

## Viterbi algorithm for CRF:

### 基本定义:

定义一个前向向量  $\alpha$  ,其物理意义是由前一时刻转向当前时刻后, 在当前时刻出现某标签的'势'(potential)。

前向向量  $\alpha$  的计算方法如下:

$$\alpha_i(x_{seq}) = \alpha_{i-1}(x_{seq}) \times M_i(x_{seq})$$

特别地, 当  $i=0$  时, 即  $i$  处于序列第一位时, 由于我们赋予了它特殊的标签值: 'start', 所以将它定义如下:

$$\alpha_0(y|x_{seq}) = \begin{cases} 1, & \text{if } y = 'start' \\ 0, & \text{otherwise} \end{cases}$$

可见是与 CRF 原文中的前向向量定义是一样的。跟原来 HMM 的解码算法不同, 原来的是 **transition probability** 乘上 **emission probability**, 但是这里由于矩阵  $M$  已经是 **conditioned on x-sequence** 的了, 所以已经不需要再乘上一个类似 **emission probability** 的东西。

### 具体操作:

解码算法主要用于 test 部分, 输入是一条 x-sequence, 一套已经 optimized 的参数集, 即模型参数, 任务是要输出 **x-sequence 相对应的标签 y-sequence**。

首先计算 matrix  $M$ , 计算方法跟论文一样, 不作多论。

然后计算一连串的  $\alpha$  , 用上述方法生成。

从每一步中产生的  $\alpha_i$  中, 查找最大值的那一维, 把其标签值加入到 y-sequence 中作为预测标签。

最后得出结果: y-sequence。

### 流程图:

