1. CRF与HMM的区别(博客解释)

说词性与词语直接的映射概率，这里体现了CRF的条件特征。如果是HMM，这里会直接统计词性-->单词的条件概率矩阵，比如 ”动词“ 生成 ”发射“ 的概率可能为1.5%，而生成”微软“ 的概率为0. 然后对于每一种可能的词性序列结合与条件概率相乘就能得到每一个候选序列的生成概率，然而取概率最高的作为标注结果即可。而CRF正好反过来，CRF通过发掘词语本身的特征（如长度，大小写，匹配特定词表等，也可以包括词语本身），把每个词语转化成为一个一维特征向量(vector)，然后对于每个特征计算特征到词性的条件概率，这样每个词语对候选词性的条件概率即为所有特征条件概率的加和。比如我们假设特征向量只有两个，且P ( ”词语长度>3" --> 名词词性）的概率为0.9， P("词语位于句子末尾“ --> 名词词性）概率为0.4，且一个词恰好满足这两个特征，则其为名词的条件概率为 (0.9 + 0.4) / 2 = 0.65. 这样，CRF根据这个条件转移数值再结合词性的马尔科夫特性，就可以使用与HMM类似的方法寻找最优的词性标注序列了。

1. CRF模型训练过程（简画）

