1、实验题目：基于HMM的拼音转汉字程序(1)

2、实验目的

理解、掌握隐马尔可夫模型，N元语法等自然语言处理的基本思想、算法，并将其应用于从汉语拼音到汉字的自动转换过程。

假定：拼音串中已经用空格进行了分隔，如“wo ai wo jia”

3、实验项目内容

（1）对训练语料及相关资源进行预处理；

（2）通过学习算法，训练HMM模型；

（3）使用合理的数据平滑方式，解决N元文法中的数据稀疏问题；

1、实验题目：基于HMM的拼音转汉字程序(2)

2、实验目的

理解、掌握隐马尔可夫模型，N元语法等自然语言处理的基本思想、算法，并将其应用于从汉语拼音到汉字的自动转换过程。

3、实验项目内容

（1）利用HMM模型和维特比算法，实现从任意拼音到汉字的自动转换，要求程序能对转换中的歧义进行正确识别和处理（如 fa piao bao xiao中bao xiao可能为报销、爆笑、报效等）。

（2）利用给定测试集，评价上述程序的转换准确率。