

高等学校教材

计算机导论 *Introduction to Computers*

宋华珠 钟 珞 主编



高等教育出版社
HIGHER EDUCATION PRESS

高等学校教材

计算机导论

Jisuanji Daolun

宋华珠 钟 珞 主编



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容提要

本书是为计算机导论课程编写的教材。全书共分7章，分别是“信息与计算机”、“从传统计算机到‘无处不在’的计算”、“走近软件”、“程序设计语言”、“计算机中的数据”、“计算机应用”和“计算机应用的主要发展方向”，内容涵盖了计算机的基本知识以及计算机科学与技术的应用领域和发展前沿。本书由浅入深，循序渐进，紧密结合新技术的发展，并配有电子教案和课程网站等教学资源。

本书可作为高等学校计算机及相关专业计算机导论类课程教材，也可作为计算机及相关领域从业人员的参考读物。

图书在版编目(CIP)数据

计算机导论 / 宋华珠, 钟珞主编. —北京: 高等教育出版社, 2013. 7

ISBN 978 - 7 - 04 - 037504 - 6

I. ①计… II. ①宋… ②钟… III. ①电子计算机 - 高等学校 - 教材 IV. ①TP3

中国版本图书馆 CIP 数据核字(2013)第 113021 号

策划编辑	刘 艳	责任编辑	刘 艳	封面设计	于 涛	版式设计	马敬茹
插图绘制	郝 林	责任校对	刘丽娴	责任印制	朱学忠		

出版发行	高等教育出版社	咨询电话	400 - 810 - 0598
社 址	北京市西城区德外大街 4 号	网 址	http://www.hep.edu.cn
邮政编码	100120		http://www.hep.com.cn
印 刷	高教社(天津)印务有限公司	网上订购	http://www.landaco.com
开 本	787mm × 1092mm 1/16		http://www.landaco.com.cn
印 张	19	版 次	2013 年 7 月第 1 版
字 数	430 千字	印 次	2013 年 7 月第 1 次印刷
购书热线	010 - 58581118	定 价	27.90 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 37504 - 00

前 言

随着计算机科学与技术领域的新成果不断涌现,计算机的发展已经进入了一个快速而又崭新的阶段。计算机性能越来越高、分布越来越广,在引领科技进步的同时不断地融入、延伸到人们日常生活的方方面面,计算机技术已成为影响广泛、极具活力的科学技术领域之一。因此,在当今社会,掌握基本的计算机技术是非常必要的。为了使读者对计算机技术有一个总体的了解,我们特意编写了本书。

本书涵盖了计算机的基础知识及新的发展技术,并按照人们一般的思维习惯科学地编排各章节内容,以帮助读者理解和探究计算机知识。本书共分7章。第1章主要介绍了信息的定义、信息的种类及信息的度量,以及计算机信息处理的过程、原理,最后简述了计算机科学的定义。第2章从实现计算所需要具备的物质条件/硬件环境(包括硬件及体系架构)的角度,依次介绍了个人计算机、计算机系统结构、移动终端、计算机网络、云计算和物联网,使读者在从传统计算机到“无处不在”的计算的发展过程中,了解计算机硬件及其体系结构的发展与变化。第3章主要介绍了运行于硬件之上的软件,首先简述了软件的定义、分类及软件架构,然后介绍了常用的计算机和智能手机操作系统软件;描述了软件开发过程,重点介绍了用户界面和用户体验,并对软件行业的相关就业岗位进行了介绍,以方便读者了解软件开发的业界需求,确定专业方向。第4章主要介绍了程序设计语言的基本知识,包括高级程序设计语言、并程序序设计语言。第5章介绍了数据、数据结构、算法、数据库以及数据中心的有关知识,反映了计算机中数据处理的不同视角。第6章介绍了计算机技术的主要应用,包括多媒体技术、人工智能、嵌入式系统、电子商务、移动互联网、信息系统集成与决策支持。第7章介绍了计算机应用的主要发展方向,包括高性能化、网络化和智能化,为读者呈现出一个更加广阔、充满生机的计算机应用场景。

本书内容由浅入深,循序渐进,紧密结合计算机新技术的发展。在体例方面,每章结尾安排有:本章小结,以便于读者回顾本章的主要知识;拓展阅读,以供读者开阔视野及培养学习兴趣;习题,以帮助读者巩固所学的知识。在配套资源方面,本书配有电子教学课件、相关视频以及课程网站(Web网址:www.computingdream.com;手机网址:www.wap.computingdream.com)等教学资源。

本书由宋华珠、钟珞主编,第1章~第3章、第5章、第7章由宋华珠、杨书勇、肖聪编写,第4

II 前 言

章由钟珞、孙骏、王小雪编写,第6章由宋华珠、孙骏、王小雪编写。武汉大学何炎祥教授审阅了本书,在此表示感谢!限于作者的学识和时间,书中不当和差错之处难免,请读者发送电子邮件至 shuaz@whut.edu.cn 指正。

编 者
2013 年 4 月

目 录

第1章 信息与计算机	1
1.1 信息	2
1.1.1 信息的定义	2
1.1.2 信息的种类	4
1.1.3 信息的度量	4
1.2 计算机信息处理	7
1.2.1 位模式的信息表示	7
1.2.2 计算机信息处理过程	12
1.3 计算机信息处理的几个例子	13
1.3.1 从方程看代数	13
1.3.2 从随机生成迷宫地图看几何	15
1.3.3 从测谎看逻辑	17
1.4 图灵机	19
1.5 计算机科学的定义	21
1.5.1 计算机科学的学科形态	21
1.5.2 计算机科学的含义和基本问题	23
本章小结	23
参考资料	24
习题	24
扩展阅读 1-1 信息战	25
扩展阅读 1-2 信息的搜索——百度使 用小技巧	26
第2章 从传统计算机到“无处不在”的计算	29
2.1 个人计算机	30
2.1.1 冯·诺伊曼体系结构	30
2.1.2 个人计算机的硬件子系统	31
2.1.3 工作站	37
2.2 计算机系统结构	40
2.2.1 计算机系统结构的基本概念	40
2.2.2 计算机组成	41
2.2.3 计算机实现	43
2.3 移动终端	44
2.4 计算机网络	45
2.4.1 计算机网络的定义与分类	45
2.4.2 计算机网络的硬件组成	47
2.4.3 网络拓扑结构	51
2.4.4 计算机网络体系结构	52
2.5 云计算与物联网	57
2.5.1 云计算	57
2.5.2 物联网	65
本章小结	69
参考资料	69
习题	70
扩展阅读 2-1 虚拟化技术	71
扩展阅读 2-2 计算机选购技巧和笔记本电脑 硬件检测及保养技巧	75
第3章 走近软件	79
3.1 软件的定义	79
3.2 软件的分类	81
3.2.1 按软件的功能分类	81
3.2.2 按软件的规模分类	82
3.2.3 按软件的工作方式分类	82
3.3 软件架构	83
3.3.1 软件架构的定义	83
3.3.2 软件架构的分层	83
3.4 操作系统	85
3.4.1 什么是操作系统	85
3.4.2 DOS	86
3.4.3 Windows 操作系统	87
3.4.4 UNIX 操作系统	90
3.4.5 Linux 操作系统	92

II 目 录

3.4.6 Mac OS	95	4.4.3 OpenMP	138
3.4.7 智能手机操作系统简介	96	本章小结	138
3.5 软件的开发	98	参考资料	139
3.5.1 软件工程	98	习题	139
3.5.2 软件开发过程	98	扩展阅读 4-1 唐纳德·克努特与程 序设计艺术	139
3.5.3 用户界面与用户体验	101	扩展阅读 4-2 Android 手机应用开发	141
3.6 软件行业职业简介	106		
3.6.1 项目管理类	106	第 5 章 计算机中的数据	144
3.6.2 开发类	106	5.1 数据	145
3.6.3 测试类	108	5.1.1 数据的定义	145
3.6.4 系统、安全及维护类	108	5.1.2 数据和信息的区别	145
3.6.5 销售类	110	5.1.3 计算机常用的编码方式	146
本章小结	111	5.2 数据结构	152
参考资料	112	5.2.1 数据结构的定义	152
习题	112	5.2.2 数据的逻辑结构	153
扩展阅读 3-1 从 Linux 看开源软件	113	5.2.3 数据的物理结构	154
扩展阅读 3-2 从 Google Chrome 浏览器 看用户体验	115	5.2.4 常见的数据结构	154
		5.3 算法	160
第 4 章 程序设计语言	118	5.3.1 算法的定义	160
4.1 程序设计语言简介	119	5.3.2 算法示例	161
4.2 程序设计语言的分类及编程模式	121	5.3.3 定义动作	162
4.2.1 机器语言	121	5.3.4 细化	162
4.2.2 汇编语言	122	5.3.5 泛化	163
4.2.3 高级语言	122	5.4 数据库	164
4.2.4 编程模式	125	5.4.1 数据库的定义及特点	164
4.3 高级编程语言介绍	130	5.4.2 数据库的体系结构	165
4.3.1 BASIC 语言	130	5.4.3 数据库模型	167
4.3.2 PASCAL 语言	130	5.4.4 关系的操作及结构化查询语言	168
4.3.3 C 语言	131	5.4.5 数据库管理系统	171
4.3.4 C++ 语言	131	5.4.6 主流数据库	172
4.3.5 Java 语言	132	5.5 数据中心	173
4.3.6 Ada 语言	134	5.5.1 数据中心的定义	173
4.3.7 Logo 语言	135	5.5.2 数据中心的组成	174
4.3.8 脚本语言	136	5.5.3 新一代数据中心	174
4.4 并行程序设计	136	本章小结	178
4.4.1 并行程序设计语言	136	参考资料	178
4.4.2 并行程序设计语言的编译过程	137	习题	179

扩展阅读 5-1 经典算法之一——汉诺塔·····	179	6.6.5 决策支持系统的目标和功能·····	247
扩展阅读 5-2 企业数据中心·····	181	6.6.6 决策支持系统的基本特征·····	248
第 6 章 计算机应用 ·····	186	6.6.7 决策支持系统与管理信息 系统的关系·····	248
6.1 多媒体技术·····	187	6.6.8 决策支持系统与专家系统·····	249
6.1.1 多媒体技术的基本概念·····	187	本章小结·····	252
6.1.2 多媒体技术的应用·····	188	参考资料·····	252
6.1.3 多媒体技术的关键技术·····	191	习题·····	253
6.1.4 常用多媒体工具软件简介·····	193	扩展阅读 6-1 从 64 位计算看 64 位 应用软件·····	254
6.2 人工智能·····	197	扩展阅读 6-2 C/S 与 B/S 分布式 应用的特点·····	255
6.2.1 人工智能简介·····	197		
6.2.2 知识表示·····	199	第 7 章 计算机应用的主要发展方向 ·····	258
6.2.3 人工智能的主要研究与应用 领域·····	213	7.1 高性能化·····	259
6.3 嵌入式系统·····	217	7.1.1 高性能计算简介·····	259
6.3.1 嵌入式系统简介·····	217	7.1.2 高性能计算机的研究现状·····	260
6.3.2 嵌入式系统的组成·····	218	7.1.3 高性能计算机的发展展望·····	261
6.3.3 嵌入式系统的应用·····	219	7.2 网络化·····	266
6.3.4 嵌入式系统的发展趋势·····	222	7.2.1 网络化定义·····	266
6.4 电子商务·····	223	7.2.2 网络化发展现状·····	266
6.4.1 电子商务的概念·····	223	7.2.3 网络化发展展望·····	269
6.4.2 电子商务的框架结构·····	224	7.3 智能化·····	274
6.4.3 电子商务的分类·····	226	7.3.1 智能化的定义·····	274
6.4.4 电子商务的应用与案例·····	229	7.3.2 智能化发展现状·····	274
6.4.5 电子商务的发展趋势·····	232	7.3.3 智能化发展展望·····	275
6.5 移动互联网·····	233	本章小结·····	279
6.5.1 移动互联网的形成及分类·····	233	参考资料·····	279
6.5.2 移动互联网的特点·····	234	习题·····	280
6.5.3 移动互联网体系及热点技术·····	236	扩展阅读 7-1 信息技术与纳米技术、生物 技术、认知学科会聚·····	280
6.5.4 移动互联网的应用·····	240	扩展阅读 7-2 信息安全·····	283
6.5.5 移动互联网的发展趋势·····	241		
6.6 信息系统集成与决策支持·····	242	附录 A 计算学科的问题空间及专业能力 培养 ·····	287
6.6.1 信息系统集成的概念·····	242	附录 B 图灵奖 ·····	292
6.6.2 信息系统集成的原则和方法·····	243		
6.6.3 信息系统集成的体系结构·····	245		
6.6.4 决策支持系统概述·····	246		

第 1 章 信息与计算机



图书阅览室里的信息

从上面这幅图中可以看到,在宁静的图书阅览室中:

一个外国女孩正在看书;

这个图书阅览室很明亮,里面的书很多,并且排列得整齐、紧密;

这个女孩可能是大学生,也可能是职员;

她可能比较高,有着大眼睛、高鼻梁,扎了个马尾辫;

她手捧一本被翻开的厚书,正在思考问题;

.....

一幅简单的图片,竟能传达给人们如此多的信息。

可见,在纷繁复杂的世界中,信息无处不在。对于事物的信息,有的人们可以通过自己的感官直接感受到,有的则需要通过对事物进行分析和研究才能得到。人们对这些信息进行了解和掌握就可以将其转变成自身的知识。但是,如何能让计算机这种只通过 0、1 存储信息的机器来记录和描绘人们精彩纷呈的世界呢?

本章主要介绍什么是信息、信息的度量和表示,以及计算机是如何存储和处理信息的,由此引入计算机科学,讲述计算机科学的含义、学科形态及其基本问题。

1.1 信息

“信息”是目前使用频率很高的一个概念,也是很难描述清楚的一个概念。据不完全统计,关于信息的定义有 100 多种,它们从不同的侧面、不同的层次揭示了信息的某些特征和性质,但至今仍没有统一的、能为各界普遍认同的定义。

1.1.1 信息的定义

1. 维纳的信息定义

1948 年,美国数学家、控制论的奠基人维纳(Norbert Wiener)在《控制论——关于在动物和机器中控制与通信的科学》中首先将“信息”上升到“最基本概念”的位置,指出“信息是信息,不是物质,也不是能量。”后来在《人有人的用处》一书中,他又提出“信息是人们在适应外部世界并使这种适应反作用于外部的过程中,同外部世界进行互相交换的内容的名称”。

2. 信息是差异类的定义

1975 年,意大利学者朗高在《信息论:新的趋势与未决问题》一书的序言中提出:“信息是反映事物的形式、关系和差别的东西。信息存在于客体间的差别之中,而不是存在于客体之中”。英国学者阿希贝则认为,信息的本性在于事物本身具有变异性。

3. 钟义信的信息定义

北京邮电大学钟义信教授分别在本体论和认识论层次上提出了信息定义体系。其中,本体论层次的信息是指事物运动的状态及其变化方式的自我表述。认识论层次的信息是指认识主体所感知或所表述的事物运动的状态和方式,包括这种状态或方式的形式、含义和效用。

① 语法信息。它是指主体所感知或所表述的事物运动状态和方式的形式化关系。

② 语义信息。它是指认识主体所感知或所表述的事物运动状态和方式的逻辑含义。

③ 语用信息。它是指认识主体所感知或所表述的事物运动状态和方式相对于某种目的的效用。

④ 先验信息。观察者关于某事物的先验信息是指他在观察之前通过某种途径所感知的该事物运动的状态和方式。

⑤ 实得信息。在观察过程中观察者关于某事物的实得信息,是指他通过观察所新

感知到的该事物运动的状态和方式。

⑥ 实在信息。它是指在理想观察条件下,观察者所获得的关于该事物实际的运动状态和方式等全部信息。

上述的语法信息、语义信息、语用信息、先验信息、实得信息和实在信息都是从不同角度来理解的认识论层次上的各种信息的概念。

因此,可以说“信息是被反映的物质属性”或者“信息是事物属性的反映”。

4. 香农的信息定义

假定事物状态可以用一个以经典集合论为基础的概率模型来描述,则信息就是用来消除不确定性的东西,或信息是事物运动状态或存在方式的不确定性描述。

但要在实际中寻找一个合适的概率模型往往是非常困难的,甚至这样一种模型是否存在还值得探讨。此外,信息有很强的主观性和实用性,但该定义没有考虑信息接收者的主观特性和主观意义,不顾信息的具体含义、具体用途、重要程度和可能引起的后果等因素,这就与实际情况不完全一致。

5. 种加属差定义

种加属差定义亦称“真实定义”、“实质定义”或“事物定义”。定义项由被定义概念的邻近的种概念和表示被定义概念属差的概念所组成。其公式是:

被定义概念(属概念)= 邻近的种概念+属差

当要给“人”下定义时,首先要找到与“人”这一概念最近的“种概念”——动物,然后可以说“人是一种动物”。但仅仅这样说的不完整,还必须找出“人”这一“属概念”与“动物”这一“种概念”的其他“属概念”(如“鸡”、“鸭”、“牛”、“羊”……)之间的“差异”(属差)来,“人”和“鸡”、“鸭”、“牛”、“羊”等之间的“属差”是什么呢?是“制造生产工具”,从而得出“人是能制造生产工具的动物”这一完整定义。

又如“记忆是人们过去的生活实际中认识过的事物或做过的事情在头脑中遗留的印迹。”其中:“记忆”是被定义概念;“印迹”是比被定义概念外延大一级的“邻近的种概念”;而“人们过去的生活实际中认识过的事物或做过的事情在头脑中遗留”则是“记忆”区别于其他“印迹”的特有属性,即差别,称为“属差”。

简单地说,种加属差就是一个信息相加的运算,是一种对事物的定义方式,在数学界中则表示为集合的交集。由此亦可看出信息是可加的。在此基础上,德雷斯基(Fred Dretske)在《知识与信息流》一书中对信息的定义是:“粗略地说,信息或信号所携带的正是我们需要知道的”。

总的来说,一切客观存在都有信息。信息一般是指事物释放出来的消息、情报、指令、数据或信号,它是人们认识客观世界的媒介。

你能说出图 1.1 所蕴涵的信息吗?

随着社会的进步与经济的发展,信息量越来越大,要求传递与处理的信息越来越

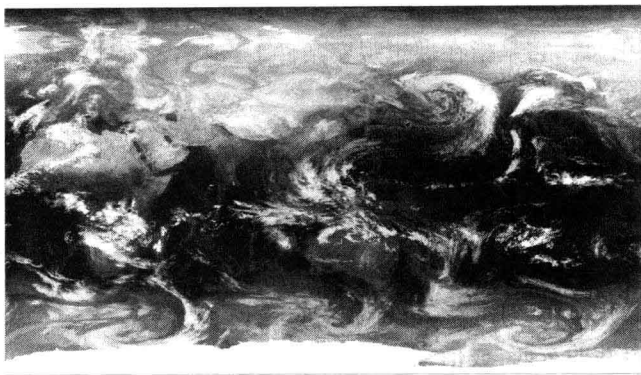


图 1.1 卫星云图

多,信息的重要性也日益凸显。信息已成为人类社会赖以存在与发展不可或缺的基本要素之一。

1.1.2 信息的种类

按照产生信息的物体的性质,可以将信息分为自然信息(声、光、热、电等),生物信息(生物为繁衍生存而表现出来的各种形态和行为,如遗传信息、生物体内的交流信息、动物种群内的交流信息等),机器信息和社会信息等。

按照人类活动领域,可以将信息分为科技信息、经济信息、政治信息、军事信息、文化信息等。

按照信息所依附的载体,可以将信息分为文献信息、口头信息、电子信息、生物信息等。

按照携带信息的信号性质,可以将信息分为连续信息、半连续信息和离散信息。

按照信息所起的作用,可将信息分为无用信息、有用信息和干扰信息。

图 1.2 显示了多种类型的信息。

1.1.3 信息的度量

信息是一个很抽象的概念。人们常常说信息多,或者信息少,但却很难说清楚信息到底有多少。例如,一本 50 万字的中文书到底有多少信息量?

一条信息的信息量大小和它的不确定性有直接的关系。比如说,我们要做一件非常不确定的事,或是要了解以前一无所知的事情,就需要大量的信息。相反,如果我们对某件事已经有了较多的了解,则不需要太多的信息就能把它搞清楚。所以从这个角度,可以认为信息量的度量就等于其不确定性的多少。

再例如,假设在甲、乙两个盒子中各装有 60 个大小、质地完全相同,颜色不同的球。

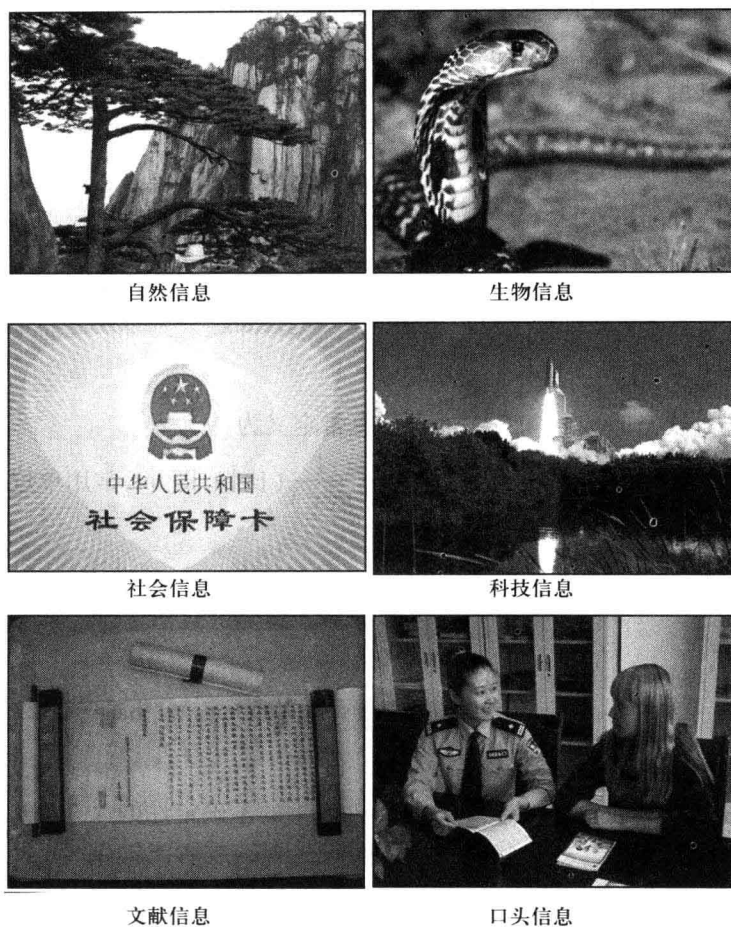


图 1.2 多种类型的信息

甲盒子中有红球、白球各 30 个,乙盒子有红球、黄球、蓝球各 20 个。现随意从甲盒子或乙盒子中取出一个球,并猜测取出的是什么颜色的球。

可见,从甲盒子中摸出红球的可能性为 $1/2$,不确定性为 $1/2$,小于从乙盒子中摸出红球的不确定性 $2/3$;同理,从甲盒子中摸出白球的不确定性为 $1/2$,也小于从乙盒子中摸出白球的不确定性 $2/3$ 。故不确定性的大小与可能发生的信息数目及各信息发生的概率有关。

由于不确定性的能够度量,因此,信息是可以度量的。下面介绍量度信息的自信息量和信息熵。

信息论的基本假设如下。

① 信息是可以用一组符号来编码的。

② 信息的产生和传输是可以用概率论和随机过程来描述的。

③ 从概率角度看,同样的信息包含同样的信息量。

自信息量是指一个事件本身所包含的信息量,它是由事件的不确定性所决定的。例如,“抛掷一枚硬币的结果是正面”就是抛掷一枚硬币所包含的信息量,即自信息量是对单个事件不确定性的度量。

假设离散随机变量 $X \sim p(x_i)$ 的数学模型为

$$X \sim \left[\begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p(x_1) & p(x_2) & \cdots & p(x_n) \end{pmatrix} \right]$$

其中 $x_i \geq 0, i = 1, 2, \dots, n$, 且 $\sum_{i=1}^n p(x_i) = 1$ 。

随机变量 X 产生的事件 x_i 的自信息量定义为

$$I(x_i) = -\log_2 p(x_i) = \log_2 \frac{1}{p(x_i)} \quad (\text{比特, 即 bit, 常用单位}) \quad (1.1)$$

或者

$$I(x_i) = -\ln p(x_i) = \ln \frac{1}{p(x_i)} \quad (\text{奈特, 即 nat, } 1 \text{ nat} \approx 1.443 \text{ bit}) \quad (1.2)$$

或者

$$I(x_i) = -\lg p(x_i) = \lg \frac{1}{p(x_i)} \quad (\text{哈特, 即 hat, } 1 \text{ hat} \approx 3.322 \text{ bit}) \quad (1.3)$$

其示意图如图 1.3 所示。

在事件发生以前,自信息量等于事件发生的不确定性的量;在事件发生以后,自信息量表示事件所含有的或所能提供的信息量。因此,自信息的含义为:收到某信息获得的信息量等于不确定性减少的量。但是,自信息量不能用来作为随机变量 X 的信息度量。

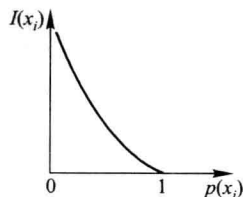


图 1.3 自信息示意图

1948 年,美国数学家、信息论的奠基人克劳德·艾尔伍德·香农(Claude Elwood Shannon)在《通信的数学理论》中提出了“信息熵”的概念——信息中排除了多余重复信息之后的平均信息量。例如,抛掷一枚硬币的试验所包含的信息量。

离散随机变量 $X \sim p(x_i)$ 的信息熵是从平均意义上对信息不确定性的度量,也称为平均自信息量,定义为

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (1.4)$$

其中,随机变量 X 由 n 个事件 x_i 构成,事件 x_i 出现的概率为 $p(x_i)$,这个公式和热力学中的熵的计算方式一样,故也称为信息熵,从式(1.4)可知,当各个事件出现的概率相等,即“不确定性”最高时,信息熵最大。故信息熵可以视为对“不确定性”或“选择的自

由度”的度量。

再来看一个例子。对于世界足球锦标赛(简称“世界杯”),人们都会关心谁是冠军。假如我错过了看“世界杯”,赛后问一个知道比赛结果的观众“哪支球队是冠军”?他不愿意直接告诉我,而要让我猜,并且我每猜一次,他都要在收一元钱后才肯告诉我是否猜对了,那么我需要付给他多少钱才能知道谁是冠军呢?我可以将球队按照从1号到32号编上号,然后提问:“获得冠军的是1~16号中的球队吗?”假如他告诉我猜对了,我会接着问:“获得冠军的是1~8号中的球队吗?”假如他告诉我猜错了,我自然知道获得冠军的是9~16号中的球队。这样只需询问五次,我就能知道哪支球队是冠军。这里“谁是世界杯冠军”这条信息的信息量值五块钱。

当然,香农不是用钱,而是用比特(bit)这个概念来度量信息熵。在计算机中所有的数据都是以二进制来表示的,一个二进制代码称为一个比特。

1.2 计算机信息处理

计算机能处理的信息一般有文本、数值、音频、图像和视频五种形态。这些形态的信息可以按照某种规律或约定俗成的程序进行转化,例如,扫描仪扫描照片,将照片存储在计算机中,这就实现了将图像信息转化为数值信息。

1.2.1 位模式的信息表示

计算机外部的各种不同类型的数据都是在按照统一的数据格式进行转换后存入计算机,并在从计算机输出时再还原回来的,这种统一的格式称为位模式。位模式是一个序列,有时也称为位串,它是0和1的组合。通常长度为8的位模式被称为1个字节(byte),一个字由若干字节组成。

下面简单介绍5种信息形态的表示。

1. 文本的表示

文本格式的信息常用编码的方法来表示,文本中每一个不同的符号(如字母表中的字母或标点符号)被指定为一个唯一的位模式。这样,文本就被表示成一个长的位串,其中相应的位模式代表了原文中相应的符号。

英文一般采用ASCII(American standard code for information interchange,美国标准信息交换代码)编码、Unicode编码等,而中文则大多采用与前者兼容的GB 2312—1980编码。

2. 数值的表示

当被记录的信息是纯数值时情况就会有所不同。假设存储数 25, 如果仍采用每个符号一个字节的 ASCII 编码存储它, 则总共需要 16 位, 这样使用 16 位能存储的最大数是 99。一个更有效的方法是以二进制形式存储数值。

二进制记数法是一种仅用数字 0 和 1 来表示数值的方法, 它是位置记数法的一种。用位置记数法表示的数值, 其表示式中的每一个位置都与一个称为“权”的量相联系。例如, 在十进制表示式 237 中, “7”的位置与量 $1(10^0)$ 相联系, “3”在与量 $10(10^1)$ 相联系的位置上, “2”在与量 $100(10^2)$ 相联系的位置(如图 1.4(a)所示)上。这些与位置相关的量是 10 的次方数, 这些量就称为相应位置的权。整个表示式所代表的值, 就是由每个位置的数字与该位置的权相乘, 然后把所得乘积相加而得到的。而在二进制表示式 1100 中, 从右至左, 最右边的“0”对应的权是 $1(2^0)$, 其左边的“0”所在的位置与权 $2(2^1)$ 相联系, 紧接着的“0”在与权 $4(2^2)$ 相联系的位置上, 最左边的“1”所在的位置对应的是权 $8(2^3)$ (如图 1.4(b)所示)。

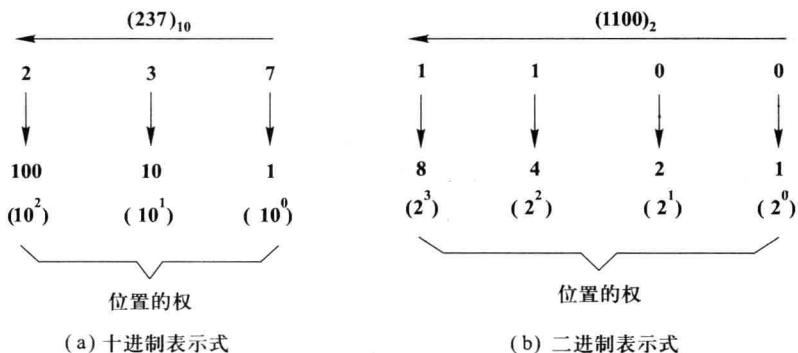


图 1.4 十进制数与二进制数的表示

将二进制表示式所代表的值转换为十进制表示式的过程, 就是将这个表示式每位数字与其所在位置所对应的权相乘并将各个乘积相加的过程。例如, 二进制 100111 所代表的十进制数是 39, 如图 1.5 所示。

对于二进制而言, 数据分为无符号数和有符号数两类。

无符号数: 就是二进制数的全部二进制位均表示数值位(没有符号位), 相当于数的绝对值。

例 1.1 `unsign1 = 0101`: 表示无符号数 5;

`unsign2 = 1001`: 表示无符号数 9。

有符号数: 约定二进制数的最高位为符号位, “0”表示正数, “1”表示负数。这种在计算机中使用的数的表示形式称为机器数。根据编码的不同, 又可将其分为原码、反码和补码。

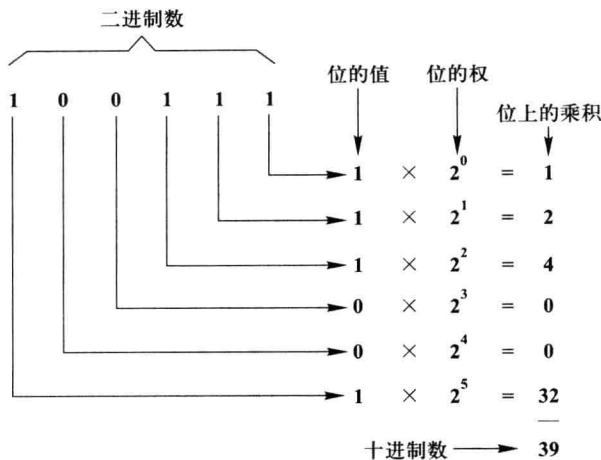


图 1.5 二进制数 100111 转换为十进制数的过程

设二进制整数的原码形式为 $x_s x_1 x_2 \cdots x_n$, 其中 x_s 表示符号位。

例 1.2 $x_1 = 1101, [x_1]_{\text{原}} = 00001101;$
 $x_2 = -1101, [x_2]_{\text{原}} = 10001101。$

在原码表示中, 0 有两种不同的表示形式: $[+0]_{\text{原}} = 00000000, [-0]_{\text{原}} = 10000000。$

在反码表示中, 对于正数, 反码表示与其原码表示形式相同; 对于负数, 将原码的数值部分按位取反。0 有两种不同的反码表示形式: $[+0]_{\text{反}} = 00000000, [-0]_{\text{反}} = 11111111。$

补码的符号位表示方法与原码相同, 其数值部分的表示与数的正负有关。

当 x 为正数时, $[x]_{\text{补}} = [x]_{\text{原}} = x。$

当 x 为负数时, 由 $[x]_{\text{原}}$ 转换为 $[x]_{\text{补}}$ 的方法有如下两种。

- ① $[x]_{\text{原}}$ 除符号位外的各位取反加“1”。
- ② 自低位向高位, 尾数的第一个“1”及其右部的“0”保持不变, 左部的各位取反, 符号位保持不变。

在补码表示中, 0 的表示形式是唯一的: $[+0]_{\text{补}} = [-0]_{\text{补}} = 00000000。$

例 1.3 $x_1 = 1101, [x_1]_{\text{补}} = 00001101;$
 $x_2 = -1101, [x_2]_{\text{补}} = 11110011。$

例如, 在 C 语言中, 将 20 定义为 int 类型, 那么它在计算机中存储的形式为二进制, 即 10100。

3. 音频的表示

音频表示声音或音乐。音频本质上与前面所讨论的文本不同。文本由可数的实体(文字)组成, 例如, 我们可以数出文本中文字的数量, 但音频是不可数的。它是随时间