

作业2 实验说明文档

练习完成方式：独立完成

保证是自己动手做出来的

请不要直接搜索或者讨论答案本身，鼓励交流讨论非答案内容

实验源码：仔细阅读此文档，编写程序完成指定功能，并运行测试通过，将修改过的源文件上传 [Gitee](#) 和学习通作业

实验文档撰写：要求使用 markdown格式书写，放在源码文件夹下面，推荐使用 [Typora](#)

实验目标

- 掌握函数指针的使用；
- 复习散列表的编程；
- 掌握链表的基础操作；
- 读写文件操作；

训练过程

- 根据老师提供的文档资料，学习散列表的基本使用方法；
- 编程实现具体要求中的所有功能；

训练过程

步骤1

user_login_short.txt文件中包含了10万条某网站的登录记录，文件中每一行的格式为：

<用户名,登录日期>,如：

Nehru,2016-01-01

Dane,2016-01-01

Walter,2016-01-01

Gloria,2016-01-01

.....

要求在一个工程文件中实现如下功能：

1. 实现完整的file_read_ht()函数，读取每一行的人名信息，构造散列表，使用链地址法处理冲突，定义数组指针存放数据信息（人名，登录次数）。其中包括lookup(), find()函数编写。
2. 实现完整的file_write_ht()函数，将散列表信息写入output.txt文件中

步骤2

使用散列表实现查找功能。

人名字符串作为关键词构造散列函数，假定关键词key包含n个字符

$a_0a_1...a_{n-1}$ ：

$$h(key) = \sum_{k=0}^{n-1} key[k] \times Base^k \mod TableSize$$

其中Base 取 31 或者35， TableSize取1000。

(1) 实现完整的search_ht(), 查找输入的人名, 显示登录信息。

(2) 实现完整的cleartable(), 销毁散列表。

完成上述功能的过程中, 要求:

- 在gitlab上完成实验, 在编写调试过程中必须向gitlab提交, 并使用commit标注进度情况, 完成情况的检查以gitlab上的提交log为准, 只有程序最终结果而无提交过程视为抄袭。在答辩过程中, 老师要看log里面的提交时间线, 以证明程序的修改和迭代过程。
- 必须撰写实验文档, 将程序编写过程记录在文档中, 并将实验文档与hashtable.c源文件一起提交到学习通。

步骤3

user_login.txt 文件中包含了100多万条登录记录, 比较不同处理量的查询运行时间。将待查询元素大小分别设置成总数量的 1/8、1/4、1/2、1倍数, 统计执行时间, 粗略了解散列查找数据大小与所需时间之间的关系。在<https://www.desmos.com/calculator> 上绘制图形进行对比。