题目一，首先利用读取函数将三个文件中的信息读入程序中，分别有transition和emission的数量，利用这些数据求得transition概率和emission的概率，形成两个概率矩阵。

利用维特比算法针对每个要求的观测序列，求出每一个state对应的最大可能的前序state和此时emission出观测序列中symbol的概率。利用动态规划求出观测序列中每个symbol的最大概率state和概率。

利用求得的概率，寻找最优路径，即可找到观测序列对应的最大概率state序列。

题目二，和题目一类似，只是每次动态规划时，保存前k个最大state和概率，最后利用图的遍历求出最大的k个序列

题目三，

首先，针对训练数据中可能缺失的symbol和state，我们需要使用平滑技术，在前面的问题中，我们使用的是add-1 smoothing，而这种平滑技术最大的问题在于，转移了太多概率，

我们尝试了不同的平滑方式，最简单直接的方式在于，将转移概率中的1改成更小的值，比如0.1或者0.000001等，但是经过我们尝试之后，发现这样的效果并不明显，甚至是负优化。

经过大量的数据分析，我们发现在训练数据中U开头的symbol会对数据预测产生非常大的影响，因此我们将这样的序列第一个状态概率设为0，改变了数据本身具有的偏差，果然，这样的方法比较简单，而且取得了明显提升效果。