**数据结构课程设计任务书**

**适用专业： 计算机科学与技术、物联网**

**适用年级： 2016级**

**华中科技大学计算机科学与技术学院**

**数据结构课程组**

**2017年12月**

**数据结构课程设计任务书**

数据结构是计算机科学技术、信息安全与物联网等专业的一门重要专业基础课，牢固掌握数据结构的基础知识，熟练地运用数据结构的思想与技术方法解决实际应用问题是是本课程学习的基本任务与目标。而课程设计是实现这一学习目标的重要环节和组成部分。通过课程设计的训练，使学生加深对数据结构知识的理解，牢固掌握其应用方法，并合理灵活地解决一定实际问题，增强和提高综合分析问题与解决问题的能力。

数据结构课程所介绍的主要数据结构包括线性表、栈与队列、串、数组与广义表、树与二叉树、图，以及静态查找表、动态查找表与HASH表等，数据结构课程设计一般要求利用其中两种以上主要结构来解决具体应用问题；另外，能够在应用中结合几种主要结构而建立新的面向应用的数据结构类型是被鼓励的。

本学期课程设计提供三个实际问题或综合实验专题，同学们可根据个人的兴趣偏好，选择其中一个作为自己数据结构课程设计的题目。

**题目一 基于图的系统日志事件关联分析**

* **设计目的：**掌握图等结构的存储结构与基本算法，通过解决较复杂的基于图模型的实际问题，提高学生对数据结构知识综合运用的技能与实践能力。
* **设计内容：**设计有效的逻辑数据结构与存储结构表示WINDOWS或LINUX或IOS系统日志事件信息及其相关关系，例如事件的级别、日期时间、来源、事件id、任务类别、用户及其它事件属性等；分析与表示事件的各种相关关系，如时间相关，数据相关，因果关系等，并用图结构表示它们；进一步在相关图上进行分析与推理，预测某些可能发生的事件，追溯已发生故障或事件的根源。
* **主要数据对象：**系统日志中的事件集；分析所产生的关联关系集或者规则集。
* **主要数据关系：**

**（1）**根据事件的发生时间先后，事件之间可能是线性关系；从单纯事件查询的角度，事件集可能构成普通查找表。

**（2）**另外，事件之间可能还有复杂的关系，如两个事件的主体可能是同一任务/进程；一个或多个事件可能直接导致另一事件的发生；多个事件中的不同主体可能对同一系统对象（文件，或设备等）进行了操作等。

* **主要运算与功能要求：**

**（1）**交互式操作界面(并非一定指图形式界面)；

**（2）**事件插入、删除、查找与检索等：支持按时间查找、事件主体与客体等查找；

**（3）**基于日志事件集进行事件关联图的建立与存储 (如主体相关、客体相关、相伴事件、因果关系等)并提供查询功能；

**（4）**应用关联关系或规则进行系统行为推理、预测；有两种解决途径，依据意愿自行选择：

（4.1）简单方案：根据新发生的事件，查找过往相似事件，并在（3）中的原始相关图上查询相似事件的邻接点事件或以相似事件为起点的路径，作为简单预测的依据。追踪索源则可以查询逆向路径。

（4.2）高级方案：基于原始事件关联图，进行统计分析获取频繁子图模式或关联规则，并表示为关联模式或规则图；并对模式或规则图进行管理：插入、删除、修改、查询、文件存储等。基于模式或规则图进行预测与追踪索源。

**（5）**事件文件的存储管理，读与写。

* **设计提示：**从WINDOWS/LINUX/IOS系统日志或事件查看器获取系统事件原始数据，抽取事件有关数据与信息并以文件形式存储，用图模型建模事件相关关系并以文件保存这种原始关系。预测与溯源可在图模型上进行简单的路径查找与统计分析而给出直接结果。如果有兴趣，可采用概率图模型及其概率推理方法或者关联规则分析方法等高级技术进行分析，但这类方法涉及一些需要自己学习的背景知识，不作强制设计要求，然而，图模型的应用则是强制性的。不允许使用数据库系统管理数据。在界面设计与其他功能上可自由发挥。
* **设计程序的评分标准：**

|  |  |  |
| --- | --- | --- |
| **评分项目** | **分值** | **评价细则** |
| **功能** | 70 | 实现（1）～（3），（4.2），（5）: 0.95～1.0  实现（1）～（3），（4.1），（5）：0.8～0.94  实现（1）～（3），（5）：0.65～0.8  功能（1）～（3），（5）中不全：0～0.65 |
| **程序规范性** | 10 | 命名规范，程序结构好，注释恰当，在程序中含所学基本ADT时调用了对应ADT实现接口：0.9～1.0  命名较规范，程序结构较好，有一定注释：0.75～0.9  规范性一般或欠缺：0～0.75 |
| **特色与创意**  **（功能与界面）** | 10 | 有创意：0.9～1.0  有一定特色：0.75～0.9  无特色：0～0.75 |
| **数据规模** | 10 | 事件记录条数>100：0.9～1.0  50≤事件记录条数<100：0.8～0.9  20≤事件记录条数<50：0.65～0.8  事件记录条数<20：0～0.65 |

* **参考文献：**

[1] 严蔚敏, 吴伟民. 数据结构（C语言版）. 北京: 清华大学出版社,1997

[2] 王晓东. 计算机算法设计与分析. 北京: 电子工业出版社, 2007

[3] 严蔚敏, 吴伟民, 米宁. 数据结构题集（C语言版）. 北京: 清华大学出版社,1999

[4] 傅成兵. 系统日志技术研究与编程实现.计算机与信息技术，2009,( Z1):94~97

**题目二 排课评估与分析助手**

* **设计目的：**掌握图、查找表等的存储结构与基本算法，通过解决较复杂的基于图等模型的实际问题，提高学生对数据结构知识综合运用的技能与实践能力。并且可应用到对排课的合理性与科学性评价，提高教学管理水平，更好地服务教师与学生。
* **设计内容：**设计有效的逻辑数据结构与存储结构表示课表信息，教师信息，课程信息，班级信息，学生信息，教室信息等，由时间、教室、课程名、班级、教师名、上课周期及课堂ID等构成课堂信息。根据课堂及其复杂相关性，依据合理的逻辑与准则评价排课课表的教师安排合理性、课程安排合理性、班级与学生课程学习合理性、教室安排合理性与利用率、能效情况等。本设计只对现有排课进行评估，不去求解最优排课方案。
* **主要数据对象：**教师；教室；课程；班级；学生；课堂等。
* **主要数据关系：**

**（1）**教师之间作为普通同事可以是一般查找表集合中元素间的关系；也可以根据上下级、部门、课程组等因素构成层次关系。教室、班级、学生可以视为普通查找表。课程之间由于先修与后继课之间具有较复杂的关系。

**（2）**课堂之间由于教师、教室、班级、上课时间、课程等因素的作用具有十分复杂的关系。

* **主要运算与功能要求：**

**（1）**交互式操作界面(并非一定指图形式界面)；

**（2）**教师、教室、班级、课程等的增加、删除、修改、查找与检索等；

**（3）**课堂的增加、删除、修改、查找与检索，包括按教师检索其任课课堂，按班级检索课堂，按教室、课程、时间等检索课堂；

**（4）**空闲教室查找与检索、教室利用率分析、能效分析；

**（5）**教师承担课堂合理性分析；//对于教师来说

**（6）**班级参与课堂的合理性分析；//对于学生来说

**（7）**主要数据对象的数据文件组织与存储。

* **设计提示：**从学校排课课表的EXCEL文件获取原始课堂排课数据，抽取其中的课堂信息，从其它资源获取教师、教室、班级、课程等数据。课堂数据可手动录入，鼓励利用相关编程接口自动抽取。使用线性表、查找表、图等结构建模所处理的数据，如有必要，鼓励使用排序算法优化操作。测试数据至少涵盖本院一学期全部课堂。合理性的逻辑与规则可访谈教师、后勤、管理、同学等相关人员。不允许使用数据库系统管理数据。在界面设计与其他功能上可自由发挥。
* **设计程序的评分标准：**

|  |  |  |
| --- | --- | --- |
| **评分项目** | **分值** | **评价细则** |
| **功能** | 70 | 实现全部功能，无一定BUG: 0.9～1.0  实现功能（1）～（5），（7）：0.8～0.9  实现功能（1）～（4），（7）：0.65～0.8  功能（1）～（3），（7）不全：0～0.65 |
| **程序规范性** | 10 | 命名规范，程序结构好，注释恰当，在程序中含所学基本ADT时调用了对应ADT实现接口：0.9～1.0  命名较规范，程序结构较好，有一定注释：0.75～0.9  规范性一般或欠缺：0～0.75 |
| **特色与创意**  **（功能与界面）** | 10 | 有创意：0.9～1.0  有一定特色：0.75～0.9  无特色：0～0.75 |
| **数据规模** | 10 | 课堂记录条数>100：0.9～1.0  50≤课堂记录条数<100：0.8～0.9  20≤课堂记录条数<50：0.65～0.8  课堂记录条数<20：0～0.65 |

* **参考文献：**

[1] 严蔚敏, 吴伟民. 数据结构（C语言版）. 北京: 清华大学出版社,1997

[2] 王晓东. 计算机算法设计与分析. 北京: 电子工业出版社, 2007

[3] 严蔚敏, 吴伟民, 米宁. 数据结构题集（C语言版）. 北京: 清华大学出版社,1999

[4] 王秀焕. 基于图论的高校排课系统优化研究.重庆大学硕士学位论文，2012

[5] 洪文，朱广斌. 排课问题及其数学模型. 安徽电力职工大学学报，2002，7（3）：74-77

[6] 陶涛，谢卫星. 课表模型及排课算法应用. 计算机系统应用，2011，20（2）：198-201

[7] 徐晓.基于本体映射和规则推理的排课模型研究. 软件导刊，2009，8（2）：1-4**题目三 基于AVL树表示的集合ADT实现与应用**

* **设计目的**

平衡二叉树(AVL)作为一种重要的查找表结构，能有效地支持数据的并行处理。本设计使学生牢固掌握AVL树及其实现方法，并应用该结构实现集合抽象数据类型，提升学生对数据结构与数据抽象的认识，提高学生的综合实践与应用能力。

* **设计内容**

本设计分为三个层次：（1）以二叉链表为存储结构，设计与实现AVL树-动态查找表及其6种基本运算；（2）以AVL树表示集合，实现集合抽象数据类型及其10种基本运算；（3）以集合表示个人微博或社交网络中好友集、粉丝集、关注人集，实现共同关注、共同喜好、二度好友等查询功能。

* **主要数据对象：**好友集、粉丝集、关注人集等。
* **主要数据关系：**

**（1）**抽象层面AVL可以表示数据元素之间层次关系或一对多关系。

**（2）**实际应用层面，所讨论的人物关系为集合内元素间的关系。立足于集合建立数据的逻辑模型。

* **主要运算与功能要求：**

**（1）**交互式操作界面(并非一定指图形式界面)；

**（2）**AVL树的6种基本运算：InitAVL、DestroyAVL、SearchAVL、InsertAVL、DeleteAVL、TraverseAVL；

**（3）**基于AVL表示及调用其6种基本运算实现集合ADT的基本运算：初始化set\_init，销毁set\_destroy，插入set\_insert，删除set\_remove，交set\_intersection，并set\_union，差set\_diffrence，成员个数set\_size，判断元素是否为集合成员的查找set\_member，判断是否为子集set\_subset，判断集合是否相等set\_equal；

**（4）**基于集合ADT实现应用层功能：好友集、粉丝集、关注人集等的初始化与对成员的增删改查，实现共同关注、共同喜好、二度好友等查询；

**（5）**主要数据对象的数据文件组织与存储。

* **设计提示**

（1）参考有关文献，实现AVL树的删除操作，维护其动态平衡，这可能是设计中较为复杂的算法；要求提供关键算法的时间与空间复杂度分析。

（2）要求从互联网上获取测试数据集或随机生成测试数据集，数据集的大小具有一定规模；数据与结果以文件保存。

* **设计程序的评分标准：**

|  |  |  |
| --- | --- | --- |
| **评分项目** | **分值** | **评价细则** |
| **功能** | 70 | 实现全部功能，无一定BUG: 0.9～1.0  只DeleteAVL功能未实现： 0.8～0.9  实现功能（1）（4-5），不用AVL实现（3）：  0.65～0.8  不实现AVL与集合ADT,以其它方式实现（1），（3），（5）：0～0.65 |
| **程序规范性** | 10 | 命名规范，程序结构好，注释恰当，在程序中含所学基本ADT时调用了对应ADT实现接口：0.9～1.0  命名较规范，程序结构较好，有一定注释：0.75～0.9  规范性一般或欠缺：0～0.75 |
| **特色与创意**  **（功能与界面）** | 10 | 有创意：0.9～1.0  有一定特色：0.75～0.9  无特色：0～0.75 |
| **数据规模** | 10 | 好友记录条数>100：0.9～1.0  50≤好友记录条数<100：0.8～0.9  20≤好友记录条数<50：0.65～0.8  好友记录条数<20：0～0.65 |

* **参考文献**

[1] 严蔚敏, 吴伟民. 数据结构（C语言版）. 北京: 清华大学出版社,1997

[2] 严蔚敏, 吴伟民, 米宁. 数据结构题集（C语言版）. 北京: 清华大学出版社,1999

[3] Lin Chen. O(1) space complexity deletion for AVL trees, Information Processing Letters, 1986, 22(3)：147-149

[4] S.H. Zweben, M. A. McDonald.**An optimal method for deletion in one-sided height-balanced trees,** Communications of the ACM, 1978, 21(6): 441-445

[5] Guy Blelloch. Principles of Parallel Algorithms and Programming, CMU, 2014

**课程设计总体要求**

1. 三个题目任选其一，计算机科学与技术专业，假期独立完成，下学期检查程序。(注意：所用的数据结构与算法应与课程紧密相关；抄袭他人程序者记零分。)
2. 各题目的难度系数

（1）基于图的系统日志事件关联分析 1.05

（2）排课评估与分析助手 1.00

（3）基于AVL树表示的集合ADT实现与应用 0.96

3. 按华中科技大学计算机科学与技术学院制定的课程设计报告文本结构及写作规范撰写课程设计报告，报告主要内容应至少涵盖如下方面(以下非报告目录)。

一、问题描述与技术现状分析

二、系统总体设计(含模块结构图)

三、数据结构和算法详细设计

四、C/C++语言程序实现的简要说明

(开发环境、支持包、函数原型与功能及调用关系；全部源程序以电子版提供，报告中只能作为附录内容之一)

五、程序测试及结果分析

六、复杂度分析

七、软件的用户使用说明

八、特色与不足

九、主要参考文献

4. 报告提交方式

（1）打印《课程设计报告》(A4纸双面打印)；

（2）以班为单位提交光盘(至少含课程设计报告、源程序和脱离开发环境独立运行的目标程序)；

（3）计算机科学与技术专业按课表上课程设计结束2-3周后提交。

**附：**

**数据结构课程设计指导用书**

——数据结构题集（C语言版），严蔚敏等编著，清华大学出版社

**数据结构课程设计参考书**

——数据结构（C语言版），严蔚敏、吴伟民编著，清华大学出版社