

高等学校计算机专业教材精选·计算机基础

计算机科学导论教程

黄思曾 主编 袁 志 黄捷迅 侯长春 编著



计算机科学导论教程

黄思曾 主编 袁 志 黄捷迅 侯长春 编著

清华大学出版社
北京

内 容 简 介

和同类教材相比,本书的特点在于以“一个核心、三条纲领”来组织教学内容。计算机的核心功能是数据处理,计算机科学要提供数据表示和数据加工表示的理论、方法、技术,并最终在计算机系统上实现。本书共分6章,内容包括认识计算机和计算机科学、计算机系统的组成、计算机软件系统、计算机通信与网络、数据表示方法、数据加工表示方法。按照本书的叙述体系,初学者容易理解后续课程中展开的专业概念,并开始关注专业课程之间的内在关联。

本教材适宜在高等学校作为教材使用。也适合有兴趣了解计算机科学的读者阅读。

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。

版权所有,侵权必究。侵权举报电话:010-62782989 13501256678 13801310933

图书在版编目(CIP)数据

计算机科学导论教程/黄思曾主编. —北京:清华大学出版社,2007.8

(高等学校计算机专业教材精选·计算机基础)

ISBN 978-7-302-15234-7

I. 计… II. 黄… III. 计算机科学—教材 IV. TP3

中国版本图书馆CIP数据核字(2007)第071263号

责任编辑:焦虹 顾冰

责任校对:时翠兰

责任印制:王秀菊

出版发行:清华大学出版社

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

c-service@tup.tsinghua.edu.cn

社总机:010-62770175

投稿咨询:010-62772015

地 址:北京清华大学学研大厦A座

邮 编:100084

邮购热线:010-62786544

客户服务:010-62776969

印 刷 者:北京国马印刷厂

装 订 者:三河市兴旺装订有限公司

经 销:全国新华书店

开 本:185×260 印 张:14.75

字 数:355千字

版 次:2007年8月第1版

印 次:2007年8月第1次印刷

印 数:1~4000

定 价:20.00元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。
联系电话:010-62770177 转 3103 产品编号:023569-01

出版说明

我国高等学校计算机教育近年来迅猛发展,应用所学计算机知识解决实际问题,已经成为当代大学生的必备能力。

时代的进步与社会的发展对高等学校计算机教育的质量提出了更高、更新的要求。现在,很多高等学校都在积极探索符合自身特点的教学模式,涌现出一大批非常优秀的精品课程。

为了适应社会的需求,满足计算机教育的发展需要,清华大学出版社在进行了大量调查研究的基础上,组织编写了《高等学校计算机专业教材精选》。本套教材从全国各高校的优秀计算机教材中精挑细选了一批很有代表性且特色鲜明的计算机精品教材,把作者们对各自所授计算机课程的独特理解和先进经验推荐给全国师生。

本系列教材特点如下。

(1) 编写目的明确。本套教材主要面向广大高校的计算机专业学生,使学生通过本套教材,学习计算机科学与技术方面的基本理论和基本知识,接受应用计算机解决实际问题的基本训练。

(2) 注重编写理念。本套教材作者群为各校相应课程的主讲,有一定经验积累,且编写思路清晰,有独特的教学思路和指导思想,其教学经验具有推广价值。本套教材中不乏各类精品课配套教材,并力图努力把不同学校的教学特点反映到每本教材中。

(3) 理论知识与实践相结合。本套教材贯彻从实践中来到实践中去的原则,书中的许多必须掌握的理论都将结合实例来讲,同时注重培养学生分析、解决问题的能力,满足社会用人要求。

(4) 易教易用,合理适当。本套教材编写时注意结合教学实际的课时数,把握教材的篇幅。同时,对一些知识点按教育部教学指导委员会的最新精神进行合理取舍与难易控制。

(5) 注重教材的立体化配套。大多数教材都将配套教师用课件、习题及其解答,学生上机实验指导、教学网站等辅助教学资源,方便教学。

随着本套教材陆续出版,相信能够得到广大读者的认可和支持,为我国计算机教材建设及计算机教学水平的提高,为计算机教育事业的发展作出应有的贡献。

清华大学出版社

2006年11月

前 言

“计算机科学导论”是大多数高等院校计算机科学与技术类专业学生必修的专业课程。课程讲授和计算机系统、计算机科学有关的基本概念,使学生对本专业的核心知识有全面的、概要的认识。

本课程的一种典型教学安排是,每周课堂讲授 3 学时、14 周共 42 学时。在相对短的时间内要完成覆盖面很广、内容极为丰富的专业知识教学,而又要使一年级学生能够理解,关键在于:

1. 要建立一个课程讲授结构体系;
2. 要把握好内容广度和深度的关系;
3. 要采用深入浅出的叙述风格。

笔者在中山大学计算机科学系任教多年,20 世纪 80 年代负责开设本课程之后,一直是本课程主讲教师。二十余年的教学实践中,笔者设计以“一个本质核心,三条内容纲领”来组织本门课程,使学生能够在短短的十几周里对计算机系统和计算机科学能够有入门性的、较全面的认识。

“一个本质核心”是指计算机的“数据处理”功能。计算机是一种数据处理机器,计算机科学是使计算机完成各行各业、形形色色的数据处理任务所需要的理论、方法和技术的知识集合。

“三条内容纲领”是:

(1) 数据的分层表示方法学,其表示层次:

- 现实世界的的数据对象;
- 数据结构和信息结构层;
- 程序设计语言层;
- 机器层;
- 物理层。

(2) 数据加工的分层表示方法学,其表示层次:

- 数据处理问题;
- 解题模型;
- 算法层;
- 程序设计层;
- 机器程序层;
- 物理层。

(3) 计算机系统的构成(数据和数据加工表示方法的最终载体):

- 硬件系统;
- 软件系统;
- 网络系统。

本教材按照上述纲领展开,包含了后续各门计算机类专业课程的主要概念,从而使学生对计算机科学的内容及其内在的关联有全面、清晰、概要的认识。

教学实践表明,本课程的学习能够激起学生新的问题和思考。课程结束后,学生的疑问通常会比课程开始前还要多。这就说明本课程已经为学生后续几年专业课程的学习奠定了良好的基础,真正起到了“导论”的作用。

由于计算机技术的发展十分迅速,作者水平所限,书中难免有错误和不足之处,期望读者在使用过程中给予指正。

黄思曾

E-mail: hsz05@126.com

2007年5月于华软园

目 录

第 1 章 认识计算机和计算机科学	1
1.1 计算机是数据处理机	1
1.1.1 数据.....	1
1.1.2 信息.....	2
1.1.3 数据处理.....	3
1.1.4 数据处理机.....	4
习题	5
1.2 计算机科学	5
1.2.1 理解计算机科学.....	6
1.2.2 算法.....	8
1.2.3 用计算机解决数据处理问题	10
1.2.4 计算机系统的发展历程	13
习题.....	14
1.3 机内信息表示基础——二进制数.....	15
1.3.1 位置记数法	15
1.3.2 P 进制数	16
1.3.3 数制的转换规则	16
习题.....	18
本章小结	18
第 2 章 计算机系统的组成	19
2.1 四大功能部件.....	19
2.1.1 内存	19
2.1.2 中央处理器	24
2.1.3 外部设备	26
2.1.4 总线	37
习题.....	39
2.2 计算机程序.....	40
2.2.1 机器指令	41
2.2.2 程序存储原理	43
2.2.3 冯·诺依曼结构体系	46
习题.....	47
2.3 计算机系统.....	48
2.3.1 硬件和软件	48

2.3.2	计算机系统的几种应用模式	50
	习题	52
	本章小结	52
第3章	计算机软件系统	54
3.1	软件系统	54
3.1.1	软件系统的组成	54
3.1.2	应用软件示例	55
3.1.3	获得软件的途径	58
3.1.4	软件版权保护	59
	习题	60
3.2	操作系统	60
3.2.1	操作系统的目标	60
3.2.2	进程和中断	61
3.2.3	操作系统的资源管理功能	62
3.2.4	作业的管理调度方式	64
3.2.5	I/O设备的输入/输出控制	66
3.2.6	操作系统的常见种类	67
3.2.7	用户界面	68
3.2.8	有代表性的操作系统产品	70
	习题	71
3.3	程序设计语言和语言处理软件	72
3.3.1	程序设计语言	72
3.3.2	程序设计语言的发展	73
3.3.3	程序设计和软件开发	76
3.3.4	程序/语言的处理过程	77
3.3.5	编译程序	78
3.3.6	上机作业的传统过程	79
3.3.7	程序作业工具	80
	习题	80
	本章小结	80
第4章	计算机通信与网络	82
4.1	数据通信概念	82
4.1.1	信息、数据与信号	82
4.1.2	数据的传输	83
4.1.3	传输介质和信道	86
4.1.4	信号传输的技术特性	87
4.1.5	数据交换	90

习题.....	91
4.2 计算机网络.....	91
4.2.1 计算机网络的组成	92
4.2.2 网络类型	94
4.2.3 网络协议和体系结构模型	96
习题.....	99
4.3 计算机局域网.....	99
4.3.1 局域网的拓扑结构	99
4.3.2 介质访问控制方法.....	101
4.3.3 局域网中常见的网络设备.....	102
习题	104
4.4 因特网	105
4.4.1 网络互联.....	105
4.4.2 因特网的地址编制和域名.....	105
4.4.3 因特网的典型应用.....	106
4.4.4 因特网的接入.....	111
习题	113
本章小结.....	113
第5章 数据表示方法.....	115
5.1 数据的分层表示方法学	115
5.1.1 现实世界层.....	115
5.1.2 信息世界层.....	115
5.1.3 高级语言层.....	117
5.1.4 机器层.....	118
5.1.5 物理层.....	118
习题	118
5.2 数的表示	119
5.2.1 无符号整数的表示.....	119
5.2.2 有符号整数的表示.....	119
5.2.3 实数的表示.....	123
5.2.4 复数的表示.....	126
习题	126
5.3 字符和汉字的表示	126
5.3.1 字符编码.....	127
5.3.2 汉字编码.....	129
5.3.3 汉字处理系统.....	133
习题	133
5.4 图像和声音的表示	133

5.4.1	图像的表示	133
5.4.2	声音的表示	135
	习题	135
5.5	高级语言的数据表示手段	136
5.5.1	常量和变量	136
5.5.2	函数和表达式	137
5.5.3	数据类型	142
	习题	148
5.6	动态数据结构	148
5.6.1	数据结构含义	148
5.6.2	数据结构基本概念	149
5.6.3	线性表	153
5.6.4	树和二叉树	158
5.6.5	图	161
	习题	163
5.7	外存数据组织	163
5.7.1	文件	164
5.7.2	数据库	169
	习题	178
	本章小结	179

第6章	数据加工表示方法	180
6.1	数据加工的表达层次	180
	习题	181
6.2	算法表示手段	181
6.2.1	算法的表示	181
6.2.2	算法流程图	182
6.2.3	类程序设计语言的伪代码	188
	习题	191
6.3	结构化设计	191
6.3.1	三种流程结构	192
6.3.2	结构化方法	196
	习题	198
6.4	算法示例	198
6.4.1	顺序结构的算法设计	198
6.4.2	分支结构的算法设计	200
6.4.3	循环结构的算法设计	203
	习题	207
6.5	软件的工程化开发方法	207

6.5.1 关于软件工程学.....	207
6.5.2 软件生命周期概念	208
6.5.3 传统的软件生命周期阶段.....	209
6.5.4 软件开发规范.....	213
6.5.5 软件开发模型.....	215
6.5.6 开发方法和工具.....	215
习题	219
本章小结.....	219
参考文献	220

第 1 章 认识计算机和计算机科学

计算机问世仅仅五十多年,给人类社会带来翻天覆地的变化。如果说,蒸汽机和电的发明带来了工业革命,极大地影响了人类的生活方式。那么,计算机引起了人类社会新一轮革命,我们称之为“信息革命”。以计算机为核心的信息处理系统以空前的广度和深度渗透到社会的每一个领域。今天,生产制造业、商业、银行金融业、交通运输业、通信业、科研、教育、医疗卫生……乃至娱乐休闲都离不开计算机。众多国计民生不可或缺的事务对计算机的依赖,达到了一旦计算机系统发生故障,业务就被迫停顿的地步。

计算机在当代社会里“无所不在、无所不能”的印象使外行人对计算机产生一种神秘感。科幻小说、科幻电影里的机器人和虚拟世界更容易使普通人想入非非。其实计算机的功能是单一的,从本质的角度来看,计算机只能做一件事,就是**数据处理**,也可以称为**信息处理**。计算机只是一种数据处理机器。当然,和历史上出现过的其他数据处理机器相比,计算机的功能要全面得多、强大得多。

所以,对计算机和计算机科学的认识要从什么是数据、什么是数据处理、什么是数据处理机开始。

1.1 计算机是数据处理机

1.1.1 数据

平时遇到“数据”(data)这个字眼的时候,容易使我们联想到“数”,如实验数据、财务数据、经济数据等。计算机科学赋予数据这个术语更本质的含义。应用中,计算机要面对客观世界形形色色的事物,一个人、一台设备、一份合同、一部动画、一门课程……,不管看得见看不见,它们都是客观存在的东西。计算机科学用数据这个概念来表示客观事物。

1. 数据的定义

数据是客观事物属性的记录表示。

计算机科学用一种简单而有效的方法来表示客观事物,就是把事物等同于由事物一组特征所组成的集合。比如用

{学号,姓名,系,专业,年级,年龄,性别,血型,口音,脸形,……}

这么一组学生的特征来表示“学生”这种客观事物。用

{课程代码,课程名,学分,规定学时,先导课程,……}

来表示“课程”这种客观事物。

一般说来,一种事物几乎会有无限个各种各样的特征,表示了事物不同方面的性质。我们从业务处理的需要出发,从事物的特征中选取有限个**数据属性(attribute)**,把这组数据属性的集合称为**数据实体(entity)**。这样,抽象的数据实体概念就可以表示具体的事物了。

即

事物 { 特征 } ⇒ (数据)实体 { (数据)属性 }

这种抽象过程不是简单地替换了术语名称,而是体现了专业人员对客观事物进行分析、分类、拆分、重组、选取等一系列的思维动作,这样才能使用数据的概念来表示客观世界的事物。

2. 数据的名和值

要从“名”(name)和“值”(occurrence)两个不同角度来认识数据概念。

我们总是选择有恰当语义的名称来描述数据实体和数据属性。比如

学生 { 学号,姓名,系,专业,年级 }

可见,数据名刻画的是一类事物的特征构成。

而一组数据值:

(07001,张三,计算机系,软件开发,一年级)

表示特定的一个学生。可见,数据值刻画的是一类事物中的个体。

通常,一个数据名会和一组数据值相对应。学生名字可以是张三、李四、王五……。

同一类数据值的集合叫做数据集(data set)。

3. 数据的形式

既然数据是事物属性的记录表示,就必须有一定的表示形式(forms)。最常见的几种数据形式是:数、文字(字符/汉字串)、图像、图形、视频和声音。

18这个数,表示了某个学生“年龄”的数据值;“张三”、“Mary”这些字符串,表示了“姓名”的数据值;一张照片是幅图像,表示了某人的外貌特征,当然是数据的一种表示形式。图形和声音亦如此。

现代的计算机可以用数、文字、图像、声音等各种数据表示形式来输入、输出、存储、加工各种数据。大家已习惯用“多媒体”这个字眼来形容计算机处理多种形式数据的能力。其实更准确地说,媒体是不同形式数据的驻留介质。

4. 数据的驻留介质

显然,一切形式的数据必须记录在某种介质(media,又称媒体)上。最天然的数据驻留介质是人的大脑皮层。最广泛使用的传统数据介质之一是纸。在计算机里,最常见的数据存储介质是磁盘、光盘、磁带、内存储器、纸质的穿孔带、穿孔卡等。它们采用电磁材料、光电材料、半导体电子电路等来制造。

1.1.2 信息

今天,“信息”(information)已经成为广泛使用的大众词语。但是在计算机科学里,信息是有严格的定义的。

1. 信息的定义

信息是客观事物给出的知识。

和数据一样,信息也和客观世界的事物密切相关。但是,信息不是事物本身的刻画表示,而是指由事物传递出来的知识。信息之所以能够产生,有三个不可或缺的要素。了解这三个要素有助于理解信息概念的含义。

2. 信息产生三要素

信息成立必须依赖三个要素的存在:(信息)源、理解规则、接收者。

先看一个例子：傍晚的天空中出现一片红霞，有人看到了，凭着他掌握的气象知识或者经验，会想到“明天天晴”。

这个例子里，“晚霞”是客观事物，看到的晚霞（形式）就是一种数据，是信息的产生源头；“看见晚霞的人”是接收者；“气象知识”就是一种理解规则，产生的信息就是“明天天晴”。缺少了上述任何一种要素，信息都无法产生。乌云满天，如何想到天晴？晚霞灿烂，就是没有人注意，或者看见晚霞的人是个3岁小孩子，只会觉得云彩像猫、像狗，不会由此产生“明天天晴”这个信息。

就是说，信息的接收者要依据某种理解规则，从数据中提取出所包含的信息。信息就是这样产生的。

3. 信息和数据

现在可以区分数据和信息这两个最基本的概念了。我们说：信息是数据的内涵，数据是信息的外表形式。

从数据中提取信息必须依靠某种理解规则。这些规则可能是非常专业的，更多的规则会是社会和文化所约定俗成的。比如，碰到“张石头”这个字眼，我们一定会认为这是某位中国人的名字，不可能认为它包含了“某块石头姓张”这样一个信息。

就是说，引用不同的理解规则必然会在同一个数据源里得到不同的信息。尽管如此，大多数时候数据里所包含的信息，内容是明确的，提取信息的理解规则是隐含的、约定俗成的。因此，即使在计算机专业的专业范畴里，往往也会不加区分地使用数据和信息这两个术语。尽管它们的定义是有根本区别的。

1.1.3 数据处理

1. 数据处理的含义

数据处理(data processing)是个过程，指的是：把收集到的源数据经过加工得到含有特定信息的结果数据。就是说，数据处理加工的对象是数据，得到的结果也是数据，特定的处理过程所产生的数据含有人们希望使用的信息。因此，数据处理也往往称为信息处理(information processing)。

数据代表万物，对数据的加工也是含义广泛的。粗略地说，可以把各种加工动作分为两大类：数值性加工和非数值性加工。前者可以理解为通常意义的数学运算，后者用来指各种非数学意义的数据变换操作。因此，把两个数加加减减，得到和差是一种数据加工；从一堆数据中查找出特定的一个数据，或者按指定的升降准则把数据排好次序是数据加工；修改、删除数据的值也是一种数据加工。可见，各个领域都存在数不胜数的数据处理要求。

2. 数据处理的基本环节

人类活动和社会事务中充满形形色色的数据处理过程。小学生计算 $2+3=?$ 是个数据处理过程。先把被加数和加数写在纸上，再输入大脑“心算”，最后指挥手拿好笔，把结果5写出来。人过马路也是一个数据处理过程。人在过马路前先站在马路边左右张望，这是用眼睛收集数据。看到一辆车飞驰而来，此信息马上输入大脑，大脑开始紧张的运算加工：先检索出一条存储过的知识，人让车碰到可不得了；然后赶快估计车的速度、方向、相会时间，决定是站住不动，还是快步走过；最后发出神经信号指挥肌肉执行。

分析各种过程，可以发现数据处理都包含下述几个基本环节：收集、录制和输入、加工、

输出、存储、传输。处理过程的数据流如图 1-1 所示。



图 1-1 数据处理的基本环节

1.1.4 数据处理机

任何数据处理过程都必须在某种物理实体上进行,称之为数据处理机(data processor)或数据处理系统。人是一种天然的数据处理机体系。但是,人类从远古时代起就发现自身数据处理能力的缺陷。“容易忘记事情”,意味容易丢失已经存储在大脑里的数据;“计算不够快、不够准”,意味着计算能力的低下。“距离远一点,就看不见、听不清楚”意味数据传送能力的局限。于是人们就致力于发明各种器械来改进和提高数据处理的能力。

长期以来,针对数据处理的单个环节不断做出改善。老祖宗想出“结绳记事”的方法来提高记忆能力,效果当然并不理想。看到绳结,提醒自己有事了,什么事呢?还是没记住。于是改进再继续,发明文字和记数的方法以记录数据,发明纸作为数据的长期存储介质,发明笔用作书写工具以完成数据输出,发明算盘、计算尺、机械式计算机以加快计数速度和提高准确度,发明照相机、录音机以记录图像形式和声音形式的数据……。努力一直持续了几千年,直到 20 世纪 50 年代电子计算机问世,人类才获得了高效率地完成数据处理全过程所有环节的手段。

1. 人作为数据处理机

人是一个数据处理系统,处理任务的各个基本环节由相应的人体器官完成。

(1) 五官负责数据输入。眼睛看到的是影像、文字、数等形式的数据,耳朵听到的是声音形式的数据,口尝到的味道、鼻嗅到的气味、皮肤触摸到的物体表面状况都是事物的各种特征,即数据属性。

(2) 大脑是最重要的数据处理器官。大脑负责存储数据、加工数据,还要负责控制和协调人体系统数据处理的全过程。大脑出了毛病,人还能活,但是就失去了数据处理的基本能力。

(3) 嘴、手、身体都是数据输出器官。说的话、写的字、做出的表情和动作都包含信息,都是数据。

(4) 而数据传输这个环节是由神经系统来完成的。

2. 计算机作为数据处理机

“计算”不是“计数”的同义词,计算机也不只是数学意义的计算设备。计算机(computer)是一种数据处理机,能够以前所未有的性能完成数据处理所有环节的任务。日常生活里专注数学计算的电子设备叫 calculator(计算器),并不等同于计算机。

1) 输入设备

计算机可以配备各种不同的输入设备(input device)来输入不同形式的数据。如使用键盘输入字符或者汉字形式的数据,用扫描仪输入图像形式的数据,用读卡机输入有特别表示形式的数据等。

2) 输出设备

计算机可以配备各种不同的输出设备(output device)来输出不同形式的数据。比如用显示器来输出字符或图像形式的数据,用打印机把数据输出到纸上等。

3) 存储器

任何数据处理系统都必须有数据存储设备。计算机的存储器(memory)是以二进制数的形式来存储各种数据的。通常,计算机有一个内存储器(也可称主存储器)来存放工作时候要用到的数据。还会使用若干种外存储器存放需要长期保存的数据,如磁盘、光盘、磁带、U盘等。

4) 中央处理器

中央处理器(central processing unit,CPU)包括:运算器、寄存器、控制器等主要部件。CPU依靠人给定的指令来完成所有数据处理动作,控制和协调操作执行的顺序。可见,CPU是计算机的核心。

5) 总线

总线(bus)是传送信息的一组信号线,是CPU内部各功能部件之间、CPU和存储器及输入输出设备之间的连接通道。按传送的信息内容,可分为数据总线、地址总线和控制总线等几部分。

如上所述,计算机和人一样都是一种数据处理系统。组成计算机的各种部件和设备负责完成数据处理各个环节的工作任务。既然大脑是人体数据处理体系的核心器官,因此把计算机称为“电脑”还是颇为贴切的。

习题

1. 举例说明数据名和数据值两个概念的差别。
2. 说明作为数据属性的“系”,和作为数据实体的“系”的差别。
3. 从数据表示类型的角度说明数123和数字串“123”的区别。
4. 讨论:“暖冬”包含着传达给人类的什么信息?
5. 举例说明数据和信息的联系和区别。
6. 举例说明为什么离开理解规则,信息就不能产生?
7. 举例说明依据不同的理解规则有可能从同一件事情的描述中抽取不同的信息。
8. 以“足球赛罚点球”为背景,分析守门员大脑中的数据处理过程。
9. 分析“到银行柜台取钱”过程中包含的数据处理基本环节。
10. 列举完成“银行取钱”数据处理过程所需要的计算机系统的典型设备。
11. 列举你知道的所有能够完成“数据加工”任务的工具。
12. 你觉得计算机和计算器有哪些主要差别?

1.2 计算机科学

计算机的发明导致人类一个新的学科门类的诞生,这就是计算机科学。数学、物理学、文学、哲学……是千百年来人类知识分门别类的积累,而计算机科学则是最近几十年以来围绕计算机的制造和应用而迅猛发展着的学科。社会的强烈需求是计算机科学快速发展的

外部诱因,而数学(特别是离散数学)、逻辑学(特别是数理逻辑)、电子学、光学等传统学科奠定了计算机科学坚实的基础。

1.2.1 理解计算机科学

可以认为:计算机科学是研制并利用计算机完成数据处理任务所涉及的理论、方法和技术的学科。完成这个学科任务,要研究基本理论、揭示基本规律,也要解决能够在计算机上实现的技术方法。理论支持技术、技术体现理论,两者相辅相成、互相融合是计算机科学的特点。因此,计算机科学的特征是科学性和工程性并重,理论性和实践性相结合。在短短几十年里,计算机科学就发展成为有众多分支领域、内容非常丰富、应用极其广泛的学科。

从学习的角度出发,必须依据理解纲领才能在比较短的时间里概要地认识计算机科学的主要内容。对于计算机科学与技术类专业学生在三四年内要学习的数十门专业课程而言,可以用“一个核心,三条纲领”来统帅之。

计算机的本质功能是数据处理。为了完成这个任务,计算机科学要解决三个方面的问题:

1. 数据的分层次表示方法学

要完成数据处理任务,首先要解决数据的表示问题。计算机科学面临的是:数据代表了现实世界的客观事物,形式多变而复杂;而计算机内部用来表示数据的符号又极其简单,就是两个二进制数字,一个0、一个1。显然无法用“平面”的方法来解决数据的表示问题。为此,计算机科学区分出数据的不同表示层次,在每个层次上定义相对独立的数据表示概念和方法,提出相邻层次数据表示手段的映射关系。这样就可以把问题“化繁为简、各个击破”了。实际上这是一种“先抽象、再具体”的思想方法。先在全局的抽象层次提出问题的解法,再在较为具体的层次上逐步求精,直到在计算机的底层得到问题的解为止。

可以把数据的表示划分为下列几个表示层次:

- (1) 现实世界层次上的数据对象;
- (2) 数据结构和信息结构层;
- (3) 程序设计语言层;
- (4) 机器层;
- (5) 物理层。

本书第5章叙述了各层次上数据表示的核心概念。而更加详细的学习就要依靠后续专业课程的展开了。

2. 数据加工的分层次表示方法学

数据的表示是为了对它们进行预定的加工,以得到人们所希望的信息。计算机绝对不是一种“自动机”,一切数据加工过程必须由人预先设定,交付计算机处理,最终机械地执行。计算机不过是台机器,它能执行的操作种类不单有限而且极其简单。不单如此,表示这百把种极其简单操作的最终符号仍然只能是二进制数字0和1,否则就无法在计算机内部实现。计算机科学还是使用分层次的方法学来解决问题,使用两个简单记号表示出千千万万种极为复杂的数据加工过程。

可以把数据加工的加工表示划分为下列几个表示层次: