

高等学校教材·计算机科学与技术

数据结构——Java 语言描述

朱战立 编著

清华大学出版社

北京

内 容 简 介

本书系统地介绍了各种类型的数据结构和查找、排序的各种方法。对于每一种类型的数据结构,都详细阐述了基本概念、各种不同的存储结构和不同存储结构上一些主要操作的实现算法,并给出了许多设计实例帮助读者理解。另外,书中还介绍了递归算法的设计方法。面向对象方法是目前软件设计的主流方法,本书用面向对象思想组织全部材料。全书采用 Java 语言作为算法描述语言。

本书既可作为大专院校计算机等专业的教科书,也可作为从事计算机应用的工程技术人员的自学参考书。

版权所有,翻印必究。举报电话:010-62782989 13501256678 13801310933

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

本书防伪标签采用特殊防伪技术,用户可通过在图案表面涂抹清水,图案消失,水干后图案复现;或将表面膜揭下,放在白纸上用彩笔涂抹,图案在白纸上再现的方法识别真伪。

图书在版编目(CIP)数据

数据结构——Java 语言描述/朱战立编著.—北京:清华大学出版社,2005.11

(高等学校教材·计算机科学与技术)

ISBN 7-302-11718-7

I. 数… II. 朱… III. ①数据结构—高等学校—教材 ②Java 语言—程序设计—高等学校—教材
IV. ①TP311.12 ②TP312

中国版本图书馆 CIP 数据核字(2005)第 099546 号

出 版 者:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客户服务:010-62776969

责任编辑:郑寅堃

印 刷 者:

装 订 者:

发 行 者:新华书店总店北京发行所

开 本:185×260 印 张:20.75 字 数:512 千字

版 次:2005 年 11 月第 1 版 2005 年 11 月第 1 次印刷

书 号:ISBN 7-302-11718-7/TP·7645

印 数:1~ 000

定 价: .00 元

编审委员会成员

(按地区排序)

清华大学	周立柱	教授
	覃 征	教授
	王建民	教授
	刘 强	副教授
北京大学	冯建华	副教授
	杨冬青	教授
	陈 钟	教授
	陈立军	副教授
北京航空航天大学	马殿富	教授
	吴超英	副教授
	姚淑珍	教授
中国人民大学	王 珊	教授
	孟小峰	教授
北京师范大学 北京交通大学 北京信息工程学院 北京科技大学 石油大学 天津大学 复旦大学	陈 红	教授
	周明全	教授
	阮秋琦	教授
	孟庆昌	教授
	杨炳儒	教授
	陈 明	教授
	艾德才	教授
	吴立德	教授
	吴百锋	教授
	杨卫东	副教授
华东理工大学	邵志清	教授
华东师范大学	杨宗源	教授
	应吉康	教授
	乐嘉锦	教授
东华大学	蒋川群	教授
上海第二工业大学	吴朝晖	教授
	李善平	教授
浙江大学	骆 斌	教授
南京大学	秦小麟	教授
南京航空航天大学	张功萱	教授
南京理工大学		

南京邮电学院	朱秀昌	教授
苏州大学	龚声蓉	教授
江苏大学	宋余庆	教授
武汉大学	何炎祥	教授
华中科技大学	刘乐善	教授
中南财经政法大学	刘腾红	教授
华中师范大学	王林平	副教授
	魏开平	教授
武汉理工大学	李中年	教授
国防科技大学	赵克佳	教授
	肖 侬	副教授
中南大学	陈松乔	教授
湖南大学	林亚平	教授
	邹北骥	教授
西安交通大学	沈钧毅	教授
	齐 勇	教授
长安大学	巨永峰	教授
西安石油学院	方 明	教授
西安邮电学院	陈莉君	副教授
哈尔滨工业大学	郭茂祖	教授
吉林大学	徐一平	教授
	毕 强	教授
长春工程学院	沙胜贤	教授
山东大学	孟祥旭	教授
	郝兴伟	教授
山东科技大学	郑永果	教授
中山大学	潘小轰	教授
厦门大学	冯少荣	教授
福州大学	林世平	副教授
云南大学	刘惟一	教授
重庆邮电学院	王国胤	教授
西南交通大学	杨 燕	副教授

改革开放以来,特别是党的十五大以来,我国教育事业取得了举世瞩目的辉煌成就,高等教育实现了历史性的跨越,已由精英教育阶段进入国际公认的大众化教育阶段。在质量不断提高的基础上,高等教育规模取得如此快速的发展,创造了世界教育发展史上的奇迹。当前,教育工作既面临着千载难逢的良好机遇,同时也面临着前所未有的严峻挑战。社会不断增长的高等教育需求同教育供给特别是优质教育供给不足的矛盾,是现阶段教育发展面临的基本矛盾。

教育部一直十分重视高等教育质量工作。2001年8月,教育部下发了《关于加强高等学校本科教学工作,提高教学质量的若干意见》,提出了十二条加强本科教学工作提高教学质量的措施和意见。2003年6月和2004年2月,教育部分别下发了《关于启动高等学校教学质量与教学改革工程精品课程建设工作的通知》和《教育部实施精品课程建设提高高校教学质量和人才培养质量》文件,指出“高等学校教学质量和教学改革工程”是教育部正在制定的《2003—2007年教育振兴行动计划》的重要组成部分,精品课程建设是“质量工程”的重要内容之一。教育部计划用五年时间(2003—2007年)建设1500门国家级精品课程,利用现代化的教育信息技术手段将精品课程的相关内容上网并免费开放,以实现优质教学资源共享,提高高等学校教学质量和人才培养质量。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上;精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展、顺应并符合新世纪教学发展的规律、代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻

性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。首批推出的特色精品教材包括:

(1) 高等学校教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 高等学校教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 高等学校教材·电子信息——高等学校电子信息相关专业的教材。

(4) 高等学校教材·软件工程——高等学校软件工程相关专业的教材。

(5) 高等学校教材·信息管理与信息系统。

清华大学出版社经过近 20 年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材经过 20 多年的精雕细刻,形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会
E-mail: dingl@tup.tsinghua.edu.cn

数据结构是计算机专业以及其他一些和计算机技术关系密切的专业的一门核心课程。数据结构课程主要讨论现实世界中数据的各种逻辑结构、在计算机中的存储结构以及各种算法的设计问题。数据结构课程的教学目的,是使学生掌握如何组织数据、如何存储数据和如何处理数据的基本概念和软件设计的基本方法,从而为进一步学习后续专业课程打下坚实的基础。

计算机软件开发方法是不断发展的,数据结构的课程内容也应随着软件开发方法的不断发展而发展。目前面向对象的软件分析和设计技术已发展成为软件开发的主流方法。因此,用面向对象的思想组织数据结构课程的材料,用面向对象的程序设计语言描述数据结构问题,就成为数据结构课程内容改革的必然。

用面向对象的程序设计语言描述算法时,首先涉及选用什么样的高级程序设计语言问题。Java 语言是一种以对象为核心的纯粹面向对象的高级程序设计语言。同 C++ 语言相比,Java 语言取消了指针,也不允许多重继承,但用对象引用的方法不仅可以实现指针的功能,而且更具安全性;在单重继承基础上的接口方法,不仅可以实现多重继承,而且效率更高。另外,Java 语言具有类库丰富、网络编程方便快捷、兼容多种应用环境等特点,是其他高级程序设计语言无法相比的。目前,国外大学的数据结构课程大部分都采用 Java 语言作为算法描述语言,国内大学虽然在这门课程的教学内容改革和教材建设方面起步较晚,但预计改革速度将会很快。因此,本书采用 Java 语言作为描述语言。

数据结构课程是一门理论和实践结合密切的课程。本书采用理论叙述简洁明了、实践应用例子丰富完整的方法编写,希望能达到理论和实践密切结合的教学目的。

本书讨论的内容主要包括:线性表、堆栈、队列、串、数组、向量、集合、矩阵、树、二叉树、图、排序、查找、哈希表以及递归。其中,线性表、堆栈、队列、串、数组、向量、集合和矩阵属于线性结构,树和二叉树属于树型结构,图属于图型结构,排序和查找是两个应用广泛的算法设计问题,哈希表是一种特殊类型的存储结构。树、二叉树和图的算法设计问题经常需要使用递归算法,考虑到有些读者对递归的概念和设计方法不太了解,因此专设一章讨论了递归。

本书给出了十多个类的设计代码,并详细讨论了这些类的设计方法。另外,本书还给出了许多简单应用系统的完整程序设计实例。本书的所有程序代码都在 JCreator

3.50 LE 中调试通过。

为了帮助教师使用本书作为教材,以及帮助读者学习理解本书讨论的内容,本书出版后,所有类的设计代码和所有应用实例的设计代码都可在清华大学出版社网站(<http://www.tup.com.cn>)相关图书的资料下载区中得到。

习题的选择和设计是教材编写的一个重要方面。作者在习题选择和设计上考虑了习题的完整性、系统性和典型性,并把所有习题按类型分成基本概念题、复杂概念题、算法设计题和上机实习题四大类。基本概念题是每章基本内容的要点,读者学习完各章后,应以这部分习题为提纲,透彻理解各章的基本内容以及相关的基本概念;复杂概念题和算法设计题既是教材内容的延伸,也是对读者学习理解程度的检查,读者要花时间认真完成这部分习题;数据结构课程要安排一定课时的上机实习,上机实习题是专为此目的设计的。习题中前面标有*号的,表示难度较大的习题。考虑到学生学习和教师教学的方便,作者在附录 B 给出了上机实习报告书写规范和一个上机实习报告书写实例,在附录 C 给出了部分典型习题的解答。

本书作者近年来一直从事数据结构课程的教学工作,讲授的数据结构课程被学校评为一类课程,编写的 C++ 语言的数据结构教材被教育部定为普通高等教育“十五”国家级规划教材。本书是作者多年教学实践的结晶。

根据作者的经验,使用本教材授课约需 54~70 课时,其中包括 10 个左右课时的课内上机实习。标有*号的内容在课时较少时可不讲授。

本书的最后成书得到了许多人的帮助。韩家新老师和作者多次讨论了书中的一些重要概念问题,研究生杨锦锋、李红雷、杨明、杨萌、秦金祥协助作者做了许多程序设计和调试以及电子图稿的制作工作。在本书完成之际,在此表示诚挚的感谢。另外,感谢在数据结构课程教学和教材编写方面富有经验的叶核亚老师不辞辛劳,认真审阅了全书,指出了书中的一些不妥之处,并提出了许多建设性的意见。

尽管作者在写作过程中非常认真和努力,但是错误和不足之处仍在所难免,敬请读者批评指正。

朱战立

2005 年 3 月

第 0 章 Java 语言基础	1
0.1 数据类型	1
0.2 类和对象	4
0.3 继承	8
0.4 多态	13
0.5 异常处理	14
0.6 包	17
第 1 章 绪论	20
1.1 数据结构的基本概念	20
1.2 抽象数据类型	23
1.3 算法和算法的时间复杂度	24
1.4 算法的空间复杂度分析	30
1.5 Java 语言的工具包	31
第 2 章 线性表	34
2.1 线性表	34
2.2 顺序表	36
2.3 单链表	44
2.4 循环单链表	52
2.5 双向链表	52
2.6 仿真链表	54
2.7 面向对象的软件设计方法	54
2.8 设计举例	55
第 3 章 堆栈和队列	63
3.1 堆栈	63

3.2	堆栈的应用	69
3.3	队列	76
3.4	优先级队列	84
第 4 章	串	90
4.1	串的基本概念及其抽象数据类型	90
4.2	串的存储结构	92
4.3	串类	93
4.4	串的模式匹配算法	101
第 5 章	数组、集合和矩阵	112
5.1	数组	112
5.2	向量类	116
5.3	集合	119
5.4	矩阵类	124
5.5	特殊矩阵	128
5.6	稀疏矩阵	131
第 6 章	递归算法	140
6.1	递归的概念	140
6.2	递归算法的执行过程	141
6.3	递归算法的设计方法	144
6.4	递归过程和运行时栈	146
6.5	递归算法的效率分析	149
6.6	递归算法到非递归算法的转换	150
6.7	设计举例	151
第 7 章	树和二叉树	160
7.1	树	160
7.2	二叉树	165

7.3	以结点类为基础的二叉树设计	171
7.4	二叉树类	181
7.5	二叉树的分步遍历	183
7.6	线索二叉树	190
7.7	哈夫曼树	192
7.8	树与二叉树的转换	199
7.9	树的遍历	201
第 8 章	图	204
8.1	概述	204
8.2	图的存储结构	207
8.3	邻接矩阵图类	209
8.4	图的遍历	214
8.5	最小生成树	218
8.6	最短路径	226
第 9 章	排序	234
9.1	排序的基本概念	234
9.2	插入排序	236
9.3	选择排序	240
9.4	交换排序	246
9.5	归并排序	250
9.6	基数排序	252
9.7	各种排序算法的性能比较	255
第 10 章	查找	257
10.1	查找的基本概念	257
10.2	静态查找	258
10.3	动态查找	263

第 11 章 哈希表	281
11.1 哈希表的基本概念	281
11.2 哈希函数构造方法	283
11.3 哈希冲突解决方法	285
11.4 哈希表类设计	286
附录 A Java 语言工具包实现的常用数据结构	292
附录 B 上机实习内容规范和实习报告范例	295
附录 C 部分习题解答	300
参考文献	313

Java语言基础

本书以纯面向对象的 Java 语言作为数据结构的描述语言,因此,掌握 Java 语言程序设计的基本概念和基本方法就是学习本课程的基础。为帮助读者学习,本章概要叙述与数据结构课程教学内容相关的一些 Java 语言基础。

本章主要知识点:

- 基本数据类型、基本数据类型的包装类、数组和字符串的用途及使用方法;
- 类和对象的概念,类和类成员的访问权限,对象引用和对象的概念;
- 继承的概念,子类的设计方法,接口的作用;
- 多态的概念,多态的作用;
- Java 的异常处理机制,两种异常的处理设计方法;
- 包的概念,包的作用。

0.1 数据类型

数据类型是 Java 语言从 C/C++ 语言中吸取的、使用非常方便的一种语言要素。为了和 Java 语言以类为语言基本要素和核心的思想一致,Java 语言又以 Java API (Java application programming interface) 的形式提供了基本数据类型的包装类。另外,在 Java 中,数组和字符串是两种系统定义的类,考虑到数组和字符串使用的频率很高,因此 Java 语言对数组类和字符串类做了特殊处理,允许用户可以像定义基本数据类型变量一样,定义数组变量和字符串变量(实际上是定义和创建数组类对象和字符串类对象)。

0.1.1 基本数据类型

Java 语言中共定义了 8 种基本数据类型,其中 4 种为整型数,2 种为浮点型数,1 种为字符型数,1 种为布尔型数。Java 语言定义的 8 种基本数据类型及相应的关键字如下。

- 整型: byte、short、int、long;
- 浮点型: float、double;
- 逻辑型: boolean;
- 字符型: char。

基本数据类型可以用来定义变量,例如,

```
int i,j,k;
```

就定义了三个 `int` 类型的变量,分别为 `i`、`j` 和 `k`。上述语句意味着,在编译完成后,系统就为 `int` 类型的变量 `i`、`j` 和 `k` 分配了所需的内存空间。

直接定义基本数据类型(如 `int`)的变量,可以简化程序设计,提高程序设计效率。但 Java 是一种纯面向对象的编程语言,程序的基本要素是类,这就需要考虑基本数据类型到类的转换,这个任务由基本数据类型的包装类完成。

0.1.2 基本数据类型的包装类

有些时候,不仅要知道程序中某个基本数据类型变量(如变量 `i`、`j` 和 `k`)的值,而且要知道某个基本数据类型变量的其他信息。另外,有时候需要把基本数据类型变量放入容器里进行某种处理,但容器只能容纳对象,不能容纳变量,因此,需要有基本数据类型变量到类类型对象的转换。这就是 Java API 的 `java.lang` 包提供的基本数据类型包装类的目的和作用。

每一种基本数据类型都对应有一个包装类。例如,`int` 类型的包装类是 `Integer`,`byte` 类型的包装类是 `Byte`。这些类都被声明为具有 `final` 属性,即不能再被其他子类继承。

利用包装类提供的方法,可以用基本数据类型变量的值来创建对象。例如,下面语句分别用 `int` 类型变量的值和字符串"100"创建了一个由对象引用 `intobj_1` 和 `intobj_2` 指示的 `Integer` 类的对象。

```
int i=10;
Integer intobj_1=new Integer(i);
Integer intobj_2=new Integer("100");
```

这样的对象创建后,就可以像使用其他对象一样来使用上述两个对象。如果想得到对象的值,可以调用相应的成员函数。例如:

```
int i1= intobj_1.intValue();
int i2= intobj_2.intValue();
```

上述语句执行后,`int` 类型变量 `i1=10`,`int` 类型变量 `i2=100`。

0.1.3 数组

数组可以用来保存一组数据类型相同的数据元素,或一组类类型相同的对象。数组中的每个数据元素称作一个数组元素。当把一维数组中的每个数组元素定义为一个一维数组时,就构成了 Java 语言的二维数组,以此类推,还可以有三维数组甚至更多维数组。

和变量的使用方法类同,数组也要定义后才能使用。但不同的是,数组在定义后,还要经过内存单元申请才能使用。例如:

```
int[] a;           //定义
```

```
a=new int[10]; //申请内存单元并赋初始值
```

由于 Java 语言实际上是把数组设计成类,所以上述语句实际上是先定义了一个数组引用 *a*,然后创建该对象(即给该对象分配内存空间并赋初始值),并让数组引用 *a* 表示该对象。

上面两句也可合写成:

```
int[] a=new int[10];
```

如果定义的是基本数据类型的数组,则数组元素的使用方法就与基本数据类型变量的使用方法类似。例如,如下语句段就为每个数组元素赋予了一个从 1 开始递增的数值。

```
int n=10;
for(int i=0; i<n; i++){
    a[i]= i+1; //给数组元素赋值
}
```

数组有一个经常使用的成员变量 `length`,其值表示了数组的长度。例如,上面的语句段可修改为:

```
int n=a.length;
for(int i=0; i<n; i++){
    a[i]= i+1;
}
```

除了可以定义基本数据类型的数组外,还可以定义类类型的数组,定义方法类同。

0.1.4 字符串

字符串是若干个字符组成的序列。为了把一个字符串和别的语言成分区分开来,Java 中的字符串用一对双引号括起来,一个字符串中的字符个数称为字符串的长度。如"abc"就是一个长度为 3、其值为 abc 的字符串。

Java API 中已经设计了字符串类,类名为 `String`,存放的包名为 `java.lang`。

由于字符串类是频繁使用的类,所以,系统给出了定义和使用 `String` 类对象的许多简单方法。

(1) 字符串常量。一对双引号括起来的任何字符序列都是一个字符串常量,如""和"sum"都是字符串常量。字符串常量""的长度为 0,字符串常量"sum"的长度为 3。

(2) 字符串对象。不像其他对象的创建,字符串对象的创建方法比较简单。

下面语句定义了两个字符串引用 `str1` 和 `str`:

```
String str1,str2;
```

在定义字符串引用的同时还可以赋值。例如,下面语句就在定义字符串引用 `str` 的同时,给 `str` 赋了初值"abc":

```
String str="abc";
```

还可以定义 `String` 数组,例如,下面语句就定义了一个有三个数组元素的 `String` 数

组 v :

```
String[] v=new String[3];
```

(3) 字符串的连接运算。Java 语言提供了特殊的字符串运算符“+”,表示把两个字符串连接起来。例如:

```
String str="abc"+"def";
```

该语句就把字符串值"abcdef",赋给了字符串引用 str。

字符串的连接运算在程序的标准输出 System.out.print()和 System.out.println()中经常使用。System.out.print()函数和 System.out.println()函数要求的参数是字符串或字符串表达式。所以,下面都是合法的输出语句:

```
String str="abc";
System.out.print(str+"def");           //输出 abcdef

int x=10;
System.out.println("x="+x);           //输出 x=10 并换行
```

02 类和对象

Java 语言是一种面向对象的程序设计语言,类和对象是面向对象程序设计的基本概念。类是相似对象中共同属性和方法的集合体,而对象是类的实例。

02.1 类的定义

类的定义包括类的声明、成员变量的定义、成员函数的定义三部分。其中,成员函数包括一个特殊的成员函数——构造函数。

构造函数的作用是在创建对象的同时给对象赋值。

构造函数有如下设计要求:

- (1) 构造函数名必须与类名完全相同(包括字符的大小写)。
- (2) 不能有返回类型,包括不能有 void 返回类型。
- (3) 允许同时设计多个构造函数(这称为构造函数的重载)。

例 0-1 设计一个 Animal 类,要求 Animal 类的成员变量包括姓名和年龄,要有一个 run()成员函数,其功能只是简单的输出。

解 Animal 类设计如下:

```
public class Animal{
    private String name;           //成员变量姓名
    private int age;               //成员变量年龄

    public Animal(String n,int a){ //构造函数 1
        name=n;
```

```
        age= a;
    }

    public Animal(String n){                //构造函数 2
        name=n;
        age=1;                             //age默认为 1
    }

    public void setName(String n){        //无返回值有形参的成员函数
        name=n;
    }

    public String getName(){              //有返回值无形参的成员函数
        return name;
    }

    public void setAge(int a){
        age= a;
    }

    public int getAge(){
        return age;
    }

    public void run(){
        System.out.println(name+" can run");
    }
}
```

类的设计有许多设计原则,在这里不便一一讨论。下面只是给出较重要的三点。

设计说明:

(1) 认为既然构造函数没有返回类型,因此就用空类型 void 作为构造函数的返回类型,是初学者最容易犯的错误。如果把构造函数 1 设计成如下形式,则编译不能通过:

```
public void Animal(String n, int a){
    name=n;
    age= a;
}
```

(2) 构造函数允许重载多个,但要求重载的构造函数要么参数的个数不同,要么参数的数据类型不同。Animal 类设计了两个构造函数,构造函数 1 有两个参数,构造函数 2 只有一个参数。

(3) 由于成员变量 name 定义为 private 权限,所以按照类的设计原则,一般应该提供一个 setName()成员函数和一个 getName()成员函数,分别用来给成员变量 name 赋值和取成员变量 name 的值。成员变量 age 类同。