Natural Language Process

条件随机场

原理介绍

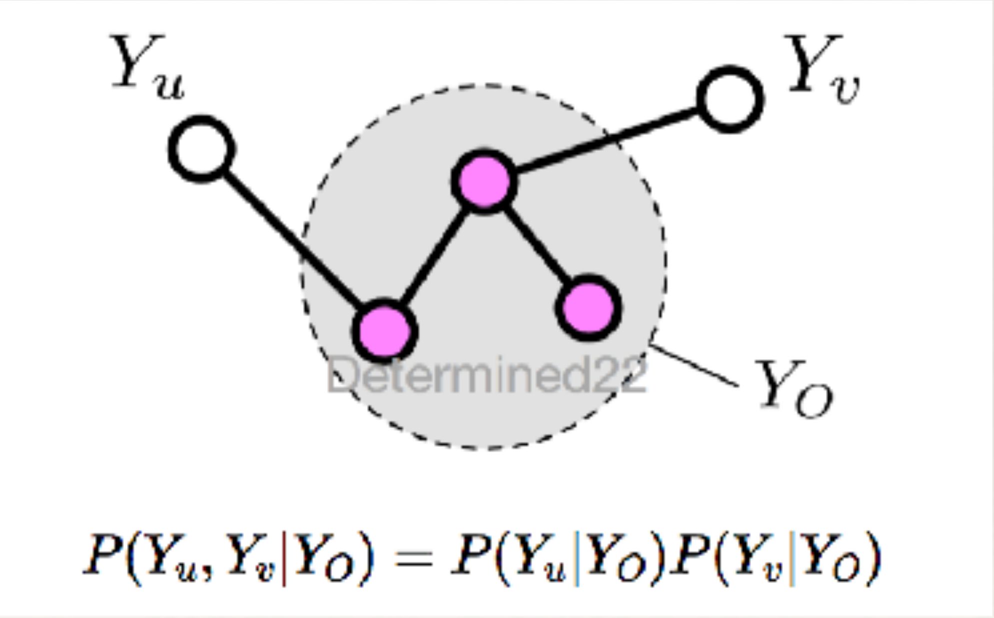
目录

1 马尔可夫随机场-马尔可夫性

- 马尔可夫随机场又叫概率无向图模型，表示一个联合概率分布。设有联合概率分布 P(V) 由无向图 G=(V, E) 表示，图 G 中的节点表示随机变量，边表示随机变量间的依赖关系。如果联合概率分布 P(V) 满足成对、局部或全局马尔可夫性，就称此联合概率分布为马尔可夫随机场。
- 设有一组随机变量 Y ，其联合分布为 P(Y) 由无向图 G=(V, E) 表示。图 G 的一个节点$v \in V$表示一个随机变量Y_v ，一条边$e \in E$ 就表示两个随机变量间的依赖关系。
- 下面分别介绍成对、局部和全局马尔可夫性。

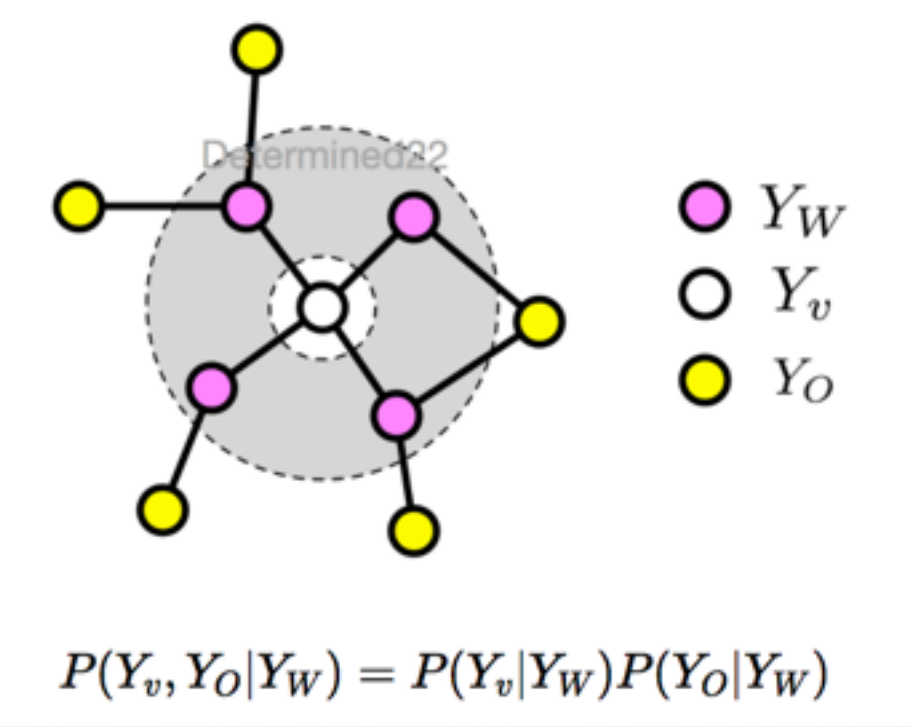
1 马尔可夫随机场-马尔可夫性

- （1）成对马尔可夫性：
- 设无向图 G 中的任意两个没有边连接的节点 u 、 v ，其他所有节点为 O ，成对马尔可夫性指：给定Y_O 的条件下，Y_u和 Y_v条件独立。



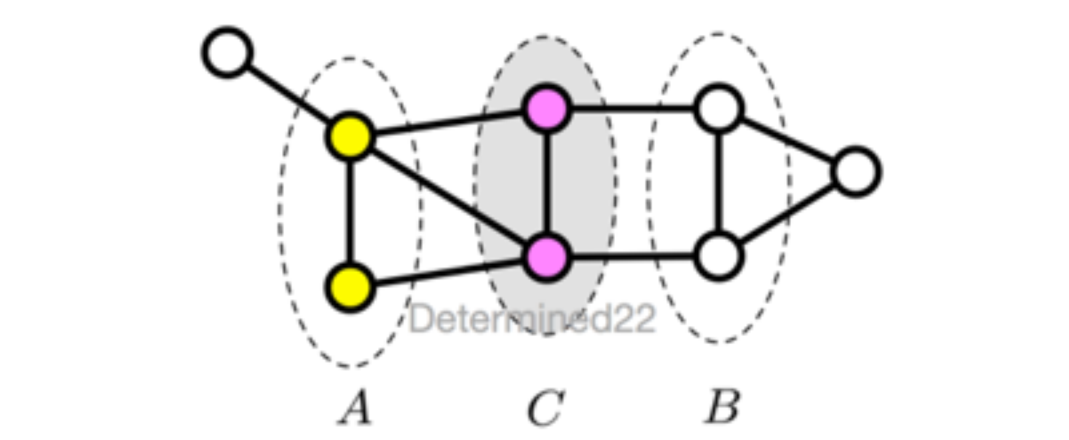
1 马尔可夫随机场-马尔可夫性

- （2） 局部马尔可夫性：
- 设无向图 G 的任一节点 v ， W 是与 v 有边相连的所有节点，O 是 v 、W 外的其他所有节点，局部马尔可夫性指：给定 Y_W 的条件下，Y_v 和 Y_O 条件独立。



1 马尔可夫随机场-马尔可夫性

- （3） 全局马尔可夫性：
- 设节点集合 A、B 是在无向图 G 中被节点集合 C 分开的任意节点集合，全局马尔可夫性指：给定 Y_C 的条件下，Y_A 和 Y_B 条件独立。



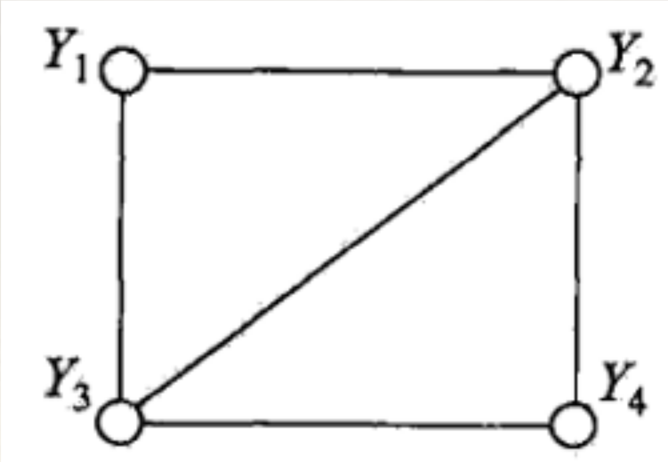
$$P(Y_A, Y_B | Y_C) = P(Y_A | Y_C) P(Y_B | Y_C)$$

1 马尔可夫随机场-因子分解

- 无向图模型的优点在于其没有隐马尔可夫模型那样严格的独立性假设，同时克服了最大熵马尔可夫模型等判别式模型的标记偏置问题。参考文献[7]中对此有详细介绍。
- 我们需要关心的问题是如何求解联合分布P(V)。对于给定的概率无向图模型，我们将整体的联合概率写出若干的子联合概率的乘积的形式，也就将联合概率进行因子分解。
- 将概率无向图模型的联合概率分布表示为其最大团上的随机变量的函数的乘积形式的操作，称为概率无向图模型的因子分解。

1 马尔可夫随机场-因子分解

- 团与最大团
- 无向图G中任何两个结点均有边连接的结点子集称为团（clique）。若C是无向图G的一个团，并且不能再加进任何一个G的结点使其成为一个更大的团，则称此C为最大团。
- 如下面的图，图中由两个结点组成的团有5个，分别是{Y1，Y2}，…{Y1,Y3}。有两个最大团，分别是{Y1,Y2,Y3}和{Y2,Y3,Y4}。{Y1,Y2,Y3,Y4}不是一个团。



1 马尔可夫随机场-因子分解

- 给定概率无向图模型，设其无向图为G，C为G上的最大团，Yc表示C对应的随机变量，那么概率图模型的联合概率分布P（Y）可以写作图中所有最大团上的势函数的乘积形式，即：

$$P(Y)=\frac{1}{Z}\prod_{C}\Psi_{C}(Y_{C})$$

- 其中Z是规范化因子，由式子 $Z=\sum_{Y}\prod_{C}\Psi_{C}(Y_{C})$ 给出。规范化因子保证P(Y)构成一个概率分布。函数 $\Psi_{C}(Y_{C})$ 表示势函数。这里要求势函数是严格为正的，通常定义为指数函数:

$$\Psi_{C}(Y_{C})=\exp\{-E(Y_{C})\}$$

- 概率无向图模型的因子分解由Hammenrsley-Clifford定理来保证。详细的信息请看参考文献。

2 条件随机场-定义

- **条件随机场（Conditional random field，CRF）**是条件概率分布模型P(Y|X)，表示的是给定一组输入随机变量 X 的条件下另一组输出随机变量 Y 的条件概率分布。条件随机场可被看作是最大熵马尔可夫模型在标注问题上的推广。若随机变量Y构成一个由无向图G=(V,E)表示的马尔可夫随机场，即

$$P(Y_v \mid X, Y_w, w \neq v) = P(Y_v \mid X, Y_w, w \sim v)$$

- 对任意的结点v成立，则称条件概率分布P(Y|X)为条件随机场。上面的式子中w~v表示在图G中与结点v有边连接的所有结点w，$w \neq v$表示结点v以外的所有结点。
- 设X和Y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，则称P(Y|X)为线性链条件随机场。

2 条件随机场-定义

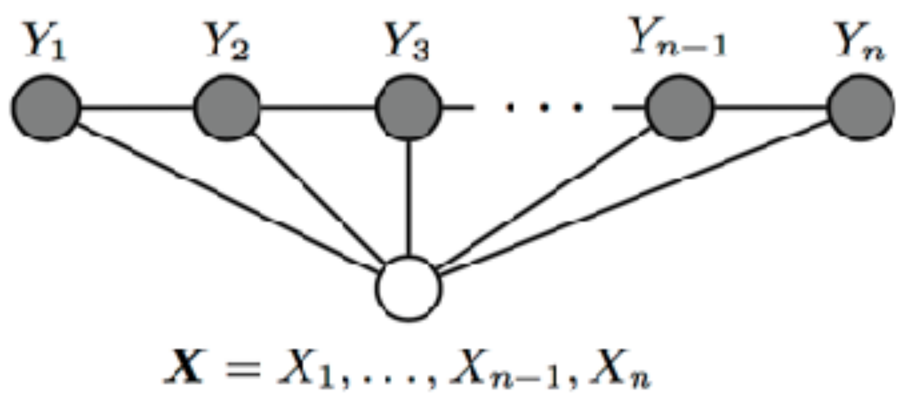


Figure 1: Graphical structure of a chain-structured CRFs for sequences. The variables corresponding to unshaded nodes are *not* generated by the model.

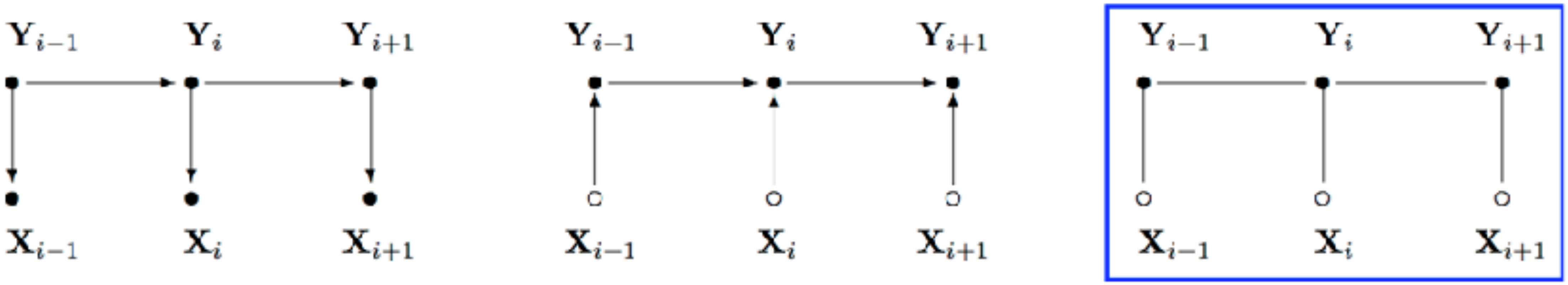


Figure 2. Graphical structures of simple HMMs (left), MEMMs (center), and the chain-structured case of CRFs (right) for sequences. An open circle indicates that the variable is not generated by the model.

2 条件随机场-参数形式

- 根据前文的介绍，可以知道线性链条件随机场的最大团由相连的两个结点组成。这样我们可以给出P(Y|X)的因子分解式，各因子是定义在相邻两个结点上的函数。
- 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的条件概率具有如下形式：

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

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

2 条件随机场-参数形式

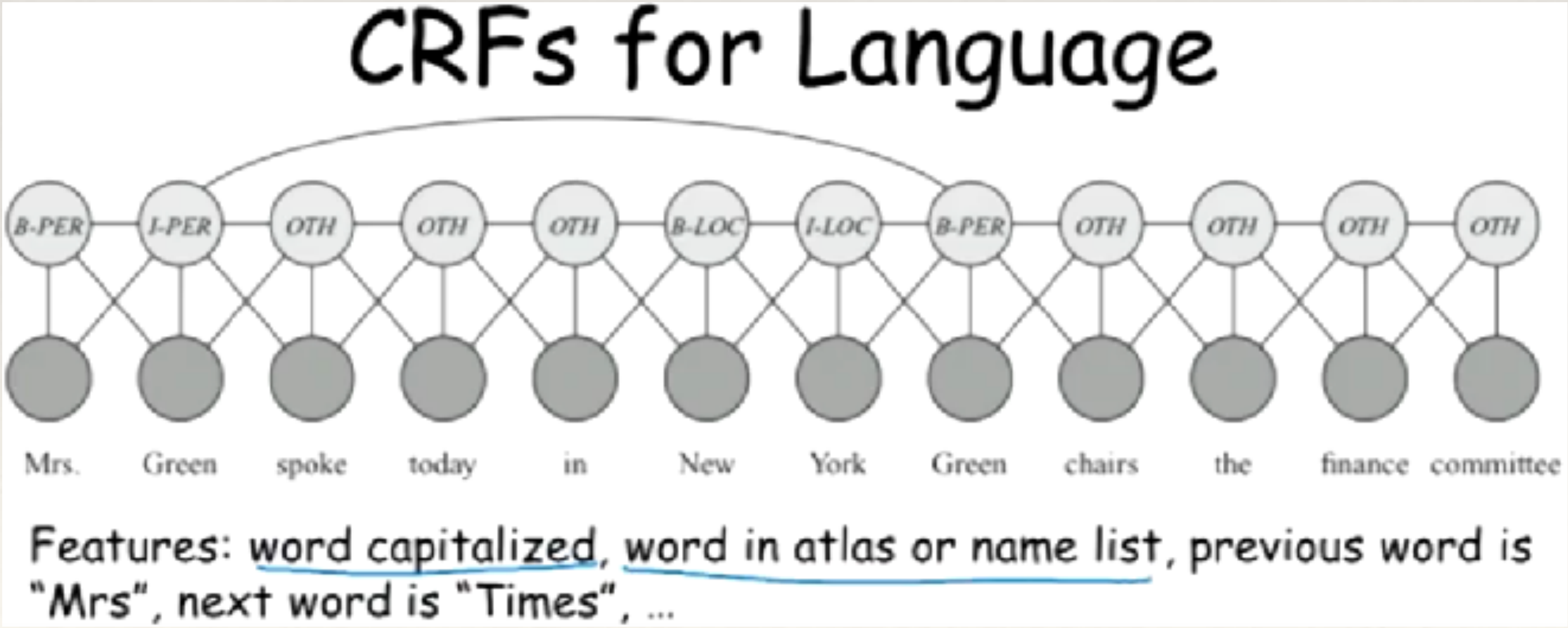
$$P(y|x)=\frac{1}{Z(x)}\exp(\sum_{i,k}\lambda_k t_k(y_{i-1},y_i,x,i)+\sum_{i,l}\mu_l s_l(y_i,x,i))$$

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

- 在上面的式子中，t是定义在边上的特征函数，称为转移特征，依赖于当前位置和前一个位置；s是定义在结点上的特征函数，称为状态特征，依赖当前位置。t和s都依赖于位置，是局部特征函数。特征函数t和s的取值为1或者0，当满足条件时取值为1，否则为0。CRF可以把整个图模型的局部特征函数转化为全局特征函数，从而写成条件随机场的权重向量和特征向量的内积形式。这样我们就可以每次只关心一个特征，再把所有全局特征集中起来一起考虑。

2 条件随机场-参数形式

- 例如词性标注问题，我们不需要详细理解每个相邻word直接的关系。可以先找出word是否是大写，然后找出word是否在姓名列表中，接着找出word的上一个word是否是“Mrs”，下一个word是否是”times”。最后把这些分布结合在一起看，就较容易判断了。



2 条件随机场-简化表示

- 由公式可知，同一个特征在各个位置都有定义，可以对同一个特征在各个位置求和，将局部特征函数转换为一个全局的特征函数。
- 设有K1个转移特征，K2个状态特征，K=K1+K2，记：

$$f_k(y_{i-1},y_i,x,i)=\begin{cases} t_k(y_{i-1},y_i,x,i),k=1,2,...,K1 \\ s_l(y_i,x,i),k=K1+l;l=1,2,...,K2 \end{cases}$$

- 然后，对转移和状态特征在各个位置i求和，记为：

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

2 条件随机场-简化表示

- 用w_k表示特征f_k(y,x)的权值，即

$$w_k = \begin{cases} \lambda_k, k = 1,2,...,K1 \\ \mu_l, k = K1 + l; l = 1,2,...,K2 \end{cases}$$

- 于是，条件随机场可以表示为：

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

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

2 条件随机场-简化表示

- 若以w表示权重向量，以F(y,x)表示全局特征向量，即：

$$F(y,x)=(f_1(y,x),f_2(y,x),\ldots,f_k(y,x))^T$$

- 则条件随机场可以写成向量w与F(y,x)的内积的形式。

$$P_w(y\mid x)=\frac{\exp(w.F(y,x))}{Z_w(x)}$$

$$Z_w(x)=\sum_y \exp(w.F(y,x))$$

2 条件随机场-矩阵形式

- 条件随机场还可以以矩阵的形式表示。条件随机场的矩阵形式其实就是一种模型按时序展开，分步计算的形式。类似于对输入观测序列X的每一个x_i：
- 1.计算出所有的可能情况（根据不同的假设y_n-1y_n组合情况，计算激活的特征函数与权值乘积的和），按照一定的顺序组成矩阵；2.在所有矩阵计算完成之后，利用这些矩阵可以完成最优序列的求解，Z(x)的计算。
- 假设现有一个线性链条件随机场，用于求给定观测序列X，相应的标记序列Y的条件概率。引进特殊的起点和终点状态标记y_0=start, y_n+1 = stop，这时条件随机场可以通过矩阵形式表示。

2 条件随机场-矩阵形式

- 引入新的量:

$$M_i(y_{i-1}, y_i|x) = \exp \sum_{k=1}^{K} w_k f_k(y_{i-1}, y_i, x, i), \quad i = 1, 2, \ldots, n+1$$

- 这个量融合了参数和特征，是一个描述模型的简单的量。这个量是针对于某个位置i以及前一个位置i+1的。那么，假设状态序列的状态存在 m 个可能的取值，对于任一位置 i = 1,2,...,n+1 ，定义一个 m 阶方阵:

$$\begin{aligned} M_i(x) &= [\exp \sum_{k=1}^{K} f_k(y_{i-1}, y_i, x, i)]_{m\times m} \\ &= [M_i(y_{i-1}, y_i|x)]_{m\times m} \end{aligned}$$

- 因为 $\prod_i [\exp \sum_{k=1}^{K} w_k f_k(y_{i-1}, y_i, x, i)] = \exp(\sum_{k=1}^{K} w_k \sum_i f_k(y_{i-1}, y_i, x, i))$ ，所以线性链条件随机场可以表述为如下的矩阵形式:

$$P_{\mathbf{w}}(Y = y|x) = \frac{1}{Z_{\mathbf{w}}(x)} \prod_{i=1}^{n+1} M_i(y_{i-1}, y_i|x)$$

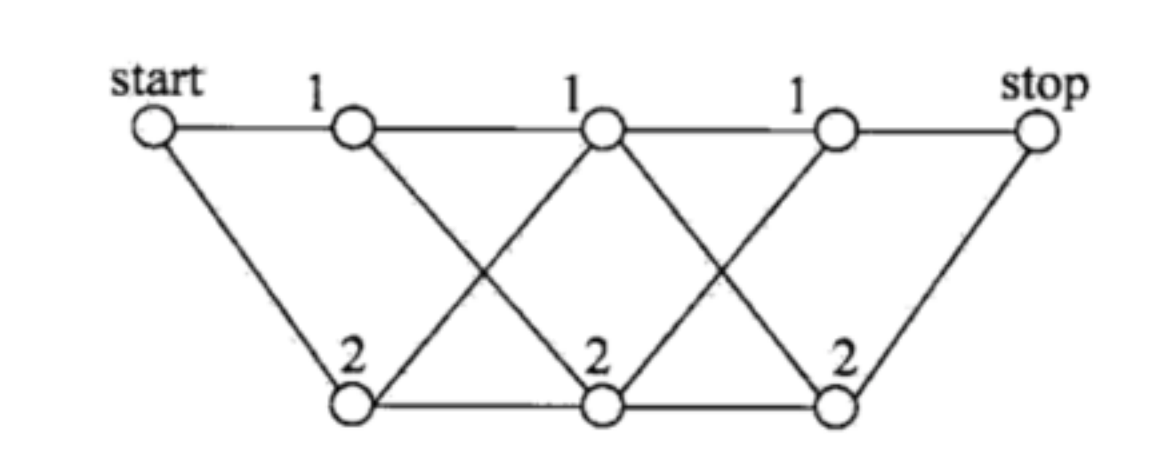
$$Z_{\mathbf{w}}(x) = (M_1(x) M_2(x) \cdots M_{n+1}(x))_{(start,stop)}$$

2 条件随机场-矩阵形式

- 给定一个线性链条件随机场，n = 3 ，状态的可能取值为 1 和 2 。设 y_0 = start = 1 、y_{n+1}=stop=1 ，且 M 矩阵在 i = 1,2,...,n+1 的值已知（如下图），求状态序列以 start 为起点、以stop为终点的所有状态路径的非规范化及规范化概率。

$$M_1(x) = \begin{pmatrix} a_{01} & a_{01} \\ 0 & 0 \end{pmatrix}, \quad M_2(x) = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$M_3(x) = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \quad M_4(x) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$



2 条件随机场-矩阵形式

- 在计算图中从start到stop对应y=(1,1,1), y=(1,1,2),…,y=(2,2,2)，各个路径的非规范化概率分别是:

$$a_{01}b_{11}c_{11},\quad a_{01}b_{11}c_{12},\quad a_{01}b_{12}c_{21},\quad a_{01}b_{12}c_{22}$$
$$a_{02}b_{21}c_{11},\quad a_{02}b_{21}c_{12},\quad a_{02}b_{22}c_{21},\quad a_{02}b_{22}c_{22}$$

- 通过计算矩阵乘积M_1 M_2 M_3 M_4，可以得到第1行第1列的元素为:

$$\begin{aligned}&a_{01}b_{11}c_{11}+a_{02}b_{21}c_{11}+a_{01}b_{12}c_{21}+a_{02}b_{22}c_{22}\\&\qquad+a_{01}b_{11}c_{12}+a_{02}b_{21}c_{12}+a_{01}b_{12}c_{22}+a_{02}b_{22}c_{21}\end{aligned}$$

- 这个值恰好等于start到stop的所有非规范化概率之和，即正则项Z。

3 条件随机场-三个问题

- 概率问题：给定一个条件随机场P(Y|X)，观察序列X和状态序列Y，计算条件概率 $P(Y_i = y_i \mid x)$、$P(Y_{i-1} = y_{i-1}, Y_i = y_i \mid x)$ 。
- 学习问题：给定观察序列X和状态序列Y，学习条件随机场模型的参数。
- 预测问题：给定条件随机场模型和观察序列x，求最有可能的状态序列y*。

3 条件随机场-概率问题

- 概率问题：给定一个条件随机场P(Y|X)，观察序列X和状态序列Y，计算条件概率 $P(Y_i = y_i | x)$ 、$P(Y_{i-1} = y_{i-1}, Y_i = y_i | x)$ 。
- 前向后项算法
- 对每个位置i=0,1,…,n+1，定义前向向量：

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

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

- alpha_i(y_i|x)的含义是在位置 i 的状态是 Y_i=y_i且到位置i的前部分状态序列的非规范化概率，y_i的取值有m个，所以alpha是m维列向量。这个递推式子可以直观地把 M_i(y_{i-1},y_i|x) 理解为"转移概率"，求和符号表示对 y_{i-1}的所有可能取值求和。写成矩阵的形式就是下式：

$$\boldsymbol{\alpha}_i^\top(x) = \boldsymbol{\alpha}_{i-1}^\top(x) M_i(x)$$

3 条件随机场-概率问题

- 概率问题：给定一个条件随机场P(Y|X)，观察序列X和状态序列Y，计算条件概率 $P(Y_i = y_i | x)$ 、$P(Y_{i-1} = y_{i-1}, Y_i = y_i | x)$ 。
- 前向后项算法
- 类似地，可以定义后项向量：

$$\beta_{n+1}(y_{n+1}|x) = \begin{cases} 1, & y_{n+1} = stop \\ 0, & otherwise \end{cases}$$

$$\beta_i(y_i|x) = \sum_{y_{i+1}} M_{i+1}(y_i, y_{i+1}|x)\beta_{i+1}(y_{i+1}|x), \quad i = 0, 1, \ldots, n$$

- beta_i(y_i|x)的含义是在位置 i 的标记 Y_i=y_i且从位置 i+1 到位置 n 的局部标记序列的非规范化概率。写成矩阵的形式就是：

$$\boldsymbol{\beta}_i^{\top}(x) = M_{i+1}(x)\boldsymbol{\beta}_{i+1}(x)$$

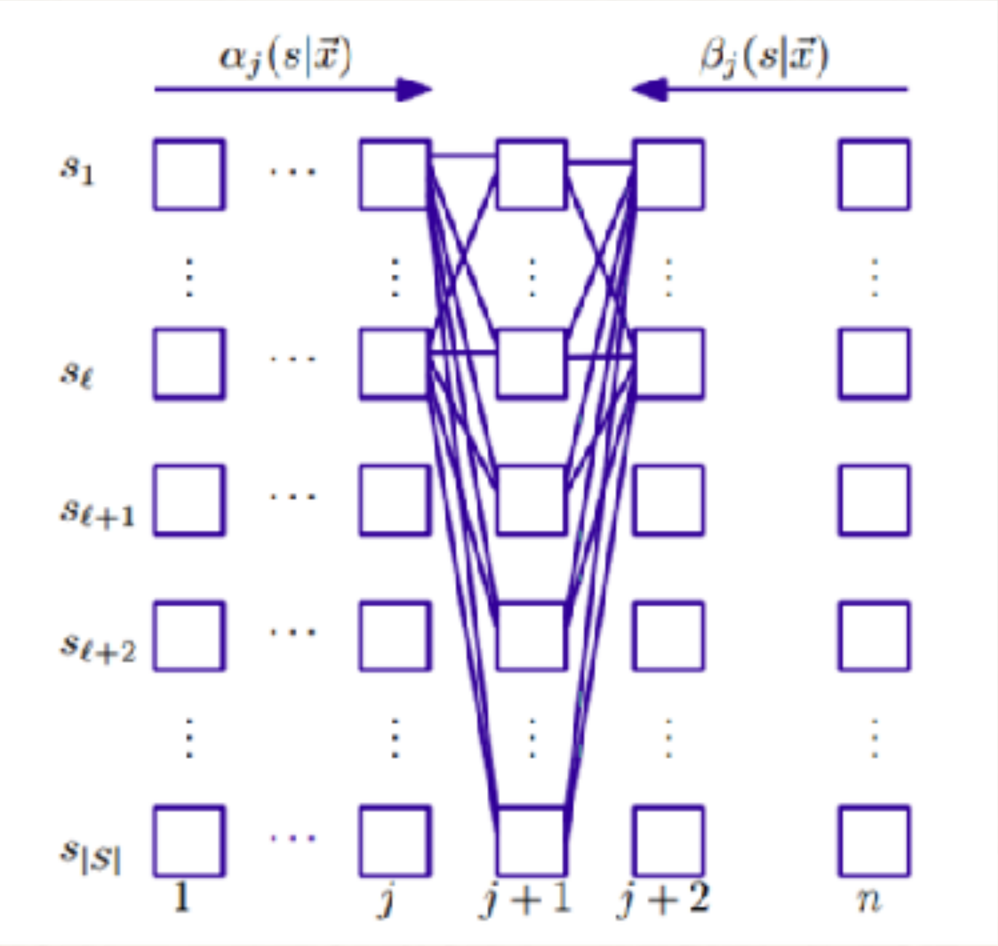
3 条件随机场-概率问题

- 概率问题：给定一个条件随机场P(Y|X)，观察序列X和状态序列Y，计算条件概率 $P(Y_i = y_i \mid x)$、$P(Y_{i-1} = y_{i-1}, Y_i = y_i \mid x)$ 。
- 概率计算
- 类给定一个CRF模型，那么 两个概率可以利用前向向量和后向向量计算为：

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

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

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



4 条件随机场-学习问题

- 学习问题：给定观察序列X和状态序列Y，学习条件随机场模型的参数。
- 已知训练数据集，可以知道经验概率分布，可以通过极大化训练数据的对数似然函数来求解模型参数训练数据的对数似然函数为：

$$
\begin{aligned}
L(\mathbf{w}) = L_{\tilde{P}}(P_{\mathbf{w}}) &= \ln \prod_{x,y} P_{\mathbf{w}}(Y=y|x)^{\tilde{P}(x,y)} \\
&= \sum_{x,y} \tilde{P}(x,y) \ln P_{\mathbf{w}}(Y=y|x) \\
&= \sum_{x,y} \tilde{P}(x,y) \ln \frac{\exp \sum_{k=1}^{K} w_k f_k(y,x)}{Z_{\mathbf{w}}(x)} \\
&= \sum_{x,y} \tilde{P}(x,y) \sum_{k=1}^{K} w_k f_k(y,x) - \sum_{x,y} \tilde{P}(x,y) \ln Z_{\mathbf{w}}(x) \\
&= \sum_{x,y} \tilde{P}(x,y) \sum_{k=1}^{K} w_k f_k(y,x) - \sum_{x} \tilde{P}(x) \ln Z_{\mathbf{w}}(x)
\end{aligned}
$$

- 可以通过IIS或者LBFGS来求解。

4 条件随机场-学习问题

- 学习问题：给定观察序列X和状态序列Y，学习条件随机场模型的参数。
- 对于条件随机场模型：

$$P_w(y|x)=\frac{\exp\left(\sum_{i=1}^{n} w_i f_i(x,y)\right)}{\sum_{y}\exp\left(\sum_{i=1}^{n} w_i f_i(x,y)\right)}$$

- 学习的优化目标是：

$$\min_{w\in\mathbf{R}^n} f(w)=\sum_{x}\tilde{P}(x)\log\sum_{y}\exp\left(\sum_{i=1}^{n} w_i f_i(x,y)\right)-\sum_{x,y}\tilde{P}(x,y)\sum_{i=1}^{n} w_i f_i(x,y)$$

- 其梯度函数是：

$$g(w)=\sum_{x,y}\tilde{P}(x)P_w(y|x)f(x,y)-E_{\tilde{P}}(f)$$

- 用LBFGS来求解，求解过程

4 条件随机场-学习问题

- 学习问题：给定观察序列X和状态序列Y，学习条件随机场模型的参数。

输入：特征函数 $f_1, f_2, \cdots, f_n$；经验分布 $\tilde{P}(X,Y)$；

输出：最优参数值 $\hat{w}$；最优模型 $P_{\hat{w}}(y|x)$.

（1）选定初始点 $w^{(0)}$，取 $\boldsymbol{B}_0$ 为正定对称矩阵，置 $k=0$

（2）计算 $g_k = g(w^{(k)})$. 若 $g_k = 0$，则停止计算；否则转 (3)

（3）由 $B_k p_k = -g_k$ 求出 p_k

（4）一维搜索：求 λ_k 使得

$$f(w^{(k)} + \lambda_k p_k) = \min_{\lambda \geq 0} f(w^{(k)} + \lambda p_k)$$

（5）置 $w^{(k+1)} = w^{(k)} + \lambda_k p_k$

（6）计算 $g_{k+1} = g(w^{(k+1)})$，若 $g_k = 0$，则停止计算；否则，按下式求出 B_{k+1}：

$$B_{k+1} = B_k + \frac{y_k y_k^{\mathrm{T}}}{y_k^{\mathrm{T}} \delta_k} - \frac{B_k \delta_k \delta_k^{\mathrm{T}} B_k}{\delta_k^{\mathrm{T}} B_k \delta_k}$$

其中，

$$y_k = g_{k+1} - g_k, \quad \delta_k = w^{(k+1)} - w^{(k)}$$

（7）置 $k = k+1$，转 (3). ■

5 条件随机场-预测问题

- 预测问题：给定条件随机场模型和观察序列x，求最有可能的状态序列y*。
- viterbi算法
- 通过推导

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

- 可以知道条件随机场的预测问题成为了求非规范化概率最大的路径优化问题：

$$\max_y \ (w \cdot F(y,x))$$

- 这里只需要计算非规范化概率，而不必计算概率（即不需要求规范化项），可以大大提高计算效率。求最优路径，可以将上面的式子写成如下形式：

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

5 条件随机场-预测问题

- 预测问题：给定条件随机场模型和观察序列x，求最有可能的状态序列y*。
- viterbi算法
- 可以知道条件随机场的预测问题成为了求非规范化概率最大的路径优化问题：

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

- 上面公式中：

$$F_i(y_{i-1}, y_i, x) = (f_1(y_{i-1}, y_i, x, i), f_2(y_{i-1}, y_i, x, i), \cdots, f_K(y_{i-1}, y_i, x, i))^{\mathrm{T}}$$

- 后面，将叙述viterbi算法的求解过程

5 条件随机场-预测问题

- 预测问题：给定条件随机场模型和观察序列x，求最有可能的状态序列y*。
- viterbi算法
- 首先求出位置1的各个状态j=1,2,…,m的非规范化概率：

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

- 由递推公式，求出位置i的各个状态l=1,2,…,m的非规范化概率的最大值，同时记录非规范化概率最大值的路径中：

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

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

- 直到i=n时终止，这时求得非规范化概率的最大值为：

$$\max_y (w \cdot F(y, x)) = \max_{1 \leqslant j \leqslant m} \delta_n(j)$$

5 条件随机场-预测问题

- 预测问题：给定条件随机场模型和观察序列x，求最有可能的状态序列y*。
- viterbi算法

输入：模型特征向量 $F(y,x)$ 和权值向量 w，观测序列 $x=(x_1,x_2,\cdots,x_n)$；

输出：最优路径 $y^*=(y_1^*,y_2^*,\cdots,y_n^*)$．

（1）初始化

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

（2）递推．对 $i=2,3,\cdots,n$

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

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

（3）终止

$$\max_y(w\cdot F(y,x))=\max_{1\leqslant j\leqslant m}\delta_n(j)$$

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

（4）返回路径

$$y_i^*=\Psi_{i+1}(y_{i+1}^*)\,,\quad i=n-1,n-2,\cdots,1$$

求得最优路径 $y^*=(y_1^*,y_2^*,\cdots,y_n^*)$．

参考文献

- 【1】统计学方法
- 【2】http://www.cnblogs.com/Determined22/p/6915730.html
- 【3】Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data
- 【4】https://disi.unitn.it/~passerini/teaching/2010-2011/SRL/slides/crf/talk.pdf
- 【5】条件随机场理论综述
- 【6】马尔可夫网络
- 【7】http://blog.csdn.net/lskyne/article/details/8669301