



哈尔滨工业大学(威海)

Harbin Institute of Technology, Weihai

# 数据结构实验报告册

学 号: \_\_\_\_\_

姓 名: \_\_\_\_\_

班 级: \_\_\_\_\_

任课教师: \_\_\_\_\_

开课学期: 2020-2021 秋季学期

哈尔滨工业大学(威海)计算机科学与技术学院

实验一	实验二	实验三	实验四	实验五

# 目 录

实验一 线性结构 .....	1
实验二 树.....	7
实验三 图.....	13
实验四 查找.....	20
实验五 排序.....	27

# 实验一 线性结构

## ——一元多项式求导

### 1. 实验目的

熟练掌握链式线性表的基本操作。

### 2. 实验内容

设计函数求一元多项式的导数。

### 3. 实验要求

#### (1) 输入格式说明：

以指数递降方式输入多项式非零项系数和指数（绝对值均为不超过 1000 的整数）。数字间以空格分隔。

#### (2) 输出格式说明：

以与输入相同的格式输出导数多项式非零项的系数和指数。数字间以空格分隔，但结尾不能有多余空格。

#### (3) 样例输入与输出：

序号	输入	输出
1	3 4 -5 2 6 1 -2 0	12 3 -10 1 6 0
2	5 20 -7 4 3 1	100 19 -28 3 3 0
3	1000 0	0 0
4	-1000 1000 999 0	-1000000 999

## 4. 实验过程

### (1) 问题描述

(问题分析及功能描述)

## (2) 数据结构与算法设计

(逻辑结构分析+存储结构设计+关键算法思路+伪代码或流程图)

### (3) 程序实现

(函数说明+函数之间的调用关系+关键算法的实现代码)

#### (4) 实验结果

(运行截图+结果分析描述+遇到的问题 and 解决办法等)

## (5) 实验总结

(实验体会、学习收获、过程总结等)



## 实验二 树

### ——还原二叉树

#### 1. 实验目的

熟练掌握二叉树存储结构、遍历及应用。

#### 2. 实验内容

给定一棵二叉树的先序遍历序列和中序遍历序列，要求计算该二叉树的高度。

#### 3. 实验要求

##### (1) 输入格式说明：

输入首先给出正整数  $N$  ( $\leq 50$ )，为树中结点总数。下面 2 行先后给出先序和中序遍历序列，均是长度为  $N$  的不包含重复英文字母（区别大小写）的字符串。

##### (2) 输出格式说明：

输出为一个整数，即该二叉树的高度。

##### (3) 样例输入与输出：

序号	输入	输出
1	9 ABDFGHIEC FDHGIBEAC	5
2	15 cdefghijklmxyz cdefghijklmxyz	15
3	7 Abcdefg gfedcbA	7
4	1 a a	1

## 4. 实验过程

### (1) 问题描述

(问题分析及功能描述)

## (2) 数据结构与算法设计

(逻辑结构分析+存储结构设计+关键算法思路+伪代码或流程图)

### (3) 程序实现

(函数说明+函数之间的调用关系+关键算法的实现代码)

#### (4) 实验结果

(运行截图+结果分析描述+遇到的问题 and 解决办法等)

## (5) 实验总结

(实验体会、学习收获、过程总结等)

## 实验三 图

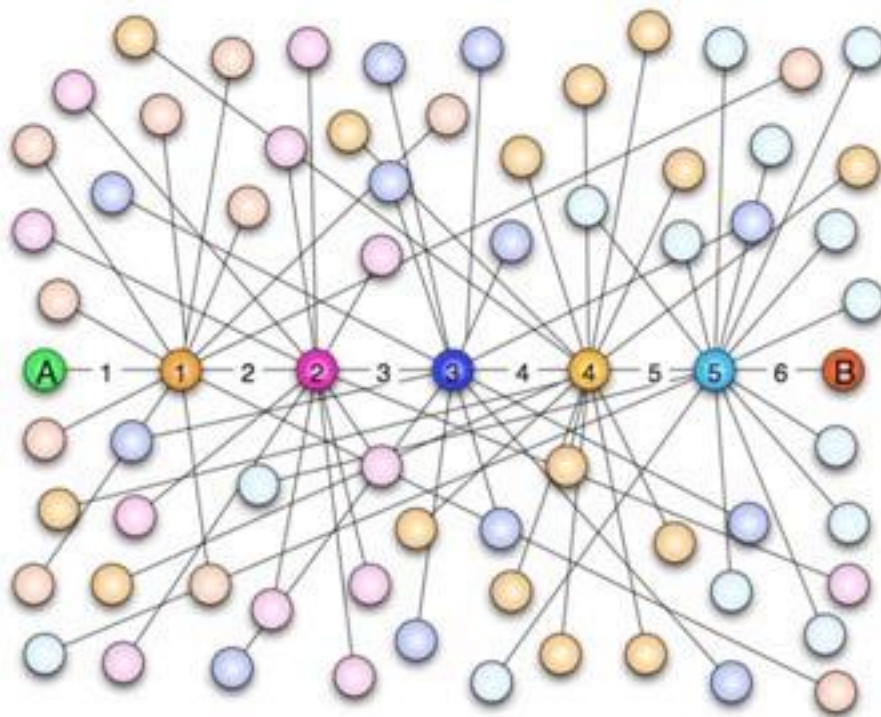
### ——六度空间

#### 1. 实验目的

熟练掌握图的存储结构及广度优先遍历方法。

#### 2. 实验内容

“六度空间”理论又称作“六度分隔（Six Degrees of Separation）”理论。这个理论可以通俗地阐述为：“你和任何一个陌生人之间所间隔的人不会超过六个，也就是说，最多通过五个人你就能够认识任何一个陌生人。”如下图所示。



“六度空间”理论虽然得到广泛的认同，并且正在得到越来越多的应用。但是数十年来，试图验证这个理论始终是很多社会学家努力追求的目标。然而由于历史的原因，这样的研究具有太大的局限性和困难。随着当代人的联络主要依赖于电话、短信、微信以及因特网上即时通信等工具，能够体现社交网络关系的一手数据已经逐渐使得“六度空间”理论的验证成为可能。

假如给你一个社交网络图，请你对每个节点计算符合“六度空间”理论的结点占结点总数的百分比。

#### 3. 实验要求

##### (1) 输入格式说明：

输入第1行给出两个正整数，分别表示社交网络图的结点数  $N$  ( $1 < N \leq 104$ ，表示人数)、边数  $M$  ( $\leq 33 * N$ ，表示社交关系数)。随后的  $M$  行对应  $M$  条边，

每行给出一对正整数，分别是该条边直接连通的两个结点的编号（节点从 1 到 N 编号）。

**(2) 输出格式说明：**

对每个结点输出与该结点距离不超过 6 的结点数占结点总数的百分比，精确到小数点后 2 位。每个结节点输出一行，格式为“结点编号: (空格) 百分比%”。

**(3) 样例输入与输出：**

序号	输入	输出
1	10 9	1: 70.00%
	1 2	2: 80.00%
	2 3	3: 90.00%
	3 4	4: 100.00%
	4 5	5: 100.00%
	5 6	6: 100.00%
	6 7	7: 100.00%
	7 8	8: 90.00%
	8 9	9: 80.00%
	9 10	10: 70.00%
2	10 8	1: 70.00%
	1 2	2: 80.00%
	2 3	3: 80.00%
	3 4	4: 80.00%
	4 5	5: 80.00%
	5 6	6: 80.00%
	6 7	7: 80.00%
	7 8	8: 70.00%
	9 10	9: 20.00%
		10: 20.00%
3	11 10	1: 100.00%
	1 2	2: 90.91%
	1 3	3: 90.91%
	1 4	4: 100.00%
	4 5	5: 100.00%
	6 5	6: 100.00%
	6 7	7: 100.00%
	6 8	8: 100.00%
	8 9	9: 100.00%
	8 10	10: 100.00%
	10 11	11: 81.82%



## 4. 实验过程

### (1) 问题描述

(问题分析及功能描述)

## (2) 数据结构与算法设计

(逻辑结构分析+存储结构设计+关键算法思路+伪代码或流程图)

### (3) 程序实现

(函数说明+函数之间的调用关系+关键算法的实现代码)

#### (4) 实验结果

(运行截图+结果分析描述+遇到的问题 and 解决办法等)

## (5) 实验总结

(实验体会、学习收获、过程总结等)

## 实验四 查找

### ——基于词频的文件相似度

#### 1. 实验目的

掌握倒排索引表的应用。

#### 2. 实验内容

实现一种简单原始的文件相似度计算，即以两文件的公共词汇占总词汇的比例来定义相似度。为简化问题，这里不考虑中文（因为分词太难了），只考虑长度不小于 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。

#### 3. 实验要求

##### (1) 输入格式说明：

输入首先给出正整数  $N$  ( $N \leq 100$ )，为文件总数。随后按以下格式给出每个文件的内容：首先给出文件正文，最后在一行中只给出一个字符“#”，表示文件结束。在  $N$  个文件内容结束之后，给出查询总数  $M$  ( $M \leq 104$ )，随后  $M$  行，每行给出一对文件编号，其间以空格分隔。这里假设文件按给出的顺序从 1 到  $N$  编号。

##### (2) 输出格式说明：

针对每一条查询，在一行中输出两文件的相似度，即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，例如“You”和“you”是同一个单词。

##### (3) 样例输入与输出：

序号	输入	输出
1	3 Aaa Bbb Ccc # Bbb Ccc Ddd # Aaa2 ccc Eee is at Ddd@Fff # 2 1 2	50.0% 33.3%

	1 3	
2	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	23.1%
3	2 This is a test to show ... # Not similar at all # 1 1 2	0.0%
4	2 These two files are the same # these.two_files are the SAME # 1 1 2	100.0%

## 4. 实验过程

### (1) 问题描述

(问题分析及功能描述)



## (2) 数据结构与算法设计

(逻辑结构分析+存储结构设计+关键算法思路+伪代码或流程图)

### (3) 程序实现

(函数说明+函数之间的调用关系+关键算法的实现代码)

#### (4) 实验结果

(运行截图+结果分析描述+遇到的问题 and 解决办法等)

## (5) 实验总结

(实验体会、学习收获、过程总结等)

## 实验五 排序

### ——模拟 EXCEL 排序

#### 1. 实验目的

熟练掌握快速排序的库函数调用。

#### 2. 实验内容

Excel 可以对一组纪录按任意指定列排序。现请编写程序实现类似功能。

#### 3. 实验要求

##### (1) 输入格式说明：

输入的第 1 行包含两个正整数  $N$  ( $\leq 105$ ) 和  $C$ ，其中  $N$  是纪录的条数， $C$  是指定排序的列号。之后有  $N$  行，每行包含一条学生纪录。每条学生纪录由学号（6 位数字，保证没有重复的学号）、姓名（不超过 8 位且不包含空格的字符串）、成绩（ $[0, 100]$  内的整数）组成，相邻属性用 1 个空格隔开。

##### (2) 输出格式说明：

在  $N$  行中输出按要求排序后的结果，即：当  $C=1$  时，按学号递增排序；当  $C=2$  时，按姓名的非递减字典序排序；当  $C=3$  时，按成绩的非递减排序。当若干学生具有相同姓名或者相同成绩时，则按他们的学号递增排序。

##### (3) 样例输入与输出：

序号	输入	输出
1	3 1 000007 James 85 000010 Amy 90 000001 Zoe 60	000001 Zoe 60 000007 James 85 000010 Amy 90
2	4 2 000007 James 85 000010 Amy 90 000001 Zoe 60 000002 James 98	000010 Amy 90 000002 James 98 000007 James 85 000001 Zoe 60
3	1 2 999999 Williams 100	999999 Williams 100

## 4. 实验过程

### (1) 问题描述

(问题分析及功能描述)

## (2) 数据结构与算法设计

(逻辑结构分析+存储结构设计+关键算法思路+伪代码或流程图)

### (3) 程序实现

(函数说明+函数之间的调用关系+关键算法的实现代码)



#### (4) 实验结果

(运行截图+结果分析描述+遇到的问题 and 解决办法等)

## (5) 实验总结

(实验体会、学习收获、过程总结等)

