max_y \sum_{i=1}^{n} w \cdot F_i\left(y_{i-1}, y_i, x\right)$$

- 其中：

$$F_i\left(y_{i-1}, y_i, x\right) = \begin{pmatrix} f_1\left(y_{i-1}, y_i, x, i\right) \\ f_2\left(y_{i-1}, y_i, x, i\right) \\ \vdots \\ f_K\left(y_{i-1}, y_i, x, i\right) \end{pmatrix}$$

状态预测：Viterbi算法

- 定义：位置i的各个标记l=1,2,…,m的非规范化概率的最大值。

$$\delta_1(l) = w \cdot F_1(y_0 = start, y_1 = l, x) \quad l = 1,2,\cdots m$$

$$\delta_i(l) = \max_{1 \le j \le m} \{\delta_{i-1}(j) + w \cdot F_i(y_{i-1} = j, y_i = l, x)\} \quad l = 1,2,\cdots m$$

$$y_n^* = \arg\max_{1 \le j \le m} \delta_n(j)$$

CRF总结

- 条件随机场是给定输入的条件下，关于输出的条件概率分布模型，根据Hammersley-Clifford定理，可以分解成若干关于最大团的非负函数的乘积，因此，常常将其表示为参数化的对数线性模型。
- 线性链条件随机场使用对数线性模型，关注无向图边的转移特征和点的状态特征，并对每个特征函数给出各自的权值。
  - 概率计算常常使用前向-后向算法；
  - 参数学习使用MLE建立目标函数，采用IIS做参数优化；
  - 线性链条件随机场的应用是标注/分类，在给定参数和观测序列(样本)的前提下，使用Viterbi算法进行标记的预测。

参考文献

- Machine Learning:A Probabilistic Perspective, Kevin P. Murphy, The MIT Press, 2012
- Conditional Random Fields: An Introduction, Hanna M. Wallach, 2004
- An Introduction to Conditional Random Fields, Charles Sutton, Andrew McCallum, 2012
- 统计学习方法，李航著，清华大学出版社，2012年
- Pattern Recognition and Machine Learning Chapter 13, Bishop M, Springer-Verlag, 2006
- Radiner L,Juang B. An introduction of hidden markov Models. IEEE ASSP Magazine, January 1986

我们在这里

7 七月算法 http://www.julyedu.com/

- 视频/课程/社区

- 七月题库APP：Android/iOS
  - http://www.julyapp.com/
- 微博
  - @研究者July
  - @七月题库
  - @邹博_机器学习
- 微信公众号
  - julyedu

感谢大家！

恳请大家批评指正！