



数据结构与算法实验

实验九：散列表应用

彭振辉

中山大学人工智能学院

2023年秋季学期



- 散列表算法的实现以库文件方式实现，不得在测试主程序中直接实现；
- 程序有较好可读性，各运算和变量的命名直观易懂，符合关键工程要求；
- 程序有适当的注释。

正确性：

- 20分 = 10 (task1) + 10 (task2)
- 助教会检查你在每个任务中输入测试用例得到的结果截图，如全部用例都得到正确的输出，该task得满分，否则每错一个用例扣3分，三个用例都错了则扣10分。
- 助教会检查**所有**同学代码，检查是否有抄袭或未实现的情况。如有，相应的task抄袭者和被抄袭者皆不得分。

时效性：

- 当堂课结束前提交不扣分
- 当天结束前提交扣2分
- 第二天结束前提交扣4分，第三天结束前提交扣6分，第四天结束前提交扣8分，第五天及以后提交扣10分。



提交格式和方式

格式：

- 请把“第九周作业”文件夹（需存放你写入代码的**cpp**文件、编译生成的**exe**文件，和包含“整体实现参考 + 2-3句简要体会（如教训、思路、拓展应用等） + 所有测试用例结果的截图”的**pdf/doc**文档）压缩成一个文件，命名为：**学号_姓名_第九周作业.zip**



210235_彭振辉_
第三周作业.zip

提交方式（具体操作教程在最后几页）：

- 对分易提交（当堂课或者课后提交均可用该方式）

截图样例和其在文档中的标记



实验任务1/2：

输入1：

5 8 13 22 50 100 -1

3 11 77 101 220 -1

在文档中标记好：任务1_测试1

输入2：

88 -1

3 7 86 88 110 -1

在文档中标记好：任务2_测试2.jpg

截图方式：

Windows: Shift +
Command + S

或者用“截图工具”
这个软件

输入3：...

输入4：...

实验任务1



示例 : Input:
3
Aaa Bbb Ccc

Bbb Ccc Ddd

Aaa2 ccc Eee
is at Ddd@Fff

2
1 2
1 3
Output:
50.0%
33.3%

- 文档相似度计算：定义该相似度为两文档共有的单词数量占两文档所有的单词量的比例。在这个任务中，我们考虑 $10 \geq \text{字符length} \geq 3$ 的英文单词，对于 $\text{length} > 10$ 的单词，我们只考虑其前10个字母
 - 输入：先输入文档总数 $N (\leq 100)$ ，在左例中是3)。随后输入每个文档的内容(如Aaa Bbb Ccc)，在一行中输入 # 表示该文档输入结束。在输入N个文档后，给出查询总数 $M (\leq 104)$ ，随后M行，每行给出一对文件编号，以空格分隔。这里假设文件按给出的顺序从1到N编号
 - 输出：针对每一条查询，在一行中输出两文件的公共词汇量占两文件总词汇量的百分比，精确到小数点后1位。注意这里的一个“单词”只包括仅由英文字母组成的、长度不小于3、且不超过10的英文单词，长度超过10的只考虑前10个字母。单词间以任何非英文字母隔开。另外，大小写不同的同一单词被认为是相同的单词



实验任务1 测试用例

输入1：

```
2
This is a test to show ...
#
Not similar at all
#
1
1 2
```

输入2：

```
2
These two files are the same
#
these.two_files are the SAME
#
1
1 2
```

输入3：

```
2
This is a test for
repeated repeated
words.
#
All repeated words shall
be counted only once.
A longlongword is the
same as this
longlongwo.
#
1
1 2
```

说明：为了实验任务1可以得到**满分10分**，你需要提交**三个截图**，每张截图分别包含以上**输入用例和对应的输出**，我们会检查你每个用例是否得到正确的输出，**每错一个，扣3分，错三个，扣10分**。同时，我们会**查看所有同学的源代码或exe程序**，如并未对任务1做出具体实现，该任务1得分为0。

实验任务2



示例：

Input:

4 500

330106199010080419 499

110108198403100012 15000

120104195510156021 800

330106199010080419 1

4

120104195510156021

110108198403100012

330106199010080419

33010619901008041x

Output:

800

15000

1000

No Info

- 航空公司如南方航空会进行会员制（如明珠会员），会员飞行里程积分可以用来兑换机票折扣或升舱等服务。现有南方航空公司所有会员的飞行记录，要求实现根据会员身份证号码快速查询其里程积分
 - 输入：先输入飞行记录条数 $N(\leq 10^5)$ 和最低里程 $K(\leq 500)$ 。其中南方航空会将航程低于 K 公里的航班也按 K 公里积分。随后 N 行，每行给出一条飞行记录，格式为：18位身份证号码(空格)飞行里程。其中身份证号码由17位数字加最后一位校验码组成，校验码的取值范围为0~9和x共11个符号；飞行里程单位为公里，是(0, 15000]区间内的整数。最后给出一个查询条数 $M(\leq 10^5)$ ，以及 M 行要查的会员的身份证号码。
 - 输出：对每个查询人，给出其当前的里程累积值。如果该人不是会员，则输出No Info。每个查询结果占一行。



实验任务2 测试用例

输入1：

5 360

18239047518293471x 360

145623849102394520 359

188329457691023457 500

18284049572934512x 1500

328841092309457613 890

3

17732495891293794x

18239047518293471x

145623849102394520

输入2：

4 300

523129857234104861 0

533123891283984571 301

758392095736123518 5102

13317512389562138x 3092

3

52312985723410486x

523129857234104861

758392095736123518

输入3：

6 555

356183920456182911 599

15671855122983457x 8153

520103578391209571 12

66618204927512951x 3937

651239384756661233 1059

522578192857264231 2092

4

520103578391209571

520103578391209571

520103578391209561

651239384756661233

说明：为了实验任务2可以得到**满分10分**，你需要提交**三个截图**，每张截图分别包含以上**输入用例和对应的输出**，我们会检查你每个用例是否得到正确的输出，**每错一个，扣3分，错三个，扣10分**。同时，我们会**查看所有同学的源代码或exe程序**，如并未对任务2做出具体实现，该任务2得分为0。



对分易提交方法

<https://www.duifene.com/Home.aspx>

微信扫码登录对分易

The screenshot displays the Duifeni website's login page. The main banner features the text "对分易" (Duifeni) in large white characters, with the tagline "让教师更轻松 让学生更主动" (Make teachers' lives easier, make students more active) below it. A navigation bar at the bottom includes links for "首页" (Home), "上门培训" (On-site training), "关于我们" (About us), "新闻动态" (News), "活动专区" (Activity area), "注册" (Register), and "登录" (Login). An inset window titled "用户登录" (User Login) is overlaid on the right side, showing a login form with fields for "登录名" (Username) and "密码" (Password), a "记住密码" (Remember password) checkbox, and a "忘记登录密码" (Forgot login password) link. A "登录" (Login) button is at the bottom of the form. To the right of the form is a QR code with the text "手机微信扫一扫登录" (Scan with mobile WeChat to login) below it.



对分易提交方法

对分易

教学平台

欢迎您, 彭振辉 同学

[切换为老师](#)

备考神器

帮助

个人中心

首

退出课程

点进数据结构与算法的班级, 点击作业

班级学生

课程资源

作业

微信消息

考勤

讨论区

分组

成绩册

在线练习

课堂提问

调查问卷

投票

活动

教学评价

弹幕讨论

学生互测

互动直播

考试



对分易提交方法

首页 > 课程：数据结构与算法 > 作业列表

21级人工智能

点进相应的作业编号，去完成，把压缩文件上传成附件，提交

截止时间：2022-09-07 17:04

已交53人

去完成

单次作业

写作业

作业文本：

限制8000个字

0/8000

作业文件：

上传附件

单个文件大小限制为100MB，文件名称请不要超过100个字，视频文件请转成H.264编码的MP4格式。

☐ 视频保存后自动转为教师可以在线播放的格式（H264编码的MP4格式）

提交作业