



算法与数据结构设计任务书

2021-2022 第一学期

班级：B190308

指导教师：柯昌博

指导单位：南京邮电大学 计算机学院 计算机科学与技术系

编程语言：可自选，例如 Visual C++，C，JAVA 等

日期：第 9-10 周（2020.10.25-2021.11.05）

周次	周一	周二	周三	周四	周五
9	2-5 节 学六-102			2-5 节 学六-104 6-7 节 学六-202	3-5 节 学六-204 6-7 节 学六-204
10	2-5 节 学六-106	2-5 节 学六-104	1-2 节 学六-206	2-5 节 学六-202	3-5 节 学六-108

一、程序设计要求

程序设计（上机）的目的是通过实践，在程序设计方法、高级语言编程环境以及程序的调试和测试方面受到比较系统和严格的训练。通过实践提高同学们自己学习、查阅资料、独立分析问题和解决问题的能力。

在两周的课程设计中，要通过查阅相关资料、与同学讨论、老师辅导答疑，设计算法，上机调试并运行程序，完成实践报告。教师在最后对学生进行简单的答辩。

二、程序设计课题内容及课题分配

本次程序设计的题目包含 A 组和 B 组，每组各 1 道题目，每位同学需完成 A 组和 B 组中各一道题目，题目分配如下：

学号	姓名	A 组	B 组	学号	姓名	A 组	B 组
B17040815	刘熙铭	1	1	B19030820	陈励勤	3	3
B18030803	刘琪宇	2	2	B19030821	陈磊	4	4
B18030809	詹之琳	3	3	B19030822	金廷枢	5	5
B18030817	杨为渊	4	4	B19030823	郑宇轩	1	1
B19030801	万芳	5	5	B19030824	茹兴奥	2	2
B19030802	朱颖	1	1	B19030825	姚余都灵	3	3
B19030803	严碧钰	2	2	B19030826	顾奕	4	4
B19030804	吴君玉	3	3	B19030827	徐书颖	5	5
B19030805	张莹	4	4	B19030828	高梦	1	1
B19030806	胡璐	5	5	B19030829	曹议	2	2
B19030807	唐文洁	1	1	B19030830	彭贝宁	3	3
B19030808	覃程程	2	2	B19030831	韩科爽	4	4
B19030809	丁昀	3	3	B19030832	路丰源	5	5
B19030810	王昱杰	4	4	B19030833	鞠帆	1	1
B19030811	扎西顿珠	5	5	B19040830	宗一凡	2	2
B19030812	全家俊	1	1	B19060713	王耀宁	3	3
B19030813	刘一凡	2	2	B19080310	马宁远	4	4
B19030815	李兆维	3	3	B19080631	廖晨杰	5	5
B19030816	李想	4	4	B19111624	陈天祥	1	1
B19030817	杨潇	5	5				
B19030818	宋春宇	1	1				
B19030819	张清乐	2	2				

A 组（程序与算法设计题）

题目A1：大素数判定问题

编程实现大素数的随机生成；快速判定任意一个大数是否是素数；验证1000以内数的哥德巴赫猜想。

注：素数即只能被1和本身整除的正整数；哥德巴赫猜想为“任何一个大于6的偶数都可以表示成两个素数之和”。

题目A2：Top-N问题

假设全国有10000人参加物理竞赛，只录取成绩优异的前10名，并将他们从高分到低分输出。而对落选的其他考生，无需排出名次，问此种情况下，试编写尽可能快的算法解决Top-10问题？

扩展：如何在10亿数中找出前1000大的数？

题目A3：众数问题

给定含有 n 个元素的多重集合 S ，每个元素在 S 中出现的次数称为该元素的**重数**， S 中重数最大的元素称为**众数**。例如， $S=\{1, 2, 2, 2, 3, 5\}$ ， S 的众数是2，该众数的重数为3。要求对于给定的由 n 个自然数组成的多重集合 S ，计算 S 的众数及其重数。

题目 A4：AVL 树的判定问题

给定一个二叉树（存储结构采用二叉链表表示），试设计算法判断该二叉树是否为 AVL 树。

题目 A5：有向图的简单路径求解问题

给定一个有向图 G 两个顶点 a 和 b ，试编写算法求 a 到 b 的简单路径的数量，并分别输出最短的简单路径和最长的简单路径。

B 组（综合设计题）

题目 B1： 四则运算自测系统

（一）课题内容

设计一款四则运算的自测软件，可用于小学生的四则运算训练。

（二）课题要求

（1）训练题型包括填写结果、对错判断、选择题等。

（2）每次训练时，用户先进行选择，包括难度（如 10 以内、100 以内、1000 以内，带负数、带小数、运算数个数、是否允许带括号等）、题型（填写结果、对错判断、选择题）、题量（各类题目的个数）等，根据用户的选择，自动生成试题（10 道题目为一组）。

（3）用户在软件系统中答题，并提交。

（4）系统根据用户提交结果进行评判、给分，并进行分析。

（5）每次训练完后，试题、试题答案、用户答案、结果及分析应保留到相应的文本文件中，供用户查看。

【其他要求】

（1）变量、函数命名符合规范。

（2）注释详细：每个变量都要求有注释说明用途；函数有注释说明功能，对参数、返回值也要以注释的形式说明用途；关键的语句段要求有注释解释。

（3）程序的层次清晰，可读性强。

（4）界面美观，交互方便。

（5）设计时可考虑一定的趣味性。

（6）可考虑设计提供多个用户的登录功能。

题目 B2：代码分析器

（一）课题内容

读入 C 源文件，统计文件中的代码、注释和空行的行数以及函数的个数和平均行数，并利用统计信息分析评价该程序的风格。

（二）课题要求

- (1) 将程序文件按字符顺序读入源程序；
- (2) 边读入程序，边识别统计代码行、注释行和空行，同时还要识别函数的开始和结束，以便统计其个数和平均行数。
- (3) 程序的风格评价分为代码（函数平均长度）、注释（占总行数比例）和空行（占总函数比例）三个方面。每个方面分为 A, B, C:和 D 四个等级。等级的划分标准如下：

	A 级	B 级	C 级	D 级
代码(函数平均长度)	10~15 行	8~9 或 16~20 行	5~7 或 21~24 行	<5 或 >24 行
注释(占总行数比率)	15~25%	10~14 或 26~30%	5~9 或 31~35%	<5%或>35%
空行(占总行数比率)	15~25%	10~14 或 26~30%	5~9 或 31~35%	<5%或>35%

- (4) 测试时先对较小程序进行分析。当你的程序能正确运行时，对你的程序本身进行分析。

【其他要求】

- (1) 变量、函数命名符合规范。
- (2) 注释详细：每个变量都要求有注释说明用途；函数有注释说明功能，对参数、返回值也要以注释的形式说明用途；关键的语句段要求有注释解释。
- (3) 程序的层次清晰，可读性强。

题目 B3： 大整数算术运算器

（一）课题内容

实现一个简单的大整数算术运算程序。主要功能计算两个大整数（ ≥ 20 位）的加减法等算术运算，按指定的格式输出结果。通过此课题，熟练掌握字符串、格式输出、文件的各种操作，以及基本的计算算法思想的应用。

（二）课题要求

1. 基本要求

(1) 输入功能：能实现从键盘或文本文件输入代表大整数的字符串并用合适的结构存储，能实现输入整个表达式（含大整数和相应的运算符）。

(2) 输出功能：将算式及计算结果按照一定格式批量输出到屏幕及另一个文本文件中。

(3) 判断功能：能够对输入的表达式判断其正确与否，如果表达式不正确则无法进行运算；表达式正确的情况下执行运算功能。

(4) 运算功能：可以实现大整数的加法、减法、乘法和除法运算。

2. 扩展要求

(1) 实现一些常用的数学函数对大整数进行运算，如：大整数的平方根运算 $\text{sqrt}(x)$ 、大整数的幂运算 $\text{pow}(x,y)$ 等。

【其他要求】

(1) 变量、函数命名符合规范。

(2) 注释详细：每个变量都要求有注释说明用途；函数有注释说明功能，对参数、返回值也要以注释的形式说明用途；关键的语句段要求有注释解释。

(3) 程序的层次清晰，可读性强。

(4) 界面美观，交互方便。

题目 B4：基于关键词的文本排序检索系统

（一）课题内容

对于给定文本库（数据集：“题目 B4 的数据集.zip”解压后即可获得数据集，该数据集为纽约时报的报刊文章），用户提交检索关键词（例如：NBA, basket, ball），在文本库中查询与检索关键词最相关的 k 个文本（例如 k=5），并根据文本与检索关键词的相关度，对这 k 个文本进行排序，将排序后的结果返回给用户。课程教学辅助平台

（二）课题要求

1. 基本要求

- （1）利用 TF-IDF 模型，为文本库中的文本创建索引（如倒排索引）；
- （2）用户输入的关键词可以是一个或多个；
- （3）对于返回的结果文本，需同时显示各检索关键词在结果文本中的出现频度信息；
- （4）系统内支持返回结果文本的查看；

2. 扩展要求

- （1）支持文本库的动态装载和处理；
- （2）支持停用词的管理和维护，停用词是指在没有检索价值的单词，如 is, am, are, a, an, the 等；
- （3）支持大规模文本库的高效排序检索（50 万级别及以上）。

【其他要求】

- （1）变量、函数命名符合规范。
- （2）注释详细：每个变量都要求有注释说明用途；函数有注释说明功能，对参数、返回值也要以注释的形式说明用途；关键的语句段要求有注释解释。
- （3）程序的层次清晰，可读性强。
- （4）界面美观，交互方便。

题目 B5： 基于角色管理的简易家谱管理系统

（一）课题内容

家谱管理系统是查询家谱信息必不可少的一部分，利用家谱管理系统可以清楚的了解家族成员信息。该家谱管理系统是通过树来实现的。要求系统界面友好，易于操作。

（二）课题要求

（1）需要设置普通用户、超级管理员不同角色，不同角色登录后的权限各不相同，普通用户可以进行查询；超级管理员有对所有成员增加、删除和修改的权限。

（2）家谱中成员的信息中均应包含姓名、出生日期、婚否、地址、健在否、死亡日期（若其已死亡）等，也可附加其它信息，并存储于文本文件中。

（3）查询功能。可按照姓名查询，输出成员信息（包括其本人、父亲、孩子的信息、所在辈分）；按照出生日期查询成员名单。

（4）按出生日期对家谱中所有人排序。

（5）打开家谱时，提示当天生日的健在成员。

（6）以图形方式显示家谱。

【其他要求】

（1）变量、函数命名符合规范。

（2）注释详细：每个变量都要求有注释说明用途；函数有注释说明功能，对参数、返回值也要以注释的形式说明用途；关键的语句段要求有注释解释。

（3）程序的层次清晰，可读性强。

（4）界面美观，交互方便。

三、考核及报告要求

1、**考核标准：**根据学生平时的上机出勤情况、为上机所作的准备情况、上机时的表现、程序的完成情况、报告的完成情况、验收答辩时的表现等综合打分。等级为：优秀、良好、中等、及格、不及格。

2、**报告要求：**请参考《程序设计报告内容与格式建议版本》

3、**验收答辩考核内容：**

- (1) 功能演示（界面友好性和功能完备性）；
- (2) 程序结构描述（如：选用的数据结构、功能模块结构关系、类与类之间的关系、函数调用关系、输入输出文件属性描述等）；
- (3) 测试数据的准备情况；
- (4) 编程过程中遇到的典型问题及解决方案。

四、参考资料

- [1] 王海艳等. 数据结构 C 语言[M]. 第 1 版. 北京：人民邮电出版社，2017.
- [2] 陈慧南. 算法设计与分析：C++语言描述[M]. 北京：电子工业出版社，2006.
- [3] 刘汝佳. 算法竞赛入门经典[M]. 北京：清华大学出版社，2009.
- [4] Thomas H. Cormen 等著，潘金贵/顾铁成等译. 算法导论[M]. 第 2 版. 北京：机械工业出版社，2006.
- [5] 郭嵩山等. 国际大学生程序设计竞赛例题解 [M]. 北京：电子工业出版社，2007.