

# 北京科技大学实验报告

---

## 实验名称:

哈利波特文章检索

## 实验目的:

1. 将给定的 7 个 txt 文件读入
2. 对输入的字符串进行查找并将查找到的结果编号并输出其出现的页码, 章节, 书名
3. 指定查找某个编号出现的前后文

## 实验仪器:

计算机: 联想小新 Air-14ARE 2020

CPU: 2.10GHZ AMD Ryzen4600U

内存: 16GB 3200MHz DDR4

硬盘: 512GB

显卡: AMD Radeon Graphics

操作系统: Windows 10

编译器: MSVC 16

集成开发环境: Visual Studio Enterprise 2019 16.7.4

## 实验原理:

首先, 需要读入 txt 中的内容。使用 `ifstream` 函数打开 txt 文件, 这里我选择的是 STL 容器的 `vector` 存储内容, 长度可以随意添加, 且可以使用下标查找。由于题中要求输出页码, 章节, 书名, 所以需要对每一个单词记录页码, 章节, 书名。我是以行为单位, 定义一个结构体, 储存某一行的内容, 页码, 章节, 书名。对 txt 中的内容逐行读入, 如果是一个数字, 就说明是页码, 如果包含 `chapter` 就说明是章节编号, 其余就是书中的内容, 都与其之前的页码和章节相同, 加入到定义的 `vector` 变量的末尾。内容用 `string` 类型存储, 对于字符串匹配较为方便。

所有内容读取完之后, 接下来就是查找了。查找使用 `string` 的 `find` 函数即可, 会返回查找到的下标, 如果找不到就会返回 `nops`。再定一个 `vector` 数组, 储存

查找到的下标，以便之后根据标号查找上下文。

最后按照 **vector** 数组中的存储顺序输出下标加 **1**，页码，章节，书名即可。  
如果要查找某一个的上下文，那么输出其下标+1，下标，下标-1 对应的内容即可。

另外，我还制作了查找程序的界面，可以实现使用鼠标点击进入想要的功能。  
使用了 **EasyX** 图形库实现，我是以背景为分隔，不同的背景对应着不同的操作，实现功能的转换。主要的操作主要有鼠标点击切换，指定位置输出，用户输入信息。然后其余的就使用图形库中的函数均可实现。

### **实验内容与步骤：**

1. 思考怎么实现检索操作
2. 将要求查找的 **txt** 文件下载下来
3. 网上查找相关资料
4. 编写简易的查找程序
5. 测试查找程序
6. 使用图形库完善程序
7. 测试程序
8. 对程序进行优化

### **实验结果与分析：**

测试用例 1：查找 **Tom** 并查看对应序号为 **53** 的详细信息

运行截图 1：

人名/地名检索

## 点击查询详细信息

序号	人名/地名	页码	章节	书名
53	Tom	267	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
54	Tom	268	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
55	Tom	269	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
56	Tom	269	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
57	Tom	269	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
58	Tom	274	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
59	Tom	274	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
60	Tom	275	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
61	Tom	275	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
62	Tom	277	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
63	Tom	277	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
64	Tom	277	13	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets
65	Tom	360	17	J. K. Rowling - HP 2 - Harry Potter and the Chamber of Secrets

[上一页](#)
[重新输入查找信息](#)
[下一页](#)
[返回](#)

人名/地名检索

Dumbledore did not press her, though Harry could tell that he was interested. She took yet another gulp of gin and her rosy cheeks grew rosier still. "Billy Stubbs's rabbit . . . well, Tom said he didn't do it and I don't see how he could have done, but even so, it didn't hang itself from the rafters, did it?"

[返回](#)

Tom 在书中出现过，可以检索到

测试用例 2：查找 Xiaoming

运行截图 2：



Xiaoming 在书中为出现，检索不到

初始载入页面



帮助



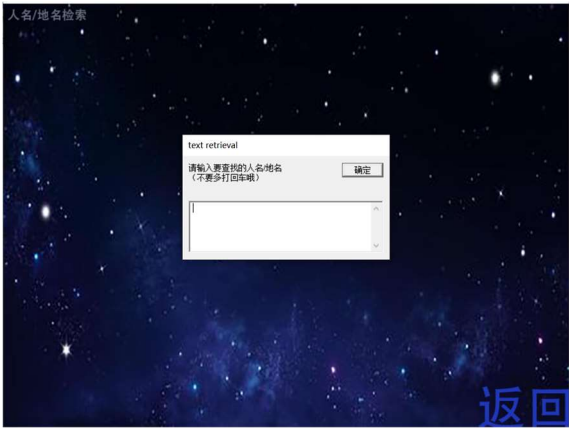
功能选择



关于



检索



所有结果



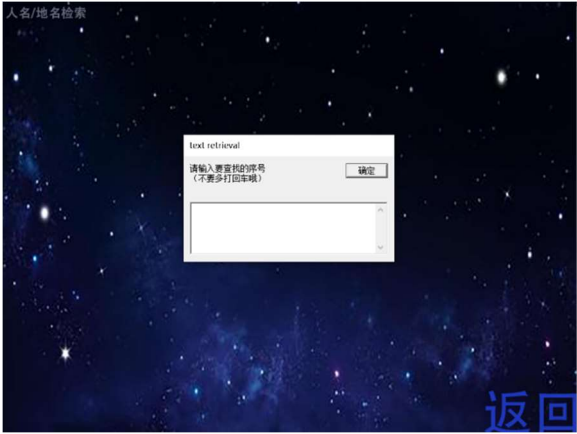
未检索到



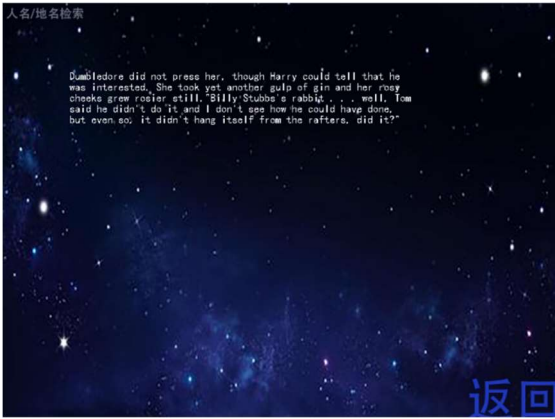
翻页过页



详细信息编号输入



详细位置显示





## 实验结论与感想

可以通过编写程序来实现对文本内容的检索。

通过这次的实验,学习到了好多之前不会的东西,大大提升了我编程实践的能力。在学习过程中,自己查找资料,查找相关函数的用法,提高了自学能力。感觉这种有实际作用的程序,相较于算法的程序更加难编。需要有对界面的设计,还要有对用户不合理操作的调整,比如用户输入对结果进行翻页时,可能翻到第一页还往前翻,这样就会使程序中的数组越界,导致程序崩溃,所以在编写程序时需要对用户可能出现的各种不合理操作进行预防,并提示用户进行正常操作。程序需要不断测试,尽可能将所有可能的操作测试一遍,以保证程序总是可以正常运行。