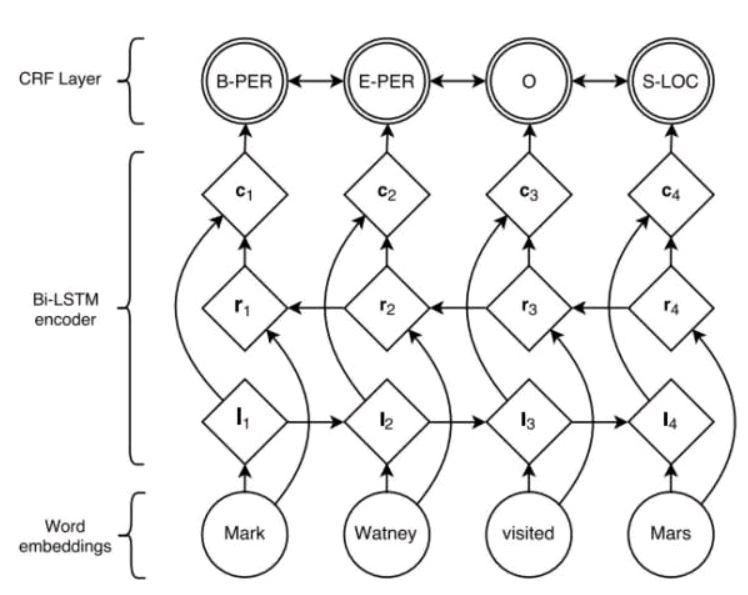
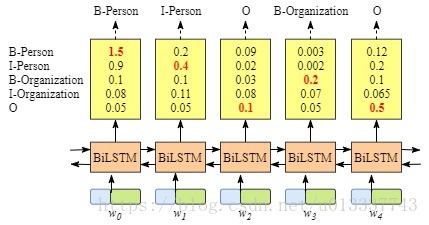
BiLSTM+CRF

整体架构图：



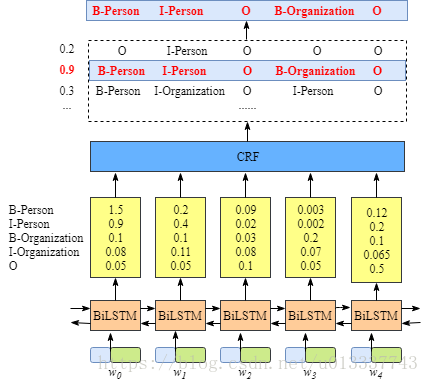
BiLSTM+Softmax：

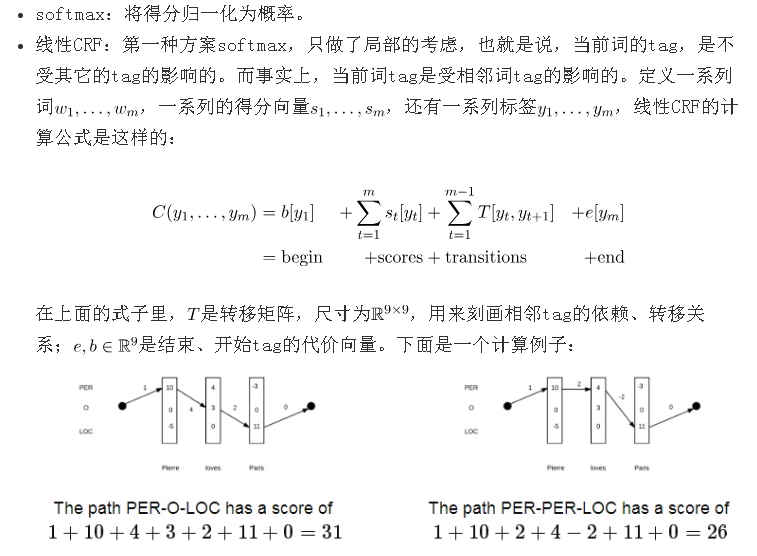


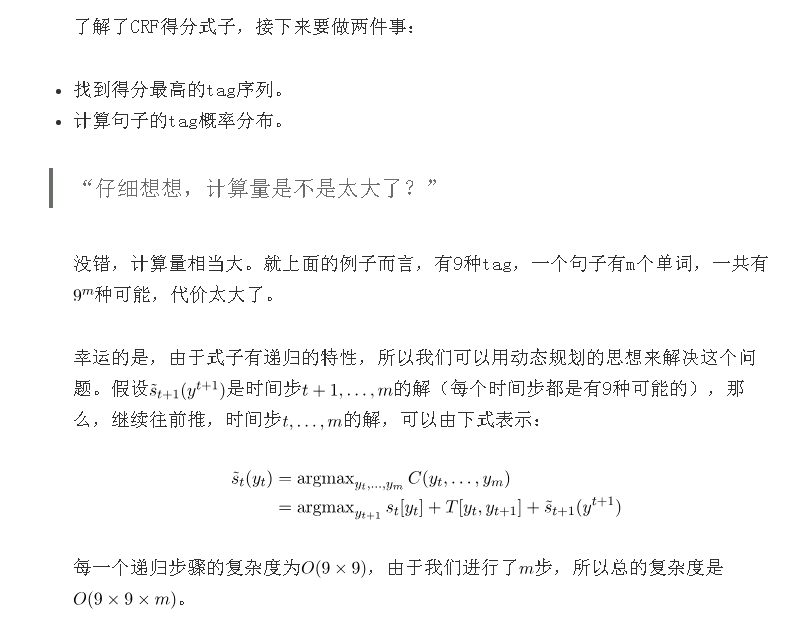
双向lstm后接一个softmax层，输出各个label的概率。那为何还要加一个crf层呢？

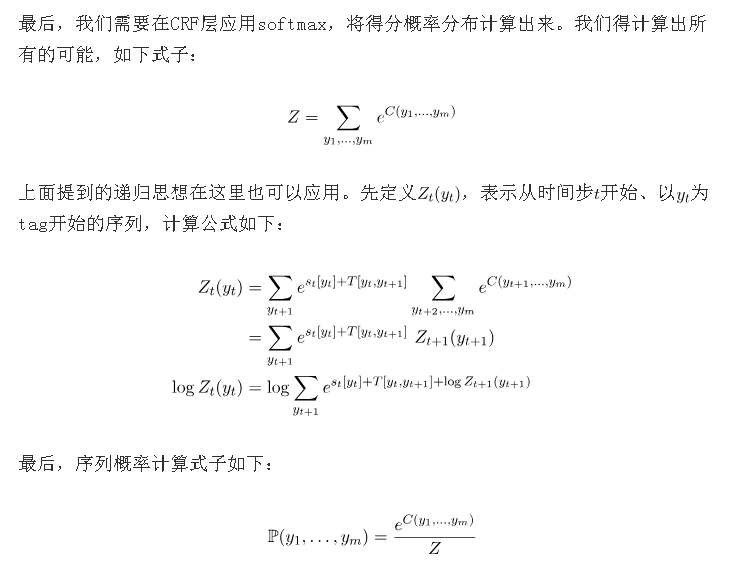
我的理解是softmax层的输出是相互独立的，即虽然BiLSTM学习到了上下文的信息，但是输出相互之间并没有影响，它只是在每一步挑选一个最大概率值的label输出。这样就会导致如B-person后再接一个B-person的问题。而crf中有转移特征，即它会考虑输出label之间的顺序性，所以考虑用crf去做BiLSTM的输出层。

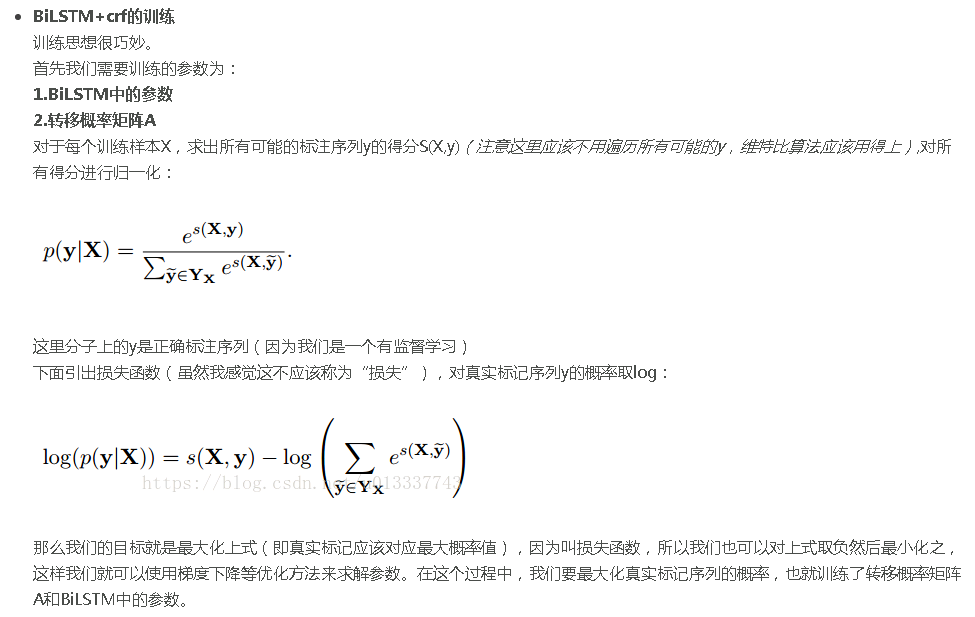
BiLSTM+CRF：

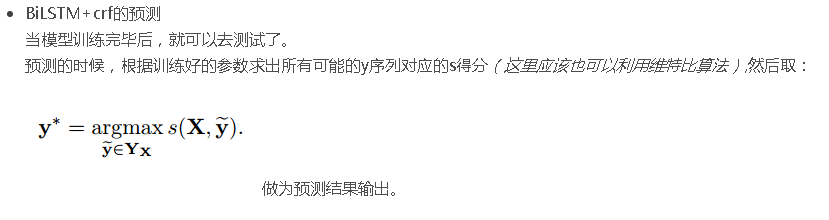












参考资料：

<https://blog.csdn.net/bobobe/article/details/80489303>

<https://www.lookfor404.com/%E7%94%A8%E5%8F%8C%E5%90%91lstmcrf%E5%81%9A%E5%91%BD%E5%90%8D%E5%AE%9E%E4%BD%93%E8%AF%86%E5%88%AB%E9%99%84tensorflow%E4%BB%A3%E7%A0%81-ner%E7%B3%BB%E5%88%97%EF%BC%88%E5%9B%9B/>