**北京科技大学实验报告**

学院： 计算机与通信工程学院 专业： 计算机类 班级： 计2401

姓名： 孙煜欣 学号： U202442404 实验日期： 2025年 5 月 26 日

**实验名称：**

哈利波特检索

**实验目的：**

1. 运行程序前，会先弹窗提示用户读入了几本书，方便后续查找
2. 当用户输入特定人名或者地名时候，程序会自动从文本中查找并显示。
3. 当用户点击查找结果的某一行时，会显示该字符的前后一部分内容，所查找的字符也会用中括号括起来方便查看

**实验仪器：**

计算机：小新16GT

CPU：3.1GHz IntelCore i5

内存：8GB 2133MHz LPDDR

硬盘：512 GB

显卡：Intel Iris Plus Graphics 650 1536 MB

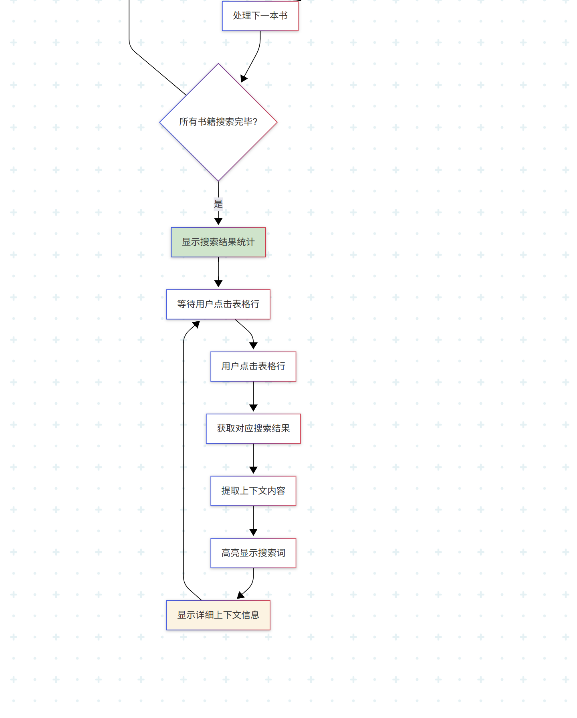
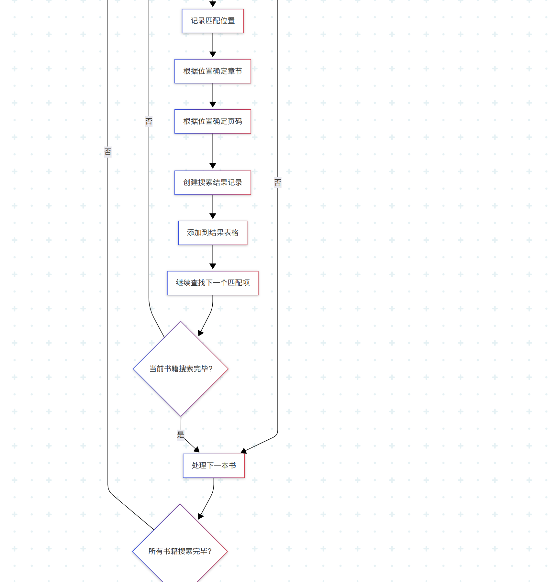
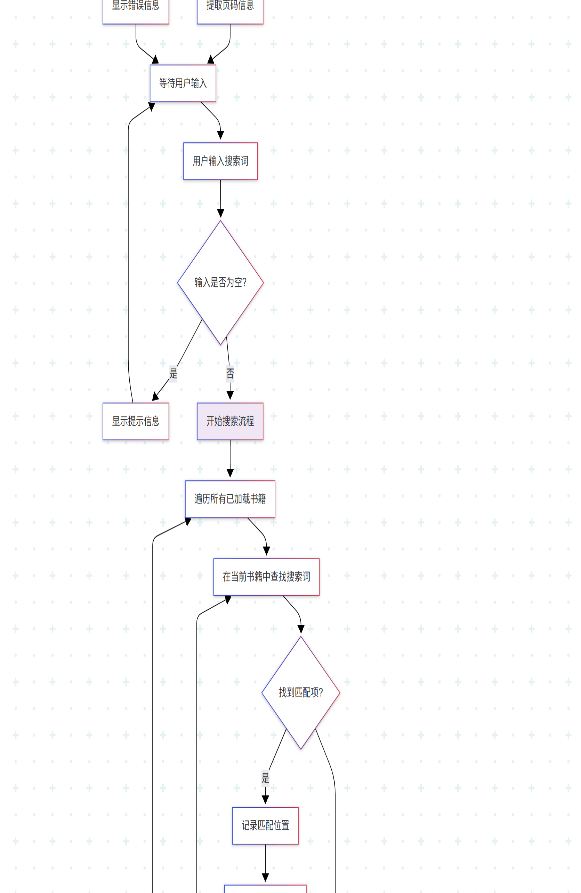
操作系统：Windows 11

编译器：gcc (x86\_64-posix-seh-rev3, Built by MinGW-W64 project) 11.2.0

**实验原理：**

**通过将多个文本文件预先加载到内存中，结合正则表达式识别章节标题和页码位置，实现了高效的全文检索功能，用户输入检索词后，系统会遍历所有已加载的文本内容进行匹配，并根据匹配位置自动判断所属章节和页码，最终以表格形式展示检索结果，同时支持点击查看详细的上下文信息，**

**图示

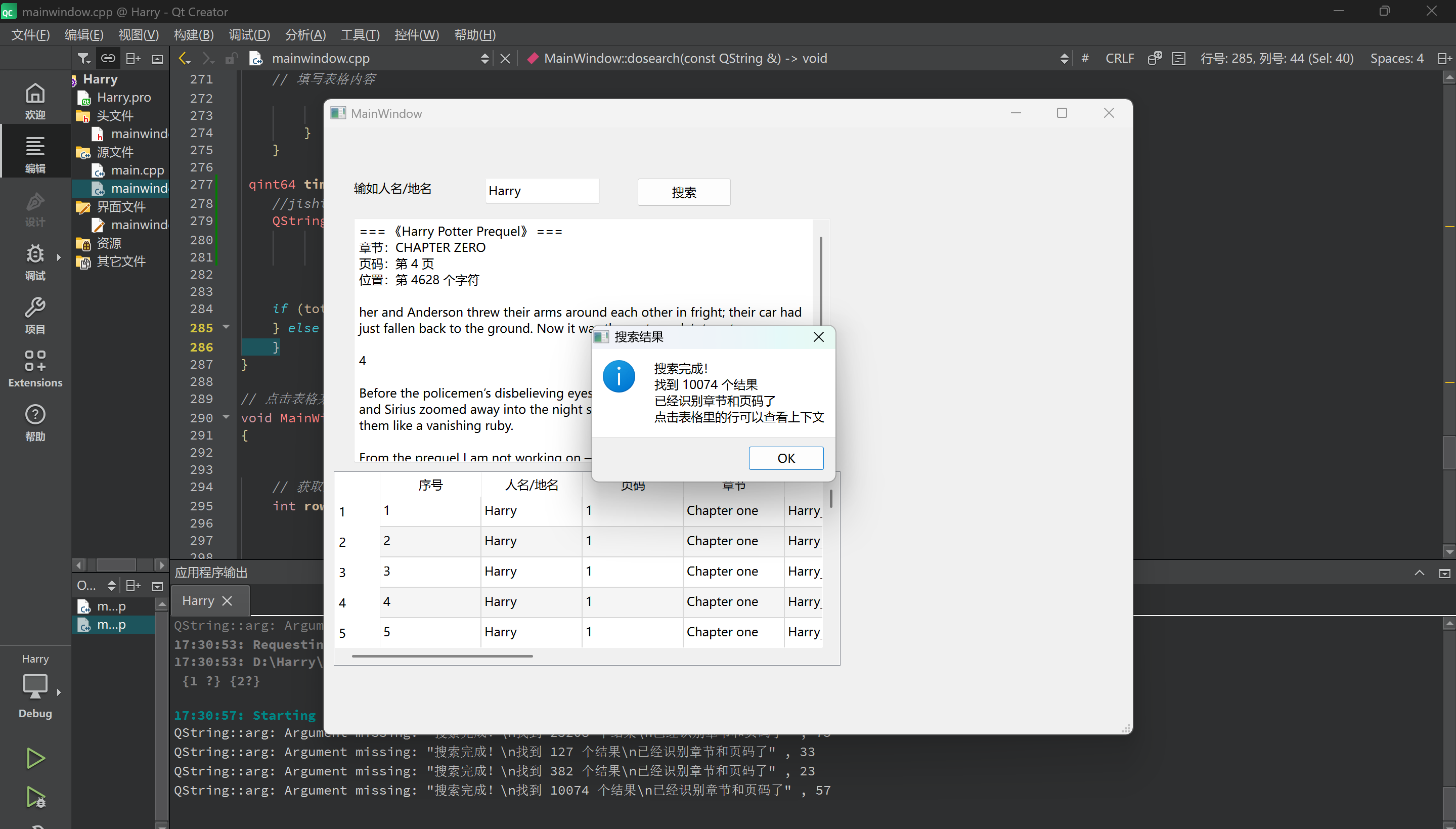
AI 生成的内容可能不正确。**

**实验内容与步骤：**

程序实现主要分为四个步骤：首先是系统初始化阶段，创建主界面并设置表格列标题，同时初始化书籍文件列表并批量加载所有文本内容到内存中；其次是文本预处理阶段，使用正则表达式从已加载的文本中提取章节信息和页码标记，；然后是检索执行阶段，接收用户输入的检索词后在所有文本中进行全文匹配，对每个匹配结果计算其在文本中的字符位置，并根据位置索引确定对应的章节和页码信息；最后是结果展示阶段，将检索结果以表格形式显示给用户，当用户点击某个结果时提取该位置的前后文内容并高亮显示检索词，完成整个检索和查看流程。

**实验结果与分析：**

测试用例1：Harry

运行截图1：表格

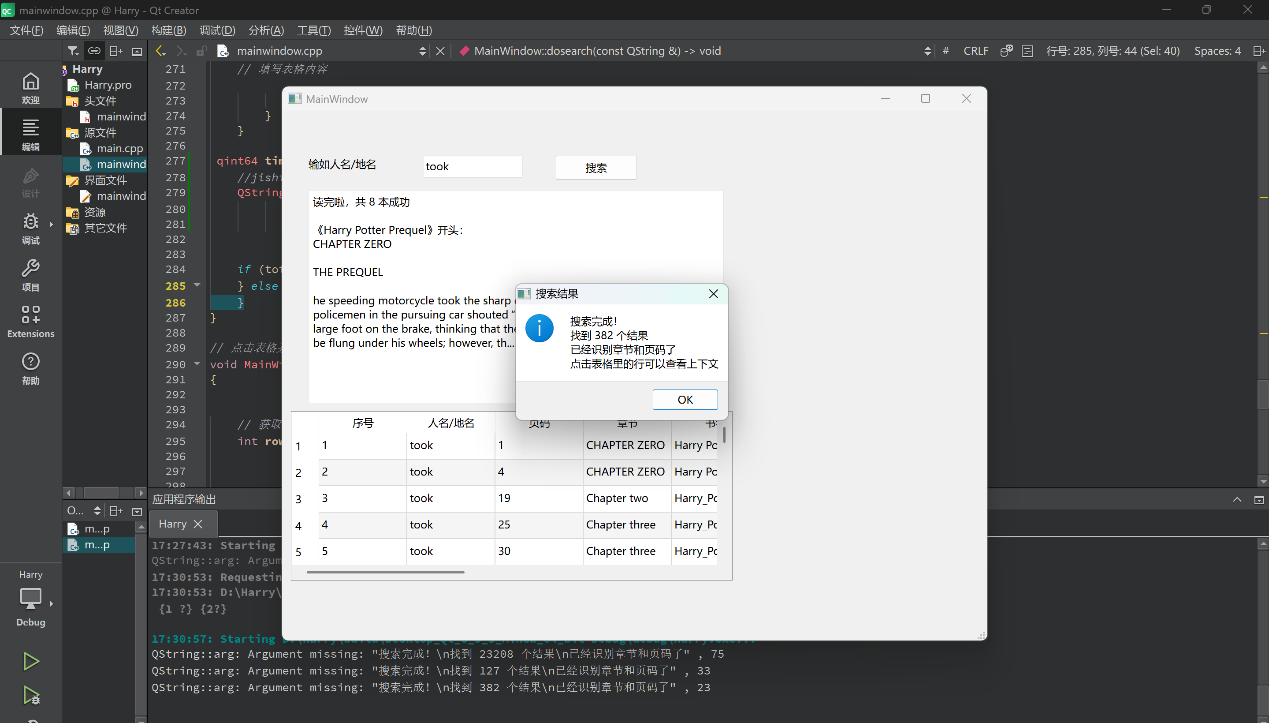
AI 生成的内容可能不正确。

当用户输入harry后，点击搜索按钮后查找出来10074个匹配项，之后会打印出搜索所用时间（如图一下面的“57”ms）之后点击第二行时，跳转到第二个Harry出现的地方，并显示出其前后的一段内容

测试用例2：took

运行截图2：表格

AI 生成的内容可能不正确。

当用户输入took后，点击搜索按钮后查找出来382个匹配项，之后会打印出搜索所用时间（如图二下面的“23”ms），之后点击第二行时，跳转到第二个took出现的地方，并显示出其前后的一段内容

**实验结论与感想**

在完成这个多文件搜索关键字的项目后，我对编程的作用有了更深入的理解。刚开始的时候，我觉得处理多个文件比较困难，不知道从哪里下手，但通过查阅资料和一步步调试，我慢慢掌握了文件读取、字符串处理等基本技能。这个实验让我体会到解决问题时的成就感。同时，我也意识到写程序不能光靠想法，更需要细心和耐心。作为大一学生，这次项目让我对编程更加有兴趣，也为以后的学习打下了基础。