

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 (第2版)

黄思曾 黄捷迅 编著



清华大学出版社

本书是根据教育部《计算机专业教学大纲》的要求，结合我国计算机专业教学改革的实际，在广泛吸收国内外优秀教材的基础上，由清华大学计算机系教师编写而成的。本书可作为高等院校计算机专业及相关专业的教材，也可供从事计算机工作的工程技术人员参考。

计算机科学导论教程

(第2版)

黄思曾 黄捷迅 编著

清华大学出版社
北京

内 容 简 介

本书共分7章,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、计算学科的知识领域。

和同类教材相比,本书的特点在于以“一个核心、三条纲领”来组织教学的内容。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中实现。按照本书叙述体系,初学者容易理解在后续课程中展开的专业概念,开始关注专业课程之间的内在关联。

本书适宜在高等学校用作“计算机科学导论”课程教材,也可供有兴趣了解计算机科学概貌的读者阅读。

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 13701121933

图书在版编目(CIP)数据

计算机科学导论教程/黄思曾,黄捷迅编著. —2版. —北京:清华大学出版社,2010.11

(高等学校计算机专业教材精选·计算机基础)

ISBN 978-7-302-23757-0

I. ①计… II. ①黄… ②黄… III. ①计算机科学—教材 IV. ①TP3

中国版本图书馆CIP数据核字(2010)第168102号

责任编辑:焦虹 顾冰

责任校对:白蕾

责任印制:杨艳

出版发行:清华大学出版社

地 址:北京清华大学学研大厦A座

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

邮 编:100084

社 总 机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62795954,jsjic@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015,zhiliang@tup.tsinghua.edu.cn

印 刷 者:北京四季青印刷厂

装 订 者:三河市漂源装订厂

经 销:全国新华书店

开 本:185×260

印 张:16.25

字 数:387千字

版 次:2010年11月第2版

印 次:2010年11月第1次印刷

印 数:1~3000

定 价:23.00元

产品编号:036857-01

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,我们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和计算机教学水平的提高,以及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

第2版前言

第1版出版后,不时会有教师给我发邮件,觉得这本书很适合用作本科学生教材,向我要教学大纲、教学进度表、考核大纲和电子课件以作参考,我乐意和同行们互相交流。这本书从来就不以“计算机科学导论”作为目标。它不过是笔者从许多年讲授“计算机科学导论”课程的过程中提炼出来的一本讲义而已。这是在书名中刻意加上“教程”二字的原因。

本课程面对一年级新生,通常只安排四五十学时,对任教教师来说,挑战在于选材、组织和演绎。讲得太“写意”或者太“工笔”都不行,太专业或者太不专业也不行。

在笔者的教学生涯中,计算机科学导论教材经历了几次变化。最早一批着眼于计算机入门操作和应用,类似今天仍然常见的以“应用基础”或“文化基础”来冠名的那些书;后来的一批有点全面介绍学科内涵的味道了,多以浓缩手法把有关领域内容写成一章一节;再往后,国内外开始出现以叙述纲领统率内容的著述,如以“算法”作为纲领贯穿教材,比如笔者用“数据表示、数据加工表示、计算机系统”三条线索来展开内容。

近年来,这个题材的新作继续出现,董荣胜的《计算机科学导论——思想与方法》一书站在“学科思想、方法这个较高的层面”来介绍学科,以此给学生提供对学科的认知基础,著述自成一格。这本书值得有关专业高年级本科生或者低年级研究生,甚至“计算机科学导论”课程的任教教师一读。

从院校总体上看,一年级大学生的逻辑思维能力和对计算机的了解程度参差不齐。不论教师选择什么样的叙述角度和层次,都受制于面对“菜鸟”们的“平均”接受能力,并非能够天马行空的。教师传道、授业、解惑之际,一定要顾及受众的感受。否则教师自我感觉讲得“天花乱坠”,而学生的反应却是“不知所云”。

这次修订增写了1.4节和第7章。希望学生在导论性课程中对学科能够有更全面的认知。大部分增加的内容可作为学生的阅读材料。但建议把7.2节列入教学计划。关于社会和职业的问题只由教师讲授,效果恐怕不会太好,应该设计更加生动的教学环节。

黄思曾

2010年8月于康乐园

第 1 版前言

“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。在较短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,一直是本课程的主讲教师。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。这是好事,说明课程为学生后续的专业学习奠定了良好基础,真正起到了“导论”应有的作用。

历届学生都表示,通过本课程的学习,他们清晰地了解到今后几年里要学什么,毕业之后会干什么,心里开始踏实。说明本课程还应该担负起学生专业教育的作用。

由于计算机技术发展十分迅速,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

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	6
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	14
习题	15
1.3 机内信息表示基础——二进制数	16
1.3.1 位置记数法	16
1.3.2 P 进制数	17
1.3.3 数制的转换规则	17
习题	18
1.4 关于计算学科及其专业	19
1.4.1 计算学科的由来	19
1.4.2 计算学科下属专业的划分	19
本章小结	20
第 2 章 计算机系统的组成	22
2.1 四大功能部件	22
2.1.1 内存	22
2.1.2 中央处理器	27
2.1.3 外部设备	30
2.1.4 总线	41
习题	43
2.2 计算机程序	44
2.2.1 机器指令	44
2.2.2 程序存储原理	47
2.2.3 冯·诺依曼结构体系	50

习题	51
2.3 计算机系统	52
2.3.1 硬件和软件	52
2.3.2 计算机系统的几种应用模式	54
习题	56
本章小结	57
第3章 计算机软件系统	58
3.1 软件系统	58
3.1.1 软件系统的组成	58
3.1.2 应用软件示例	59
3.1.3 获得软件的途径	63
3.1.4 软件版权保护	64
习题	64
3.2 操作系统	64
3.2.1 操作系统的目标	65
3.2.2 进程和中断	65
3.2.3 操作系统的资源管理功能	66
3.2.4 作业的管理调度方式	69
3.2.5 I/O 设备的输入输出控制	71
3.2.6 操作系统的常见种类	72
3.2.7 用户界面	73
3.2.8 有代表性的操作系统产品	74
习题	76
3.3 程序设计语言和语言处理软件	77
3.3.1 程序设计语言	77
3.3.2 程序设计语言的发展	78
3.3.3 程序设计和软件开发	82
3.3.4 程序/语言的处理过程	82
3.3.5 编译程序	83
3.3.6 上机作业的传统过程	85
3.3.7 程序作业工具	85
习题	86
本章小结	86
第4章 计算机通信与网络	87
4.1 数据通信概念	87
4.1.1 信息、数据与信号	87
4.1.2 数据的传输	88

4.1.3	传输介质和信道	91
4.1.4	信号传输的技术特性	92
4.1.5	数据交换	95
习题		96
4.2	计算机网络	97
4.2.1	计算机网络的组成	98
4.2.2	网络类型	99
4.2.3	网络协议和体系结构模型	100
习题		104
4.3	计算机局域网	104
4.3.1	局域网的拓扑结构	105
4.3.2	介质访问控制方法	106
4.3.3	局域网中常见的网络设备	107
习题		110
4.4	因特网	110
4.4.1	网络互联	110
4.4.2	因特网的地址编制和域名	111
4.4.3	因特网的典型应用	111
4.4.4	因特网的接入	117
习题		118
本章小结		119
第5章	数据表示方法	120
5.1	数据的分层表示方法学	120
5.1.1	现实世界层	120
5.1.2	信息世界层	121
5.1.3	高级语言层	122
5.1.4	机器层	122
5.1.5	物理层	123
习题		123
5.2	数的表示	123
5.2.1	无符号整数的表示	123
5.2.2	有符号整数的表示	124
5.2.3	实数的表示	128
5.2.4	复数的表示	130
习题		131
5.3	字符和汉字的表示	131
5.3.1	字符编码	132
5.3.2	汉字编码	134

5.3.3	汉字处理系统	137
	习题	138
5.4	图像和声音的表示	138
5.4.1	图像的表示	138
5.4.2	声音的表示	139
	习题	140
5.5	高级语言的数据表示手段	140
5.5.1	常量和变量	141
5.5.2	函数和表达式	142
5.5.3	数据类型	147
	习题	153
5.6	动态数据结构	154
5.6.1	数据结构的含义	154
5.6.2	数据结构的基本概念	155
5.6.3	线性表	158
5.6.4	树和二叉树	164
5.6.5	图	167
	习题	168
5.7	外存数据组织	169
5.7.1	文件	169
5.7.2	数据库	175
	习题	184
	本章小结	185
第 6 章 数据加工表示方法		
6.1	数据加工的表达层次	186
	习题	187
6.2	算法表示手段	188
6.2.1	算法和算法的表示	188
6.2.2	算法流程图	189
6.2.3	类程序设计语言的伪代码	194
	习题	197
6.3	结构化设计	197
6.3.1	三种流程结构	198
6.3.2	结构化方法	202
	习题	204
6.4	算法示例	204
6.4.1	顺序结构的算法设计	204
6.4.2	分支结构的算法设计	206

6.4.3 循环结构的算法设计.....	209
习题.....	212
6.5 软件的工程化开发方法	213
6.5.1 关于软件工程学.....	213
6.5.2 软件生命周期概念.....	214
6.5.3 传统的软件生命周期阶段.....	215
6.5.4 软件开发规范.....	220
6.5.5 软件开发模型.....	221
6.5.6 开发方法和工具.....	222
习题.....	225
本章小结.....	226
第7章 计算学科的知识领域	227
7.1 计算学科知识分类	227
7.1.1 数学.....	227
7.1.2 电子学.....	228
7.1.3 算法.....	229
7.1.4 计算机体系结构和组织.....	230
7.1.5 计算机软件系统.....	231
7.1.6 人机交互.....	232
7.1.7 程序设计.....	233
7.1.8 软件工程学.....	234
7.1.9 特定的应用系统.....	235
7.1.10 计算机网络	236
7.1.11 信息技术	236
7.1.12 职业素养	237
7.2 计算学科的社会与职业问题	238
7.2.1 计算技术对社会的冲击.....	238
7.2.2 相关法律.....	238
7.2.3 职业道德和职业素养.....	239
习题.....	240
本章小结.....	241
参考文献	242

第1章 认识计算机和计算机科学

计算机问世仅仅六十多年,已给人类社会带来了翻天覆地的变化。如果说,蒸汽机和电的发明带来工业革命,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,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,已达到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“无所不在、无所不能”的印象使外行人对计算机产生了一种神秘感,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。

因此,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)这个词的时候,很容易使我们联想到“数”,如实验数据、财务数据和经济数据等。计算机科学赋予“数据”这个术语更加本质的含义。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……不管你看得见看不见,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的一组**特征**所组成的集合。例如:

{学号,姓名,系,专业,年级,年龄,性别,血型,口音,脸形,……}

这样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。

一般来说,一种事物几乎会有无数个各种各样的特征,它们表示了事物各个不同方面的性质。由业务处理的需要出发,从事物众多特征中选取出有限个**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,而是体现专业人员对客观事物进行分析、分类、拆分、重组和选取等一系列的思维动作,这样才能用数据的概念来表示客观世界事物。

2. 数据的名和值

要从**数据名**(name)和**数据值**(occurrence)两个不同的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。例如:

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:

(09001, 张三, 计算机系, 软件开发, 一年级)

表示了特定的一个学生。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。比如学生的名字可以是张三、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)。常见的几种数据形式是**数**、**文字**(字符/汉字的串)、**图像**、**图形**、**视频**(活动的图像)和**音频**(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;一张照片是一幅图像,表示了某人的外貌特征,当然是数据的一种表示形式。图形和声音也如此。

现代计算机可以用数、文字、图像和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”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**都必须记录在某种**介质**(media, 又称**媒体**)上。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。计算机里,常见的数据存储介质是**磁盘**、**光盘**、**磁带**、**内存储器**、早期用过的**纸质穿孔带**和**穿孔卡**。它们采用**电磁材料**、**光电材料**和**半导体电子电路**等制造。

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,信息是有严格定义的。

1. 信息的定义

信息是由客观事物传递的知识。

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和表示,而是指透过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

信息必须依赖三个要素而存在:(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(景象形式)就是一种数据,是信息的产生源头;“看见晚霞的人”是个接收者;“气象知识”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,如何能想到天晴?晚霞灿烂,就是没有人注意,或者看见晚霞的人是个小孩子,只会觉得云彩像猫、像狗,肯定不会由此产生“明天天晴”这个信息。

就是说,信息的接收者必须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。因此,有人对你说了一大段话,你的回应是:“什么意思呀?”。这就表示你没法理解说话里蕴藏着的信息。

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,数据是信息的外在形式。

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,我们的第一反应,这是某一位中国人的名字,不可能认为它包含了“某块石头姓张”这样一个信息。

当然,引用不同的理解规则必然会在同一个数据源中得到不同的信息。在一辆汽车上面看到 BMW,按照常规理解,意味着这是一辆“宝马”。如果按照搞笑规则去理解,BMW 的信息含义会是“别摸我”。

此外,信息理解规则往往会有局部性,局限在一定的地域、民族、国家、文化、行业当中。在广东的饭馆里,食客把茶壶盖揭开,服务员就会接收到其中的信息:“客人要我续水呐”。搁在北方,恐怕没有人能够从这个动作中提取出这个信息。

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。人们关注的是数据蕴含的信息,没有信息内容的数据形同垃圾。因此,即使在计算机专业的专业范畴里,往往也可以不加区分地交替使用数据和信息这两个术语。虽然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)是一个过程,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,数据处理也往往被称为信息处理(information processing)。

数据代表万物,对数据的加工动作也是含义广泛的。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,后者用来描述各种非数学意义的数据变换操作。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,各个应用领域都会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,然后输入大脑“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。过马路之前先站在边上左右张望,这是用眼睛收集数据。看到一辆车飞驰而来,信息马上输入大脑,大脑开始紧张的运算加工:先检索出一条小时候存储过的知识,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,再决定是站住不动,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会包含几个基本环节: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。

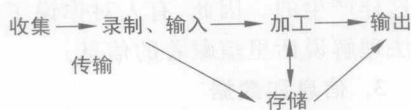


图 1-1 数据处理的基本环节

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。但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”意味着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”意味着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。于是改进不断继续,发明文字和记数的方法以记录数据,发明纸用作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) 五官负责数据的输入。眼睛看到的是影像、文字和数等形式的数据,耳朵听到的是声音形式的数据,嘴尝到的味道,鼻子嗅到的气味,皮肤触摸到的物体表面状况都是客观事物的各种特征,即数据属性。

(2) 大脑是最重要的数据处理器官。大脑要负责存储数据、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正常能力。

(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) 至于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,计算机也不只是一种数学意义上的计算设备。计算机

(computer)是一种数据处理机,能够以前所未有的性能来完成数据处理所有环节的任务。日常生活中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)输入不同形式的数。比如使用键盘输入字符或汉字形式的数,用扫描仪输入图像形式的数,用读卡机输入有特别表示形式的数等。

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。比如用显示器来输出字符或图像形式的数,用打印机把数输出到纸上等。

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(也可称为主存储器)用来存放工作时要用到的数。还会使用若干种外存储器来存放需要长期保存的数,如磁盘、光盘、U盘和磁带等。

4) 中央处理器

中央处理器(Central Processing Unit,CPU)包括运算器、寄存器和控制器等主要部件。CPU依靠人给定的指令来完成所有的数据处理动作,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路,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,可以分为数据总线、地址总线和控制总线等几部分。

如上面所述,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,那么许多人喜欢把计算机称为“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. 列举你知道的完成“银行取钱”数据处理过程所需要的计算机系统典型设备。
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?