



计算机专业硕士研究生入学考试全程领跑

计算机专业硕士研究生入学考试 数据结构分册

SHUJUJIEGOUFENCE



国防科技大学 张银福 主 编
邓劲生 张炜承 等编著

- 全面分析近年试题
- 重点涵盖名校真题
- 展望考研未来趋势

2-3



中国水利水电出版社
www.waterpub.com.cn

计算机专业硕士研究生入学考试全程领跑

计算机专业硕士研究生入学考试

数据结构分册

国防科技大学 张银福 主编

邓劲生 张炜承 等编著

中国水利水电出版社

内 容 提 要

本书是根据各大院校计算机考研专业课“数据结构”试题和教材编写的考试辅导书。全书每章以综述导航、考点研究、重点提炼和试题解析、全真检测四个部分为框架,为考生全程领跑。

全书由国内数十家重点院校自 2000 年以来的考研试卷精选试题分析而成,涵盖概论、线性表、栈和队列、串、数组和广义表、树和二叉树、图、动态存储管理、查找、内部排序、外部排序、文件等内容。书前有专家对考研复习的建议,书中每章都配备有复习进度时间表。书后附有全真检测的部分提示及解答,以及历年试题学校与年份分类索引表。

建议准备参加数据结构专业课考试的人员精心阅读本书,相信能够通过对本书的学习加快复习的行程。对于各类考研培训机构来说,本书也是一本不错的教学辅导书。

图书在版编目(CIP)数据

计算机专业硕士研究生入学考试·数据结构分册 / 张银福主编. —北京:中国水利水电出版社, 2004

(计算机专业硕士研究生入学考试全程领跑)

ISBN 7-5084-2511-1

I. 计… II. 张… III. ①电子计算机—研究生—入学考试—自学参考资料②数据结构—研究生—入学考试—自学参考资料 IV. TP3

中国版本图书馆 CIP 数据核字 (2004) 第 120583 号

书 名	计算机专业硕士研究生入学考试·数据结构分册
作 者	国防科技大学 张银福 主编 邓劲生 张炜承 等编著
出版 发行	中国水利水电出版社 (北京市三里河路 6 号 100044) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 63202266 (总机)、68331835 (营销中心)、82562819 (万水) 全国各地新华书店和相关出版物销售网点
经 售	
排 版	北京万水电子信息有限公司
印 刷	北京市天竺颖华印刷厂
规 格	787mm×1092mm 16 开本 21 印张 467 千字
版 次	2004 年 11 月第 1 版 2004 年 11 月第 1 次印刷
印 数	0001—5000 册
定 价	32.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

院士寄语

科学日新月异，计算机已经成为人们生活和工作中必备的工具与手段。随着信息化和科学化不断提高，社会对于博士、硕士等高素质人才的需求也越来越大，从而牵引着大批有志青年投入到更高层次的深造队伍中来。很高兴看到知识能够得到社会如此广泛的重视，也希冀更多的硕士报考者能够如愿以偿，脱颖而出。

在计算机的教学和科研领域，国防科技大学一直有着深刻的见解和独到之处。我总是希望能够将教学与科研的经验和心得总结出来，供各兄弟院校和广大青年学子共享，但是一直没有合适的机会。得知在副校长张银福同志的组织下，有十余名奋斗在教学科研第一线的教员、博士和硕士全力投入，辛勤工作半年有余终于完成一套考研辅导丛书，我也感到很欣慰。

这套精炼各个重点院校考研试题、凝结着一线同志多年经验的心血之作快要面世了。我有幸应邀写几句，首先是对编者这么长时间的辛勤工作表示肯定，更重要的是希望读者能够借这套丛书之力得以助跑，顺利踏上新的征程。

周兴铭

2004年9月

《计算机专业硕士研究生入学考试》全程领跑

丛书编委会

主 编：张银福

副主编：刘 波

执 行：邓劲生

编 审：徐 剑 张炜承 陈曙晖 白立军

虞万荣 王晓东 赵振宇 李 蕾

邓云佳 张 茜

组 稿：林 栋 唐建伟 周建飞 唐 琳

张 卫 张 兵 易小玲 胡贵军

何卫华 邹明亮 李海军 陈 军

录 排：文 茜 王 璐 刘伟文 陈彩虹

丛书前言

计算机专业是当今最热门也是发展最迅速的学科之一。很多学生为了进一步提高专业水平和应用能力,纷纷报考计算机专业研究生。据统计,近几年报考计算机软件与理论、计算机应用、计算机与通信等专业硕士研究生的考生数量一直位于考研科目分类排行的前列。尽管有越来越多的幸运儿踏上硕士生学习的征程,但也有大批报考者含泪折戟。究竟怎样才能更为有效地进行考研专业课程的复习呢?

专业课复习离不开基础知识的记忆、大量的模拟试题和平时的学习钻研等,但是我们认为最有效最根本的方法就是研究真题,最好的复习资料就是各个院校的历年试题。当今市场上充斥的大量的复习资料、模拟题往往忽视了这样一个问题,那就是无法判断与真题的贴近度。因为一切的复习过程,最终都是为了提高考试时做题的准确率和速度这个目的;若是偏离了这个目的,复习的努力便毫无意义。

但是考研的专业课有一个大的特点,那就是没有全国统一的试卷,而是由各个院校自主命题。经过对大量试卷进行分析发现,尽管院校采用的教材各异,命题方式和风格不同,但是考试的重点知识点基本上都保持一致。于是我们本着做深做透的原则,花费大量人力物力,从比较具有代表性的数十所学校的近几年考研真题(基本上是2000年之后的试题)中进行筛选,取出典型考题汇集而成这套丛书。

本丛书摒弃以往考研辅导书按部就班,亦步亦趋的指导方式,以一种全新的面孔呈现在读者面前。它的显著特点有:

- 在每章开始,由“综述”带出本章内容在各个院校历年考题中的重要性和出题特点,再串讲本章内容。目的在于提纲挈领,初步理解内容的结构。
- 把重点实时分配到各章节中,经过对知识点分布的特点和规律进行分析和研究,提炼出“常考”和“选考”知识点。目的在于使考生有的放矢,少走弯路,节约复习时间。
- 对于每章的学习,精心准备了一份翔实的时间进度安排表。目的在于帮助考生合理高效地利用有限的时间,按时完成考试的复习任务。
- 我们为每道题目准备了原创性的分析,并在适当的时候提出扩展。

在对不同学校的考研报考及试题情况进行详细的分析之后,我们从众多考试科目中首批选取操作系统、数据结构和计算机组成原理三科成书,希望能满足大多数读者的需要。根据考研趋势发展,我们将在后继版本中包含其他科目。

总之，“立足考研，贴近考试”是我们的宗旨，希望以“全程领跑”的方式，助你一臂之力，以顺利进入更高层次的学习领域。

丛书编委会

2004年8月于国防科技大学

前 言

考研的路，是艰辛的。如果没有志同道合者相互鼓励，没有前人为你领跑，这段路将是漫长的。

目前考研的专业课资料十分匮乏。考生选定了报考的院校之后，往往买本教材，几份往年试卷，再好点能借到本讲义，凭借口口相传的“知识点”，就这么两眼一抹黑地开始复习了。

若谈起多年前的考研，我们都已经淡忘；但要谈起近几年的考研，我们却能如数家珍。经过几个月来对 2000 年以来多家院校计算机专业硕士研究生入学考试试卷的分析，从数百份试卷中统计出各个知识点的重要性，精选出合适的试题并进行透彻分析后，我们惊奇地发现，原来在这看似题型变化多端、内容天马行空的试题中，竟然隐藏着几乎同样的规律。有些知识点，几乎从来没有考过；而有些知识点，却以各种形式出现在不同院校的试卷中，有些甚至每年都考。而这些规律的总结，正是目前市面上大部分参考书所普遍欠缺的。

我们所做的工作，就是尽量将重要知识点在试题分析和实战演练中覆盖无遗。通过对各个院校的考题进行全方位的透析总结而提炼出的常考知识点，旨在指导考生有针对性地进行复习备考，引领考生纵观考研命题特征，把握考研命题趋势，帮助大家深入掌握考研命题的特点和规律，使备考事半功倍，极大地提高学习的投入产出比。

但是，光有分类还是远远不够的。多年从教的经验告诉我们：针对不同的知识背景，需要因材施教；而面对不同的学习内容，又要因势循导。从这个意义上来说，如果没有一个很好的学习指导计划以产生尽可能高效的学习曲线，复习起来也很困难。因此，我们将多年来的教学经验融合到本书的编写中，合理安排学习时间并提出复习建议，以期伴随大家共度这段历程。

本书的编写是一项繁杂而谨慎的工作。各个院校一般都不提供考研专业课试题的标准答案，需要首先将题目自己做出来，然后请多个院校的一些同行老师进行审校，同时广泛参考各个院校的教材。在分析一些目前学术界还没有统一认识的问题时，会列出目前各种教科书的多种意见，以便大家根据实际报考情况有的放矢。

尽管我们在相关领域有着丰富的教学和实践经验，但是仍然战战兢兢如履薄冰，惟恐答案有错而误人子弟。由于水平有限，时间也比较紧迫，尽管经过多次校对和反复修改，书中仍然会存在不妥乃至错误之处。恳请广大读者批评指正，以期在后续的版本中得以修正。意见、建议和批评请发电子邮件至 huben@nudt.edu.cn，作者将不胜感谢。

编者

2004 年 8 月

目 录

丛书前言

前言

第0章 试题分析	1
第1章 概论	3
1.1 综述导航	3
1.2 考点研究	3
1.2.1 常考知识点	3
1.2.2 复习进度建议	4
1.3 重点提炼和试题解析	5
1.3.1 数据结构概念	5
1.3.2 存储方式	5
1.3.3 算法及评价	6
1.4 全真检测	8
第2章 线性表	10
2.1 综述导航	10
2.2 考点研究	11
2.2.1 常考知识点	11
2.2.2 复习进度建议	12
2.3 重点提炼和试题解析	13
2.3.1 线性表的类型定义	13
2.3.2 线性表的顺序表示和实现	14
2.3.3 线性表的链式表示和实现	16
2.3.4 线性表的应用	24
2.4 全真检测	27
第3章 栈和队列	33
3.1 综述导航	33
3.2 考点研究	33
3.2.1 常考知识点	33
3.2.2 复习进度建议	35
3.3 重点提炼和试题解析	36
3.3.1 栈	36
3.3.2 栈的应用举例	41
3.3.3 栈与递归的实现	44

3.3.4	队列	45
3.4	全真检测	49
第4章	串	51
4.1	综述导航	51
4.2	考点研究	51
4.2.1	常考知识点	51
4.2.2	复习进度建议	56
4.3	重点提炼和试题解析	56
4.3.1	串的基本概念	56
4.3.2	串的操作	57
4.3.3	KMP 算法	60
4.4	全真检测	63
第5章	数组和广义表	65
5.1	综述导航	65
5.2	考点研究	66
5.2.1	常考知识点	66
5.2.2	复习进度建议	68
5.3	重点提炼和试题解析	69
5.3.1	数组的定义和运算以及数组的顺序存储结构	69
5.3.2	矩阵的压缩存储(特殊矩阵、稀疏矩阵)	75
5.3.3	广义表的定义、存储结构	77
5.4	全真检测	81
第6章	树和二叉树	84
6.1	综述导航	84
6.2	考点研究	84
6.2.1	常考知识点	85
6.2.2	复习进度建议	87
6.3	重点提炼和试题解析	88
6.3.1	树定义和基本术语	88
6.3.2	二叉树的定义和性质	89
6.3.3	二叉树的存储结构	92
6.3.4	遍历二叉树	95
6.3.5	线索二叉树	107
6.3.6	树和森林	109
6.3.7	赫夫曼树及应用	110
6.4	全真检测	113
第7章	图	120

7.1	综述导航	120
7.2	考点研究	120
7.2.1	常考知识点	120
7.2.2	复习进度建议	122
7.3	重点提炼和试题解析	122
7.3.1	图的概念	122
7.3.2	图的存储结构	123
7.3.3	图的遍历	129
7.3.4	图的连通性问题	135
7.3.5	最短路径	144
7.3.6	有向无环图及其应用	154
7.4	全真检测	159
第 8 章	动态存储管理	166
8.1	综述导航	166
8.2	考点研究	167
8.2.1	常考知识点	167
8.2.2	复习进度建议	170
8.3	重点提炼和试题解析	170
8.4	全真检测	176
第 9 章	查找	177
9.1	综述导航	177
9.2	考点研究	177
9.2.1	常考知识点	178
9.2.2	复习进度建议	182
9.3	重点提炼和试题解析	182
9.3.1	静态表查找	182
9.3.2	动态表查找	189
9.3.3	哈希表的查找	200
9.4	全真检测	209
第 10 章	内部排序	213
10.1	综述导航	213
10.2	考点研究	214
10.2.1	常考知识点	214
10.2.2	复习进度建议	217
10.3	重点提炼和试题解析	218
10.3.1	概述	218
10.3.2	插入排序	219

10.3.3	快速排序	224
10.3.4	选择排序	235
10.3.5	归并排序	244
10.3.6	基数排序	247
10.3.7	各种内部排序方法的比较讨论	249
10.4	全真检测	251
第 11 章	外部排序	254
11.1	综述导航	254
11.2	考点研究	254
11.2.1	常考知识点	254
11.2.2	复习进度建议	258
11.3	重点提炼和试题解析	258
11.3.1	外部排序的方法	258
11.3.2	多路平衡归并的实现	258
11.3.3	置换-选择排序	260
11.3.4	最佳归并树	262
11.4	全真检测	264
第 12 章	文件	266
12.1	综述导航	266
12.2	考点研究	266
12.2.1	常考知识点	267
12.2.2	复习进度建议	269
12.3	重点提炼和试题解析	269
12.4	全真检测	275
附录 A	全真检测部分提示及解答	276
附录 B	样卷试题分析	302
附录 C	历年试题学校与年份分类索引	321

第0章 试题分析

数据结构是计算机专业的核心课程之一。数据结构课程不仅为用计算机语言进行程序设计提供方法,同时更是在一个更高的层次上提出问题、分析问题、解决问题,因此绝大多数院校将此列为研究生入学考试的必考科目。

1. 数据结构试题的特点

数据结构的题型一般分为概念题与算法题:概念题要求牢固掌握各种概念、定义;算法题要求熟记各种算法并熟练加以应用,此外还能对算法的复杂性加以定量分析。考试的特点也很明显:

- 概念多,需要记忆的东西较多。
- 算法较多,个性突出,难把握。
- 与离散数学有一定联系,抽象性强,过于形式化,与实际相联系有时较难。
- 对于编程能力有较强要求。
- 历年试题、各校试题多有重复。

本书遵循数据结构课程的教学大纲要求,从内容上分为13章:第0章为试题分析;第1章为概述,讨论数据的基本概念;第2、3章分别介绍线性表、栈以及队列;第4章讨论串;第5章重点介绍数组、稀疏矩阵、广义表等知识点;第6章介绍树型结构与二叉树;第7章介绍图的知识点;第8章介绍动态存储管理的知识点;第9、10、11章分别介绍查找、内部排序、外部排序;第12章介绍文件。

下面将从历年试题的分布情况来讲一下难点重点的分布。

数据结构的基本概念:难点不多,只要掌握基本的概念、原理即可。

线性表:线性表一章要求在深入理解线性表的类型定义的基础上,熟练掌握有关顺序表和链表的一些常规算法,包括计算长度、查找、插入、删除、逆序、集合运算等。重点是链表,包括线性链表、循环链表和双向链表。

栈和队列:这一章的主要内容有栈和队列的定义和应用。常考的地方有对栈的定义的理解,对栈的应用,循环队列的应用和算法描述。本章的重点是对栈的定义的理解和循环队列,这两个部分是历年出题最多的地方。读者应该对本章的算法确切掌握,清楚算法的原理和运行过程。

串:考题总是围绕串的基本操作来出的,选择、填空主要考察串类型的定义;对于串的实现及串的模式匹配算法主要以问答题或算法设计题出现。其中串的模式匹配算法是常考点也是本章的难点,尤其是KMP算法,读者要熟练掌握NEXT函数和改进NEXT函数的定义和计算。

树型结构与二叉树:需要熟练掌握树和二叉树的相关概念。最重要的一部分是遍历二

叉树和线索二叉树，这也是难点。这部分会涉及一些典型的算法，包括递归与非递归的遍历算法，二叉树的三种线索化方法等等。最后一部分是赫夫曼树及其应用，要求在掌握路径长度和带权路径长度概念的基础上，理解最优二叉树的含义，重点是掌握赫夫曼算法。

数组、稀疏矩阵、广义表：重点较集中，一个是数组元素地址的计算；另一个特殊矩阵的一维存储以及稀疏矩阵的压缩存储，重点掌握三元组方法；最后是掌握广义表的概念及链式存储方式。

图：概念多、算法多、难度大、分值高。题的存储结构、图的遍历、图的连通性、最短路径、有向无环图等方面几乎都是重点，该章要花大力气进行复习。

动态存储管理：虽然本章不是考试的重点内容，但概念、原理较多，要求理解、记牢。本章考点多在描述如何动态存储分配的过程，主要的方法有边界标识法和伙伴系统，要熟练掌握这两种方法的特点、分配过程和回收过程。

查找：本章是考试的重点，内容比较多，难度较大，要求在看懂书的前提下，要多做题，多实践。静态表查找一般难度相对较小，考题形式主要以填空和选择为主；本章的重点和难点集中在动态查找表中，包括：二叉排序树、平衡二叉树、B-树、B+树和键树的查找，考题形式以问答和算法设计题为主，尤其是算法设计题往往难度较大；哈希表的查找也是常考点，但往往难度不大，只要理清哈希函数的原理，这一部分考题应该不难解答。

内部排序：内部排序是数据结构的重点内容之一，内容主要包括插入排序、选择排序、快速排序、归并排序、基数排序以及各个算法的性能分析。在历年考题中，各部分都有体现，快速排序和堆排序的考题相对多一点。本章的重点是插入排序、快速排序和堆排序，另外，归并排序和基数排序的出题大多是对教材上算法的应用考查。掌握本章的内容，应该对各个排序算法的原理有深刻的认识。

外部排序：外部排序在数据结构中不属于重点内容，内容主要包括多路平衡归并、置换选择排序、最佳归并树。历年考题大都是应用型的题目，即是对置换选择排序算法和最佳归并树的应用。应该熟练掌握利用置换选择排序产生初始归并树的算法、生成最佳归并树的过程以及利用最佳归并树进行的读写纪录数的计算。

文件：内容不是很多，并不是考试重点，2000年以后考题中很少涉及本章。其中要重点掌握各种文件组织形式，包括：顺序文件、索引文件、ISAM和VSAM文件、散列文件和多关键字文件，掌握每种文件形式的优缺点、适用领域和基本操作等。

2. 复习方法

针对数据结构该门课程的特点，我们提倡按以下步骤进行复习。

- 通读教材，要非常熟悉每一个定义、算法甚至算法的背景。
- 整理出总体框架，画出框架图。
- 联系历年真题，有重点地对常考知识点加以复习强化。
- 反复阅读已经做过、看过的试题，尤其是曾经做错的题目，以避免下次犯同样的错误。

只要踏踏实实看书、总结、复习、做题、看题，相信能较全面掌握数据结构的各个知识点，从而能够从容地应付考场上的各种难题。

第1章 概论

本章知识点学习进度

	重要级别	难度系数	建议复习用时
数据结构的概念	★	☆	1 小时
存储方式	★★	☆	1 小时
算法及评价	★★★	☆☆	2 小时

本章要求掌握一些数据结构的基本概念。内容不多，但历年试题均有涉及。大家要以本章为统领，联系后面的各章，融会贯通。

1.1 综述导航

数据结构的原理与算法较抽象，要从概念、算法、复杂度分析中去领会并在实际应用中一步步掌握。本章的作用在于统领全书，大家要从数据结构的概念、存储方式、算法及评价等方面总体掌握数据结构要研究的内容，为后面复习打下基础。

1.2 考点研究

1.2.1 常考知识点

1. 数据结构的概念

(1) 数据结构：指同一数据元素类中各数据元素之间存在的关系。数据结构分为以下三部分：逻辑结构、存储结构以及数据的运算。

(2) 逻辑结构是对数据之间关系的描述。形式上可以使用一个二元组 $B = (K, R)$ 来表示，其中 K 是数据元素的有穷集合， R 则是由有限的关系所组成的集合。从逻辑上可以把数据结构分为线性结构和非线性结构。

➤ 线性结构：线性结构有且只有一个终端结点与一个开始结点，并且所有结点都最多只有一个前驱与后继，如线性表。

➤ 非线性结构：非线性结构可能有多个终端结点与多个开始结点，每个结点都可能有多多个前驱与后继。

(3) 存储结构：指数据的逻辑结构在计算机存储器中的实现，存储结构是依赖于计算机的。通常有四种基本的存储映像方法，即：顺序方法、链接方法、索引方法和散列方法。

(4) 数据的运算：在数据逻辑结构上定义的操作算法。

2. 存储方式

(1) 线性结构的存储方式

- 顺序方式：把逻辑上相邻的结点存储在物理上相邻的存储单元里，结点之间的关系由存储单元的邻接关系体现。优点是占用最少的存储空间；缺点是插入、删除操作麻烦。
- 链接存储方式：结点所占的存储单元分为两部分，一部分存放结点本身的数据，另一部分存储后继结点的存储地址即指针项。优点是插入、删除操作方便，缺点是所需的存储空间较大。
- 索引方式：用结点的索引号来确定结点的存储地址。优点是速度快；缺点是增加了额外的索引表，插入、删除操作需要修改索引表。
- 散列方式：根据结点的值确定它的地址。优点是查找、删除、插入操作方便，缺点是可能出现结点存储单元冲突。

(2) 树型结构存储

每个结点可能有多个后继结点，一般情况下采用链接的方式进行存储，链接方式正好能表达树型结构中的父子和兄弟两种层次关系。

(3) 图形结构存储方式

每个结点可能有多个前驱结点与后继结点，所以采用链接方式。具体实现请参见第 7 章，此外也可以使用矩阵来存储图。

3. 算法及评价

算法是解决某一特定类型问题的有限运算序列。本书算法采用 C 或 C++ 语言。

算法特性：有穷性、确定性、可行性、输入与输出。

算法评价：

- 正确性：能够确保对于某种相对程度的随机输入都有正确的输出。
- 运行时间：指一个算法在计算机上运行所花费的时间，采用时间复杂度来表示。时间复杂度是指算法中包含简单操作的次数，只需大致计算其量级便可，如 $O(1)$ 、 $O(\log_2 n)$ 、 $O(n \log_2 n)$ 等。
- 占用空间：指在计算机存储器上所占用的存储空间，包括算法运行过程临时占用的存储空间的大小，称之为空间复杂度。
- 简单性：指算法的易读性。

1.2.2 复习进度建议

本章知识点其实不多，概念容易掌握，但要注意时间复杂度、空间复杂度的分析，最为重要的是要与后面几章联系起来。建议该章概念不要死记硬背，理解为主，花较短的时间去复习。

1.3 重点提炼和试题解析

1.3.1 数据结构概念

【例 1-1】(中山大学 2001 年试题)

按照数据的逻辑结构的不同,可以将数据结构分成_____。

- A. 动态结构和静态结构 B. 紧凑结构和非紧凑结构
C. 线性结构和非线性结构 D. 内部结构和外部结构

【解答】

C

【例 1-2】(中山大学 2002 年试题)

以下数据结构中,是非线性数据结构_____。

- A. 树 B. 字符串 C. 数组 D. 栈

【解答】

A

【例 1-3】(武汉大学 2000 年试题)

下面关于数据结构的叙述中,正确的是_____。

- A. 数组是同类型值的集合。
B. 递归算法的程序结构比迭代算法的程序结构更为精炼。
C. 广义表是一种线性结构。
D. 用一维数组存储二叉树,总是以先(前)序顺序遍历各结点。

【解答】

A、B

1.3.2 存储方式

【例 1-4】

线性结构的顺序存储结构是一种_____的存储结构,线性表的链式存储结构是一种_____的存储结构。

- A. 顺序方式 B. 链接存储方式 C. 索引方式 D. 散列方式

【解答】

A B

【例 1-5】(哈尔滨工业大学 2002 年试题)

在计算机的存储器中表示时,物理地址和逻辑地址相同并且是连续的,称之为()。

- A. 逻辑结构 B. 顺序存储结构
C. 链式存储结构 D. 以上都不对

【分析】