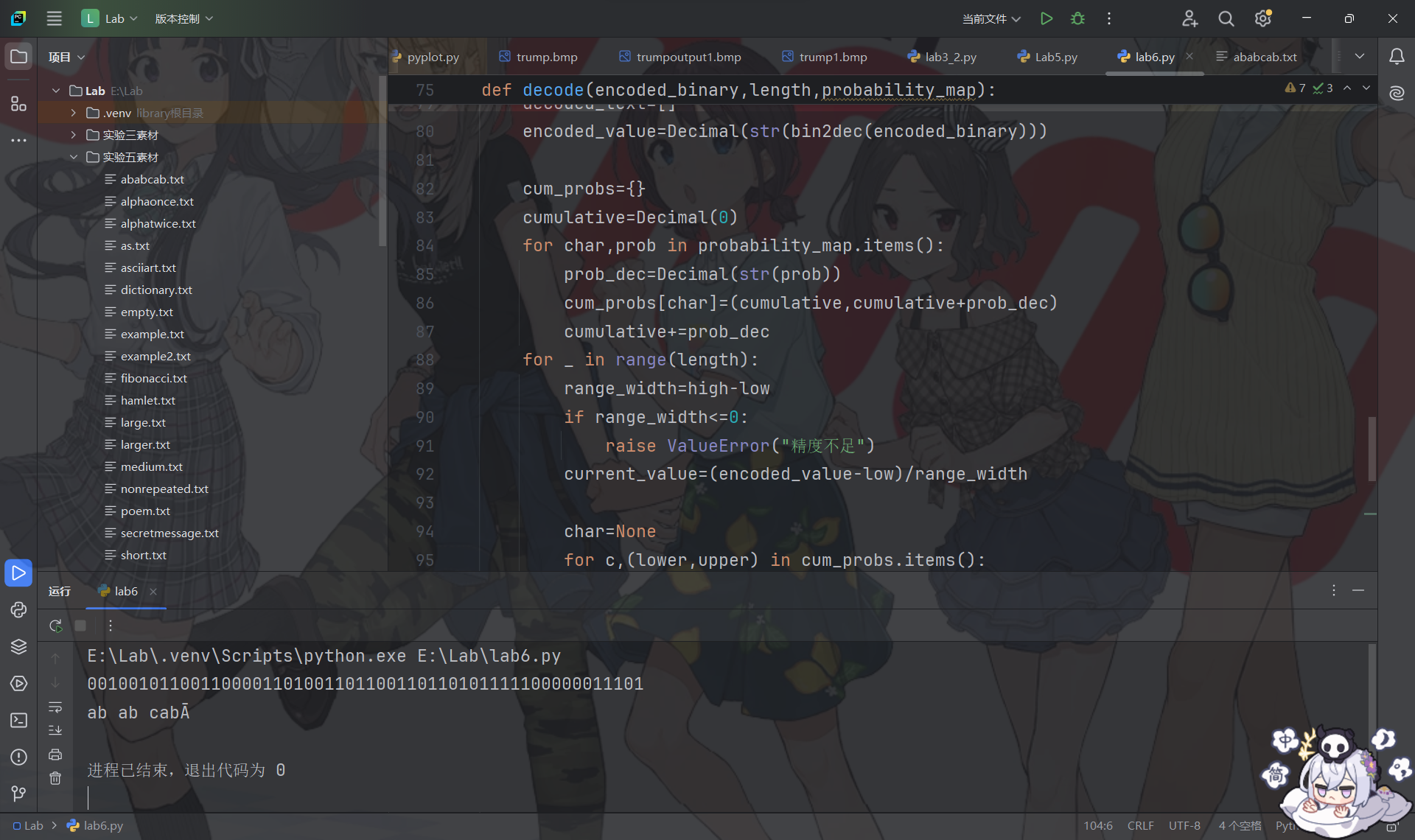
**《多媒体技术》实验报告**

**1.完成本次实验任务的思路**

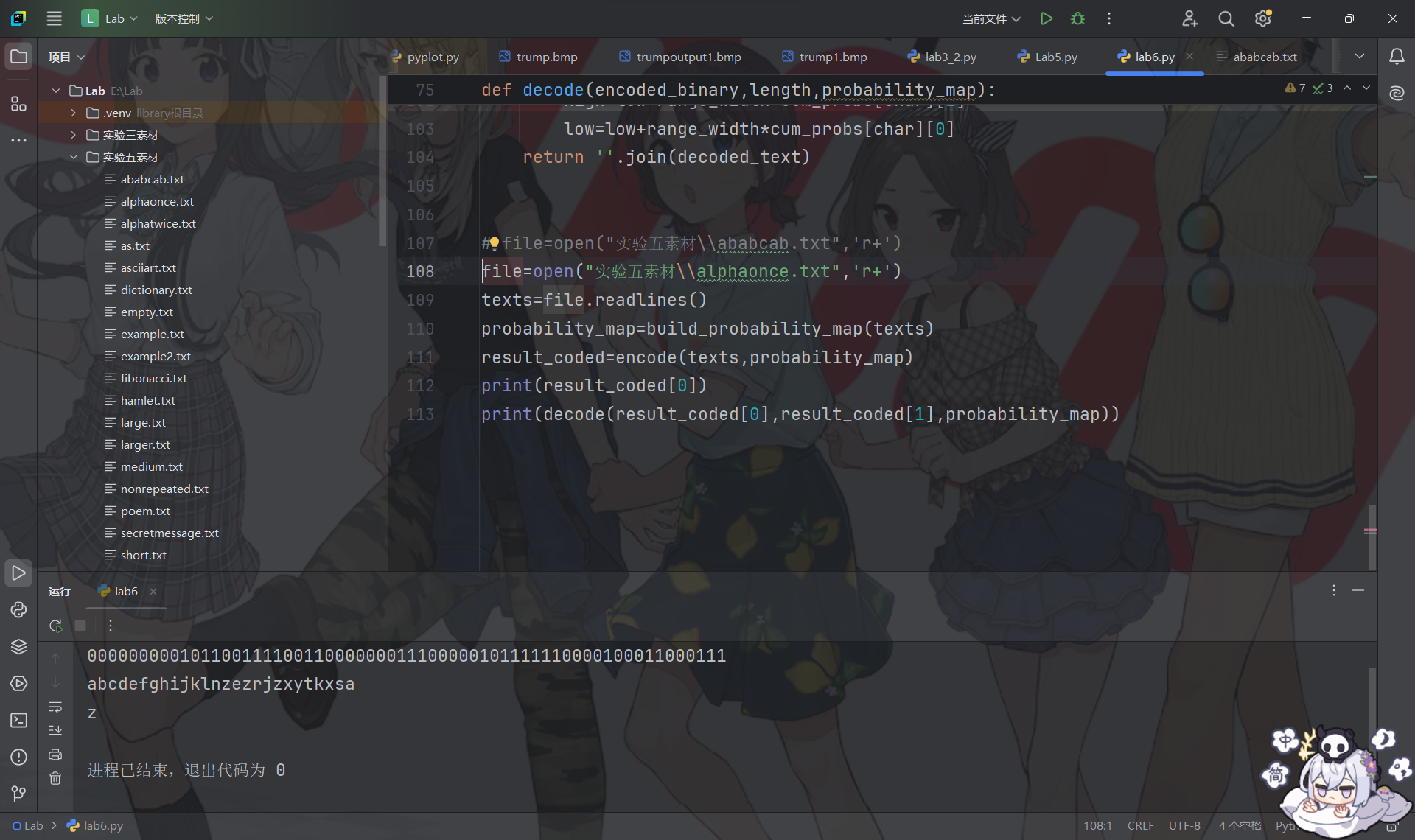
先统计整个文本文件的字符，将每个字符的频率放入字典中。根据频率，将[0,1)区间按比例分配给各个字符，扫描文本字符串，每个字符根据其分配到的区间比例不断缩小编码区间，同时统计总长度，将编码后的字符串转化为二进制后返回。在解码时，先将二进制的编码转化为小数，循环总长度次，每次找到缩小后区间包含编码的字符作为解码后的字符。

**2.运行程序截图和简要说明**

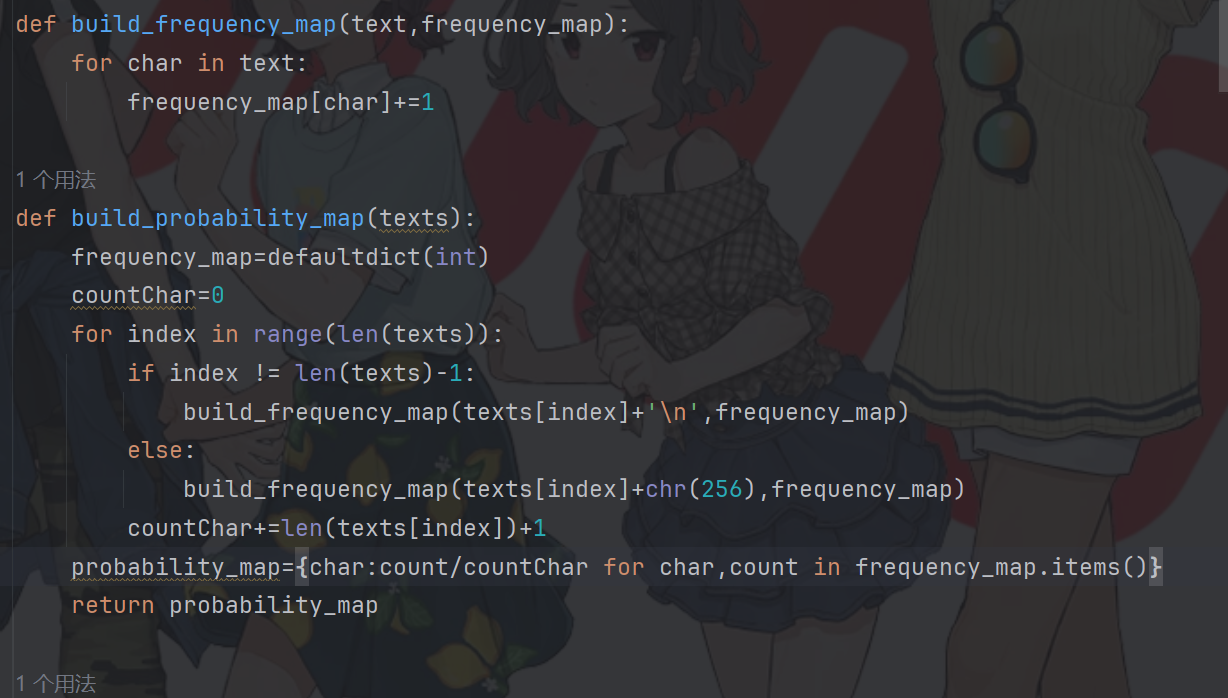
当字符数量不多时，正常运行。



当字符数量过多会则无法正常运行。



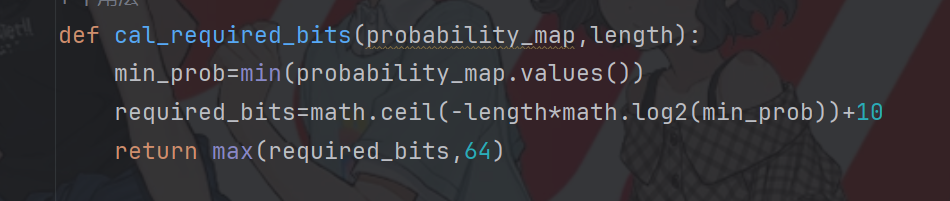
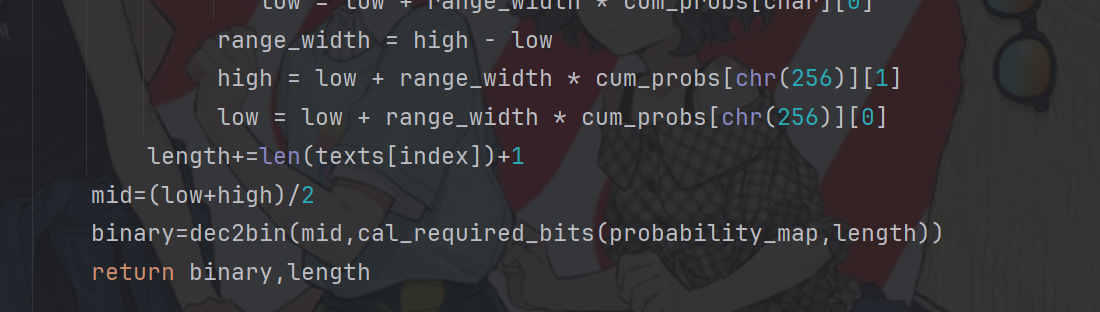
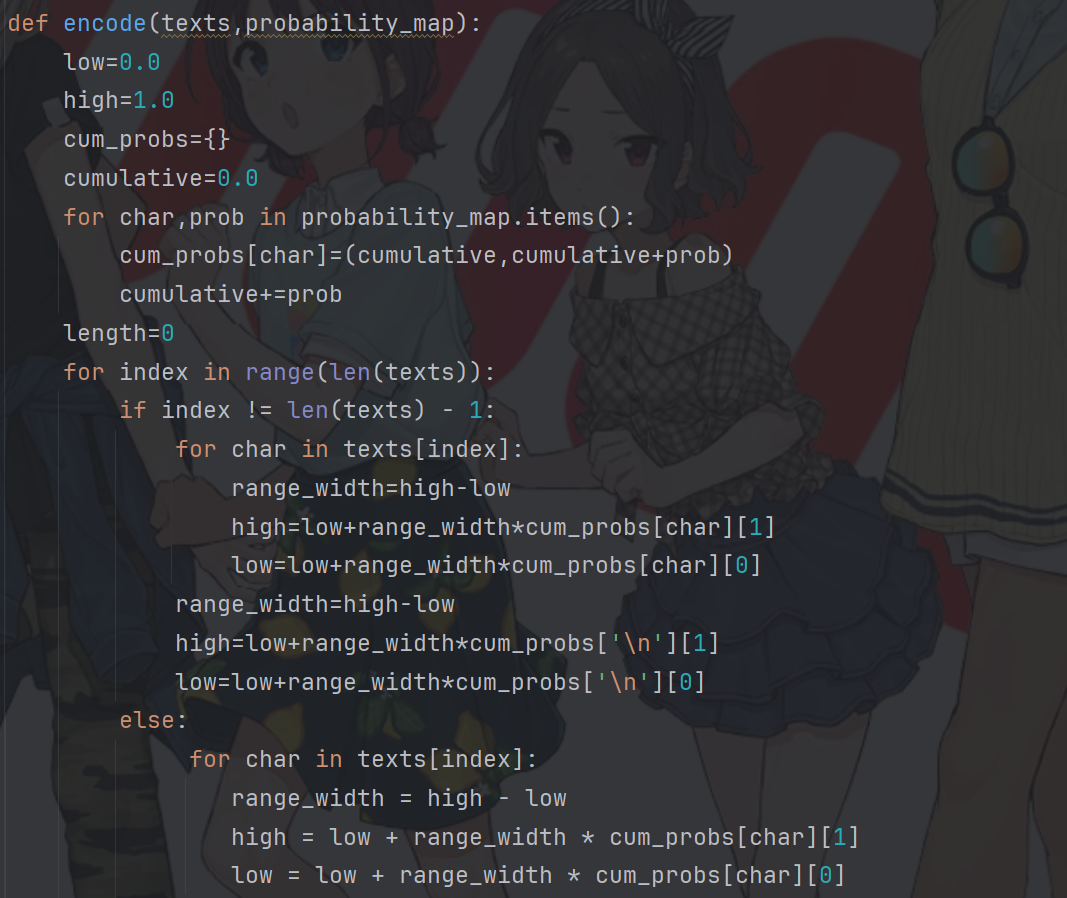
**3.核心代码展示和分析**



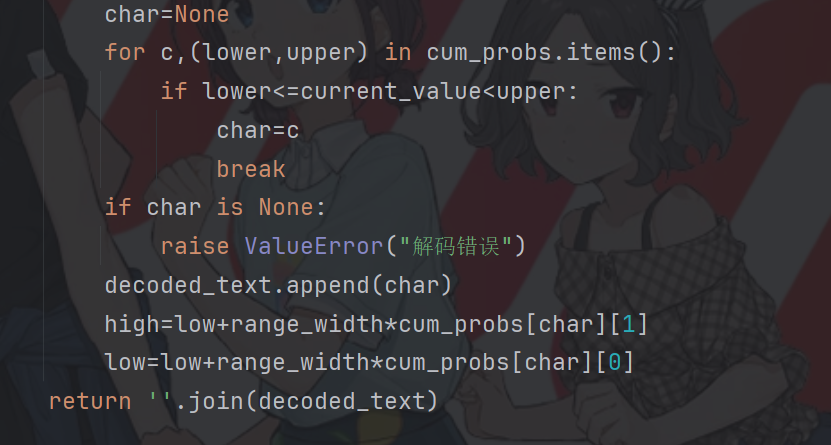
先通过遍历字符串，调用build\_frequency\_map统计每个字符的频数，同时统计总共的字符串长度，每个频数除以字符串长度即为频率。

十进制转二进制，由于值肯定是在[0,1)区间上的，直接考虑小数部分即可，每次乘以2，取整数部分，直到变为0或到达精度限制。

二进制转十进制，同样只考虑小数部分，求每个位数乘以2的-（下标+1）次方之和。



先将每个字符的区间分配给cum\_probs，遍历求出编码后，取区间的中间值为输出结果，cal\_required\_bits求得转化为二进制至少需要的精度，至少为64位，然后转化为二进制，输出二进制转化后的结果和字符串长度。

为了增大精度范围，使用Decimal作为参量，在循环过程中当精度不够导致range\_width<=0或者找不到对应的字符时会进行报错。将二进制的编码先转化成十进制小数，再进行后续的循环过程，获得解码后的字符串。