



高职高专“十一五”规划教材

计算机

信息技术基础

狄寿霞 主编



化学工业出版社

高职高专“十一五”规划教材

计算机信息技术基础

狄寿霞 主编



化学工业出版社

· 北京 ·

本书是信息技术基础的普及教材,基于工作过程为导向的指导思想组织编写。详细介绍了信息技术基本知识及计算机常用的一些基本操作,包括信息技术基础知识、计算机系统应用、Windows XP 系统应用、Word 2003 文字处理软件、Excel 2003 电子表格、PowerPoint 2003 演示文稿制作、FrontPage 2003 网页制作、计算机网络基础知识、Internet 基础知识等基本操作技能。本书结合劳动部办公自动化操作员级和高级操作员的要求及计算机等级考试的要求进行编写,开设了相应的问题情境,采用任务驱动的方式提出问题、解决问题,强调实践操作,突出应用技能的训练。

本书注重知识的实用性,重视理论概念与操作应用的结合,既可作为大学本科计算机基础教材,又可作为高职高专院校、中职学校各专业的计算机教材和计算机爱好者的参考书。

图书在版编目(CIP)数据

计算机信息技术基础 / 狄寿霞主编. —北京: 化学工业出版社, 2010.6

高职高专“十一五”规划教材

ISBN 978-7-122-08427-9

I. 计… II. 狄… III. 电子计算机-高等学校: 技术学院-教材 IV. TP3

中国版本图书馆 CIP 数据核字 (2010) 第 077042 号

责任编辑: 高 钰

文字编辑: 高 霞

责任校对: 陈 静

装帧设计: 尹琳琳

出版发行: 化学工业出版社 (北京市东城区青年湖南街 13 号 邮政编码 100011)

印 刷: 北京永鑫印刷有限责任公司

装 订: 三河市万龙印装有限公司

787mm×1092mm 1/16 印张 23¼ 字数 606 千字 2010 年 7 月北京第 1 版第 1 次印刷

购书咨询: 010-64518888 (传真: 010-64519686) 售后服务: 010-64518899

网 址: <http://www.cip.com.cn>

凡购买本书,如有缺损质量问题,本社销售中心负责调换。

定 价: 39.00 元

版权所有 违者必究

编写人员名单

主 编 狄寿霞

编写人员 狄寿霞 牛 潇 张甲杰 姜丹丹 王一如
姜学峰 仇玉斌 杨 敏 窦 磊 耿贵芝
张安国 魏 鹏 解英杰

前 言

本书出版的目的是让读者熟练掌握从计算机基础到办公软件 Office 到网络的全面知识,成为真正的办公专家。本书的内容由浅入深、由易到难,帮助读者快速、轻松地学会计算机基础以及 Office 应用程序的基本技巧,然后通过对高级概念、高级功能的深入探讨巩固和拓展计算机信息技术新知识、新技巧。

实行职业资格证书制度是国家提高劳动者素质、增强劳动者就业能力的一项重要举措。《计算机信息技术基础》根据职业院校学生的专业学习与鉴定考核要求,针对参加职业技能鉴定的学生群体,根据参编人员的多年教学、考核经验,以提高学生的职业能力水平为宗旨编写本书。对取得“计算机操作员”、“高级计算机操作员”等计算机职业资格有很高的参考价值。本书讲述了计算机信息技术基础知识、计算机系统应用、Windows XP 系统应用、文字处理软件 Word2003、Word2003 的高级应用、Excel2003 基本操作、Excel2003 数据处理、PowerPoint2003 演示文稿制作、FrontPage2003 网页制作、计算机网络基础知识及 Internet 基础知识等。

本书由张甲杰、解英杰、窦磊负责第一章、第二章、第三章的编写,姜学峰、耿贵芝、张安国负责第四章、第五章的编写,牛潇、魏鹏负责第六章、第七章的编写,狄寿霞、王一如负责第八章的编写,姜丹丹、仇玉斌、杨敏负责第九章、第十章、第十一章的编写。

教材编写是一项系统工程,需要不断地改进。希望广大师生在使用本教材的过程中,积极提出修改意见,以使其不断得到提高和完善,不足之处恳请批评指正。

编 者

2010年4月

目 录

第 1 章 信息技术基础知识	1
1.1 信息与信息技术	1
1.1.1 信息的概念	1
1.1.2 信息的特征	1
1.1.3 信息与数据	2
1.2 计算机内信息的表示与编码的基础知识	2
1.2.1 二进制	2
1.2.2 数在计算机内的表示方法	4
1.2.3 编码	5
1.3 信息安全的基础知识	8
1.3.1 信息安全的基本概念	8
1.3.2 信息安全的基本内容	8
1.3.3 信息安全技术发展趋势	10
1.3.4 信息安全的技术体系结构	10
1.4 计算机病毒	13
1.4.1 计算机病毒及其防治	13
1.4.2 计算机网络病毒和黑客	14
1.4.3 常见几种病毒	14
1.4.4 计算机病毒的防治	15
第 2 章 计算机系统应用	17
2.1 计算机的发展概述	17
2.1.1 计算机分类	18
2.1.2 21 世纪计算机发展趋势	19
2.1.3 计算机应用领域	19
2.2 微型计算机系统的基本组成	21
2.3 计算机硬件组成及其工作原理	22
2.3.1 微型计算机硬件系统及其工作原理	22
2.3.2 微型计算机的硬件组成	22
2.4 计算机软件系统	25
2.4.1 系统软件	26
2.4.2 应用软件	28
第 3 章 Windows XP 系统应用	29
3.1 Windows XP 的桌面	29
3.2 资源管理器系统	31
3.3 更改中文 Windows 的设置	36
3.3.1 设置系统的日期、时间、语言和区域 设置	36
3.3.2 设置显示屏幕	38
3.3.3 辅助功能选项	39
3.3.4 设置电源使用方案	40
3.3.5 设置系统声音方案	40
3.3.6 设置打印机和其他硬件	40
3.4 注册表	43
3.4.1 注册表的定义	43
3.4.2 注册表操作	43
3.4.3 注册表工具	44
第 4 章 文字处理软件 Word 2003	45
4.1 认识 Word 2003	45
4.1.1 启动 Word 2003	45
4.1.2 Word 2003 的窗口组成	45
4.1.3 输入文本	46
4.1.4 设置字体、字形、字号	47
4.1.5 添加图形	53

4.1.6 设置页面属性	54	4.3.3 设置边框和底纹	68
4.2 文档的编辑	57	4.3.4 页眉和页脚	69
4.2.1 文本的选定	57	4.3.5 项目符号和编号	72
4.2.2 文字的编辑修改	58	4.3.6 索引和目录	74
4.2.3 撤销操作和重复操作	59	4.4 表格的创建与处理	76
4.2.4 查找和替换文本	60	4.4.1 表格制作的一般方法	76
4.2.5 拼写和语法检查	63	4.4.2 编辑表格	78
4.2.6 插入符号与特殊字符	63	4.4.3 选择表格中的文本	79
4.2.7 插入日期和时间	64	4.4.4 格式化表格	80
4.2.8 插入文件	64	4.4.5 制作复杂表格	82
4.3 段落格式和文档版式	65	4.4.6 文字和表格的相互转换	85
4.3.1 设置段落格式	65	4.4.7 表格中数据的计算	86
4.3.2 分页和分栏	67		
第 5 章 Word 2003 的高级应用	87		
5.1 图文混排及艺术字的使用	87	5.2.2 利用模板提高创建文档的效率	99
5.1.1 插入剪贴画	87	5.3 域和宏的应用	102
5.1.2 插入图形文件	88	5.3.1 使用域在文档中添加可变内容	102
5.1.3 图片的处理	89	5.3.2 使用宏加速文档编辑进程	106
5.2 样式和模板	94	5.4 邮件合并	111
5.2.1 应用样式统一文档格式	94	5.5 超链接的建立和编辑	114
第 6 章 Excel 2003 基本操作	116		
6.1 认识中文 Excel 2003	116	6.4 设置工作表格式	132
6.1.1 Excel 2003 的主界面	116	6.4.1 设置字符格式	132
6.1.2 Excel 2003 的新功能	117	6.4.2 设置数字格式	132
6.2 工作簿基本操作	118	6.4.3 设置对齐方式	134
6.2.1 新建工作簿	118	6.4.4 添加边框和底纹	136
6.2.2 打开与关闭工作簿	118	6.4.5 自动套用格式	138
6.2.3 保存工作簿	119	6.4.6 条件格式	139
6.2.4 保护与隐藏工作簿	119	6.4.7 设置页面	141
6.2.5 共享工作簿	120	6.5 应用图表	144
6.3 工作表基本操作	122	6.5.1 常用图表类型	144
6.3.1 建立工作表	122	6.5.2 创建图表	147
6.3.2 工作表之间的切换与选定	122	6.5.3 图表的基本操作	148
6.3.3 处理工作簿中的工作表	122	6.5.4 格式化图表	150
6.3.4 保护和隐藏工作表	126	6.5.5 为数据系列添加趋势线	152
6.3.5 单元格以及行、列的操作	126	6.5.6 添加误差线	152
6.3.6 输入数据	128	6.6 使用模板	153
6.3.7 使用批注	131	6.6.1 使用模板建立工作簿文件	153

6.6.2 建立模板	154	6.7.1 操作实例一	155
6.7 操作实例	155	6.7.2 操作实例二	157
第 7 章 Excel 2003 数据处理	160		
7.1 使用公式	160	7.6 合并计算	175
7.1.1 公式的基本概念	160	7.7 数据透视表	177
7.1.2 输入公式	162	7.7.1 创建数据透视表	177
7.1.3 编辑公式	163	7.7.2 使用数据透视表	180
7.2 使用函数	164	7.8 用假设方法求解问题	181
7.2.1 函数的结构	164	7.8.1 单变量求解	181
7.2.2 函数的使用	164	7.8.2 模拟运算表	182
7.2.3 常用函数及应用	165	7.8.3 方案摘要	183
7.2.4 使用数组	168	7.9 设置数据有效性	185
7.3 建立和使用数据清单	170	7.9.1 显示出错提示信息	185
7.4 数据的排序与筛选	171	7.9.2 指定数据的有效范围	187
7.4.1 数据排序	171	7.9.3 提供下拉列表	187
7.4.2 数据筛选	172	7.9.4 删除单元格的有效数据设置	188
7.5 分类汇总	174	7.10 Excel 宏操作	188
7.5.1 创建分类汇总	174	7.10.1 录制宏	189
7.5.2 创建嵌套分类汇总	175	7.10.2 运行宏	189
7.5.3 删除分类汇总	175		
第 8 章 PowerPoint 2003 演示文稿制作	190		
8.1 初识 PowerPoint 2003	190	8.3.3 应用设计模板	205
8.1.1 PowerPoint 2003 启动	190	8.3.4 更改背景	205
8.1.2 PowerPoint 2003 工作界面	192	8.4 设置动画效果及放映方式	206
8.1.3 新建、打开、保存演示文稿	192	8.4.1 设置动画效果	207
8.2 编辑演示文稿	193	8.4.2 设置幻灯片切换方式	210
8.2.1 演示文稿的制作	193	8.4.3 排练计时	210
8.2.2 在大纲视图下创建演示文稿	200	8.4.4 创建超级链接	211
8.2.3 PowerPoint 2003 与其他应用软件共享信息	202	8.4.5 设置放映方式	212
8.3 美化演示文稿外观	203	8.4.6 放映幻灯片	213
8.3.1 母版	203	8.4.7 打印演示文稿	214
8.3.2 应用配色方案	205	8.5 习题	215
第 9 章 FrontPage 2003 网页制作	218		
9.1 一个简单网页的制作	218	9.2 编辑网页	221
9.1.1 初识 FrontPage 2003	218	9.2.1 视图	221
9.1.2 建立一个简单的网页	219	9.2.2 网站的创建	226

9.2.3 网页的基本操作	229	9.5.4 布局表格	270
9.2.4 文本的输入与格式化	233	9.6 组件的使用	276
9.2.5 对象的插入	237	9.6.1 Web 组件的使用	276
9.3 使用图形	240	9.6.2 DHTML 效果的设置	282
9.3.1 插入图片	240	9.6.3 在网页中插入多媒体文件	284
9.3.2 设置图片属性	245	9.6.4 网页的过渡效果	290
9.3.3 编辑图片	247	9.6.5 ActiveX 控件和 Java 小程序	290
9.4 超链接	251	9.7 使用框架	293
9.4.1 创建超链接	251	9.7.1 框架网页的基本操作	293
9.4.2 编辑超链接	256	9.7.2 编辑框架网页	298
9.4.3 创建和使用书签	258	9.7.3 设置框架网页属性	300
9.5 使用表格	259	9.7.4 框架网页中的超链接	302
9.5.1 表格的基本操作	259	9.8 网站的发布与管理	303
9.5.2 表格的属性设置	262	9.8.1 发布网站	303
9.5.3 编辑表格	266	9.8.2 网站的更新与管理	306
第 10 章 计算机网络基础知识	310		
10.1 计算机网络分类	310	10.5 集线器	331
10.2 局域网的拓扑结构与工作模式	315	10.6 路由器	336
10.3 网卡	319	10.7 路由器配置	339
10.4 网络体系结构与协议	325		
第 11 章 Internet 基础知识	345		
11.1 Internet 的产生、发展	345	11.5 电子邮件	354
11.2 Internet 的基本概念	345	11.6 网络信息下载	358
11.3 如何与 Internet 连接	348	11.7 网络论坛	359
11.4 漫游万维网 WWW	349		
参考文献	361		

事物及其运动状态的一种普遍形式。对某事物及其运动,可以用事件发生来描述,事件发生前的平均信息量可用美国科学家申侬的信息熵公式计算。可见,信息不仅表征事物,而且信息量(表征的能力)也是可以度量的。

(7) 可处理性。对信息可进行存储、分析、转换、传递、压缩和再生等处理。

1.1.3 信息与数据

计算机实际上是一台信息(数据)处理机。也就是说,用计算机处理信息时,必须将现实世界中的信息转换为计算机能够识别、存储和处理的形式,即二进制的 0 和 1,以及其他各种经过转换的数据,然后经过加工处理,再将结果(新的信息)提供给外界。例如,数字视频技术就是将通过摄像机等获得的光学运动图像进行处理,转变为数字化图像,然后进行压缩,以便存储在磁盘、光盘等介质上,或通过电缆等其他方式传送出去。

这里所谈的数据,不仅仅是指一些数字,还包括各种数值数据、字符数据(如英文字母、汉字、标点符号和运算符等)、文字数据(字符串数据)、图形图像数据和声音数据等。通常所谓数据处理,指对各种数据进行如采集、存储、传送、转换、分类、排序、计算和图像显示等工作。在计算机内,任何形式的数据都用二进制数表示。

信息与数据既有联系又有区别。数据是计算机化的信息,是信息的载体和表达形式;而信息是数据表达的内涵或解释。数据是具体的物理形式,而信息则是抽象出来的逻辑意义。由于信息与数据的关系如此紧密,因此在很多场合下,人们通常不区分“信息”与“数据”、“信息处理”与“数据处理”两对概念。

世界上浩如烟海的信息都用“0”和“1”来表示,要存储如此多的二进制位数据显然需要存储量特别大的存储介质,同时要提高计算机的信息处理速度,提高信息传输能力(即提高网络的信息通信能力)等。这既是计算机科学技术研究的基本内容,也是计算机特有的功能,更是信息社会的逻辑基础。

1.2 计算机内信息的表示与编码的基础知识

二进制数是计算机表示信息的基础。本节首先引入二进制数的概念,再介绍数值型数据在计算机内的表示方式和字符(包括英文字符和汉字)在计算机内的表示方式。

1.2.1 二进制

人类习惯使用十进制表示数。十进制有 10 个不同的数字(表记符号),分别是 0, 1, 2, ..., 8, 9。两个十进制数进行运算时遵循“逢 10 进 1”的规则。在进位计数制中所用不同数字的个数称为该进位计数制的基数,那么十进制的基数是 10。

在人类社会的发展过程中,人类还创造了各种不同的进位计数制。如一天有 24 小时,即二十四进制,逢 24 进 1;一小时为 60 分钟,即六十进制,逢 60 进 1;一周有 7 天,即七进制,逢 7 进 1。在上述进位计数制中,有的有自己的表记符号,有的借用其他进位计数制的表记符号。用什么表记符号并不重要,只要使用方便即可。不同进位计数制之间的区别在于它们的基数和表记符号不同,进位规则不同。

一、二进制数

二进制数只有 0 和 1 两个记数符,其进位的基数是 2,遵循“逢 2 进 1”的进位规则。在计算

机中采用二进制数表示数据的原因如下。

① 计算机科学理论已经证明：计算机中使用 e 进制 ($e \approx 2.718$) 最合理，取整数，可以使用二进制。

② 二进制数值可以采用的电子元器件（开关特性）很容易实现计算机内数值 1 和 0 的存储、计算和表达。如果使用三进制则需要 3 个电压量来表达其 3 个表记符号。显然，其可靠性比使用二进制更容易受电压波动的影响。

③ 运算简单。 $0+0=0$, $0+1=1$, $1+0=1$, $1+1=10$ 。数值量与逻辑量共存，便于用逻辑运算器件实现算术运算。

二、八进制数与十六进制数

人类使用二进制表达一个比较大的数值时，书写长，看起来不直观，很容易出错。由于八进制和十六进制与二进制有运算上的完全对应关系，所以常常采用八进制和十六进制记数法（便于人类阅读）。

八进制的基数为 8，有 0~7 共 8 个表记符号，运算遵循“逢 8 进 1”的规则。一位八进制数正好用 3 位二进制数表达，它们的对应关系如表 1-1 所示。

表 1-1 八进制数与二进制数的对应关系

八进制	0	1	2	3	4	5	6	7
二进制	000	001	010	011	100	101	110	111

八进制数与二进制数的转换很容易，按照表 1-1，每一位八进制数写成对应的 3 位二进制数，即完成八进制数到二进制数的转换；从低位到高位每 3 位二进制数写成对应的 1 位八进制数，即完成二进制数到八进制数的转换。如 $(157)_8 = (001\ 101\ 111)_2$ 。

十六进制的基数为 16，共有 0、1、2、3、4、5、6、7、8、9、A、B、C、D、E、F 16 个表记符号，运算遵循“逢 16 进 1”的规则。一位十六进制数正好用 4 位二进制数表达，它们的对应关系如表 1-2 所示。

表 1-2 十六进制数与二进制数的对应关系

十六进制	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
二进制	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111

十六进制数与二进制数的转换同样很容易，按照表 1-2，每 1 位十六进制数写成对应的 4 位二进制数，即完成十六进制数到二进制数的转换；从低位到高位每 4 位二进制数写成对应的 1 位十六进制数，即完成二进制数到十六进制数的转换。如 $(FD57)_{16} = (1111\ 1101\ 0101\ 0111)_2$ 。

三、二进制数与十进制数的转换

十进制毕竟是人们最熟悉的数制。在计算机的输入输出中采用十进制，即人类使用十进制数输入数据，计算机输出数据时也使用十进制数，而计算机内部存储数据使用二进制。数据的输入、输出过程中的十进制数到二进制数，二进制数到十进制数的转换由机器自动完成。

1. 二进制数向十进制数转换

一个二进制数按其权位（权位用十进制表示）展开求和，即可得到相应的十进制数。如

$$\begin{aligned}(110.101)_2 &= (1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3})_{10} \\ &= (4 + 2 + 0.5 + 0.125)_{10} = (6.625)_{10}\end{aligned}$$

2. 十进制数向二进制数转换

整数部分的转换采用“除 2 取余”法。十进制数整数部分除以 2，余数作为相应二进制数整

数部分的最低位；用上步的商再除以 2，余数作为二进制数的次低位……一直除到商为 0，最后一步的余数作为二进制数的最高位。

例如，将十进制数 11 转换为二进制数的过程如下：

除法	商	余数
$11 \div 2$	5	1
$5 \div 2$	2	1
$2 \div 2$	1	0
$1 \div 2$	0	1

故 $(11)_{10} = (1011)_2$

小数部分的转换采用“乘 2 取整”法。十进制小数部分乘 2，积的整数部分为相应二进制数小数部分的最高位；用上一步积的小数部分再乘 2，同样取积的整数部分作为相应二进制数小数部分的次高位……一直乘到积的小数部分为 0 或达到所要求的精度为止。

例如，将十进制数 0.625 转换为二进制数的过程如下。

乘法	积的整数部分	积的小数部分
0.625×2	1	0.25
0.25×2	0	0.5
0.5×2	1	0

故 $(0.625)_{10} = (0.101)_2$

1.2.2 数在计算机内的表示方法

一、信息度量单位

无论是数值型数据，还是字符（包括英文字符、汉字或其他符号）都是存储在一个称为“字节”的单元中。

一个二进制位称为“位”（bit），8 个二进制位组成一个字节（byte），更大的度量单位是 KB、MB、GB、TB。1KB=1024 字节，1MB=1024KB，1GB=1024MB，1TB=1024GB。计算机信息处理的最小单位是位，而计算机寻址的单位是字节。

二、原码

二进制数在计算机内可以使用原码、反码和补码表示。

以下假设数据使用 8 位二进制表示，用其中一位表示数的符号（0 表示正数，1 表示负数），其余 7 位表示数值。

例如，设十进制数 $A=+27$ ，其原码简记为 $[A]_{\text{原}}=00011011$ ；设十进制数 $B=-37$ ，其原码简记为 $[B]_{\text{原}}=10100101$ 。

在原码表示法中，数值 0 有两种表示方法：正 0 和负 0，简记为 $[+0]=00000000$ ， $[-0]=10000000$ 。

三、反码

反码又称“对 1 的补码”。符号位与原码约定相同。正数的反码与原码相同，负数的反码是在原码的基础上按位取反。例如 $[+27]_{\text{反}}=[+27]_{\text{原}}=00011011$ ，而 $[-37]_{\text{反}}=11011010$ 。

在反码表示法中，数值 0 也有两种表示方法：正 0 和负 0，简记为 $[+0]=00000000$ ， $[-0]=11111111$ 。

反码表示法的优点是统一了加减法运算，只需要计算加法；反码表示法的缺点是运算时会引起循环进位，这既占用机器计算时间，又给机器设计带来麻烦。因此，人们又寻求另一种表示方法，即补码。

四、补码

补码又称“对 2 的补码”。符号位与原码约定相同。正数的补码与原码相同，负数的补码是在原码的基础上按位取反后，最后位加 1。例如 $[+27]_{补}=[+27]_{原}=00011011$ ， $[-37]_{补}=11011011$ 。

在补码表示法中，0 有唯一的表示方法： $[+0]_{补}=[-0]_{补}=00000000$ 。

引入补码概念后，加法、减法都可用加法实现。因此，现代计算机多采用补码运算。

在讨论原码时已经假设使用 8 位二进制表示数，即假设计算机的字长是 8 位。因为要用一位表示符号，所以只有 7 位用来表示数据，那么在这台计算机上数据的表示范围是 $+127\sim-128$ 。当实际数据不在这个范围内时就会出错，这就是溢出。当然，现代计算机的字长大于 8 位（16 位、32 位或 64 位），但终究有一个范围。对于溢出问题，计算机要作相应处理。

前面讨论的三种表示方法只能表示单纯整数或小数，称其为数的定点表示法。在计算机中，参与运算的数一般是实数，既有整数部分又有小数部分。为了表示实数，使用数的浮点表示方法。任何一个实数可以表示成 $A = 2^i \times S$ ，其中 S 是实数 A 的尾数，S 的符号可用 CS 表示，0 表示正数，1 表示负数；i 是用二进制表示的阶码，i 的符号可用 Ci 表示，0 表示正数，1 表示负数。例如，实数 110.101 可表示成 $2^{11} \times 0.110101$ 。浮点表示方法的格式如下。

Ci	i	CS	S
----	---	----	---

Ci、CS 各只用一位，i 的位数决定实数的表示范围，S 的位数决定实数的精度。

计算机内的数值运算以加法为基础，其他运算都可以变成加法实现。然而，现代的微处理器内已集成了浮点运算部件，其中包括了乘法器等，以提高运算速度。

1.2.3 编码

编码即字符在计算机内的表示方法，是用二进制形式表示的，其中包括字符与汉字。

一、ASCII 编码

当今的计算机普遍采用 ASCII 编码，即美国标准信息交换码（American Standard Code for Information Interchange）。ASCII 编码集中的字符用 8 位二进制数表示，但只用低 7 位，共可表示 128 个字符。编码为 0~127 称为 ASCII 码基本集；128~255 的编码称为 ASCII 码扩展集，留作他用，如字符 A 的编码表示如下：

b7	b6	b5	b4	b3	b2	b1	b0
0	1	0	0	0	0	0	1

字符的 ASCII 编码如表 1-3 所示。其中，编码 0~31 表示控制字符。控制字符对照表如表 1-4 所示。

表 1-3 ASCII 字符编码表

b6~b4 b3~b0	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	"	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t

续表

b6~b4 b3~b0	000	001	010	011	100	101	110	111
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	'	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	R	[	k	{
1100	FF	FS	,	<	L	\	l	
1101	CR	GS	-	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	Del

表 1-4 控制字符对照表

ASCII 码	控制字符	主要作用	输入字符	ASCII 码	控制字符	主要作用	输入字符
00000	NUL	空白	Ctrl+@	10000	DLE	数据链路换码	Ctrl+P
00001	SOH	序始	Ctrl+A	10001	DC1		Ctrl+Q
00010	STX	文始	Ctrl+B	10010	DC2	设备控制 1	Ctrl+R
00011	ETX	文终	Ctrl+C	10011	DC3	设备控制 2	Ctrl+S
00100	EOT	送毕	Ctrl+D	10100	DC4	设备控制 3	Ctrl+T
00101	ENQ	询问	Ctrl+E	10101	NAK	设备控制 4	Ctrl+U
00110	ACK	承认	Ctrl+F	10110	SYN	否认	Ctrl+V
00111	BEL	警铃	Ctrl+G	10111	ETB	同步空转	Ctrl+W
01000	BS	退格	Ctrl+H	11000	CAN	发送块终	Ctrl+X
01001	HT	横表	Ctrl+I	11001	EM	取消	Ctrl+Y
01010	LF	换行	Ctrl+J	11010	SUB	载体终	Ctrl+Z
01011	VT	纵表	Ctrl+K	11011	ESC	取代	Esc
01100	FF	换页	Ctrl+L	11100	FS	扩展	Ctrl+\
01101	CR	回车	Ctrl+M	11101	GS	文件间隔	Ctrl+]
01110	SO	移出	Ctrl+N	11110	RS	组间隔	Ctrl+6
01111	SI	移入	Ctrl+O	11111	US	记录间隔	Ctrl+_

二、汉字编码

汉字编码包括汉字内码、汉字输入编码（外码）和汉字输出编码（字模）三个主要内容。

1. 汉字内码

汉字内码是汉字在计算机内的存储表示。汉字数量庞大，只能选取部分汉字用于计算机汉字信息处理。国家标准 GB2312—80 定义了国标码（区位码），它可容纳 7445 个汉字和符号，分为 94 个区，区号为 1~94，每区容纳 94 个汉字或符号（汉字在区内的排列位置称为位），位号为 1~94，可对应 94 个汉字或 94 个符号。其中 682 个符号在第 1 区~第 9 区，第 10 区~第 15 区未用；6763 个汉字分为两级，一级字库为第 16 区~第 55 区，包括 3755 个常用汉字，按音序排列；二级字库为第 56 区~第 87 区，包括 3008 个次常用汉字，按笔画数排列。

注意：国标码是4位十六进制数。按上面的定义，“啊”字在第16区第01位上，因此十进制数1601是“啊”字的区位码。国标码是将某字的区位码之区、位分别转换成十六进制数，分别加上十六进制数20。故“啊”字的国标码为3021H。每一个汉字或符号在区位码代码集或国标码代码集中都有唯一的代码。

汉字内码可使用ASCII码扩展集的代码(128~255)，但因汉字数量众多，用一个字节无法表示，所以用两个连续的字节来存放一个汉字的内码。汉字字符(用汉字内码表示)必须与英文字符(用ASCII码表示)能相互区别，以免造成混淆，这也就是所谓的中西文兼容问题。英文字符的机内代码是7位ASCII码，最高位为0(即b7=0)，汉字内码中两个字节的最高位均为1。将某汉字的国标码分别加上80H，作为汉字内码。以汉字“啊”为例，国标码为3021H，汉字内码为B0A1H。

2. 汉字输入编码

汉字的输入不能像英文字符那样，一键对应一个字符，只能多键输入一个汉字，这里的多键就是一个汉字的输入编码。可见，汉字输入编码是将汉字输入到计算机内(变成汉字内码)的编码，所以有人称之为外码。目前有几百种汉字输入编码，下面列举具有代表性的几种输入编码。

① 区位码。一个汉字的区位码是一个4位十进制数，如汉字“啊”的区位码为1601。用区位码输入汉字实际上是将区位码转换成汉字内码。这是一种无重码(重码是一个输入编码对应多个汉字)输入方法，即一个输入编码对应一个汉字(内码)。

② 拼音输入编码。用汉字的拼音符号作为输入编码，如汉字“湖”的拼音是hu，这就是其拼音输入编码。用拼音输入方法输入汉字就是把像hu这样的输入编码变成所表达的汉字的内码。显然，拼音输入方法是一种有重码的输入方法。

③ 字形输入编码。是一种以汉字的偏旁部首作基本键位的输入编码，即把键盘上的某一键位当作偏旁部首(当然，某一键位可能代表多个偏旁部首)，多个键位(对应多个偏旁部首)的组合就是汉字的字形输入编码。五笔字型输入编码属于这一类输入编码，它是目前用得相当广泛的输入编码。例如在五笔字型输入方案中，“湖”字的三个偏旁部首“氵”、“古”和“月”分别安排在键盘的i、d和e三个键位上，那么ide字串就是“湖”字的五笔字型输入编码。

一般来说，字形输入编码输入方法的重码少于拼音输入方法，输入速度快；而拼音输入方法易学，输入速度相对较慢。

3. 汉字输出编码

把某一个汉字当作一幅平面图画。分别从纵、横两个方向等距离在画上作 $(N-1)$ 条直线，这样就把该幅画分成 $N \times N$ 个小方块(点矩阵)，我们会发现有的小方块内有汉字的笔画，有的则没有。把有笔画的小方块记上1，没有笔画的小方块记上0，就得到了一幅由1组成的该汉字的轮廓画，这就是一幅数值化了的图形，如图1-1给出了一个 16×16 点阵“大”字数值化图形，当然此图中的汉字比较粗糙，按照此构思可以为每个汉字构造这样的图形。

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0
0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0
0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

图1-1 16×16 点阵模型

对于任何一个汉字,可以把其数值化了的平面 0、1 图形按先行后列(也可先列后行)的顺序编成二进制代码串存入计算机中(每一个 0 或 1 用一个二进制位存储)。把这样数值化的图形叫做某汉字的 $N \times N$ 点阵字模。对于一个 16×16 点阵的汉字字模,需要用 16×16 个二进制位来存储,即需要 32 个字节。图 1-1 中“大”字的 32 个字节数据(按先行后列顺序,十六进制)是: 00, 00, 03, 80, 03, 04, 7F, FE, 7F, FE, 27, 00, 06, 80, 0C, 40, 0C, 60, 18, 30, 18, 30, 30, 18, 30, 18, 20, 0C, 40, 0E, 00, 00。

所有汉字字模的集合称为字库。对于用 16×16 点阵字模组成的字库需要大约 220KB 存储容量。汉字字模在字库中的位置按汉字内码升序存于字库中。

1.3 信息安全的基础知识

一般意义上,信息(Information)是指事物运动的状态和方式,是事物的一种属性,在引入必要的约束条件后可以形成特定的概念体系。对现代社会来说,信息也是一种资产,包括计算机和网络中的数据,还包括专利、标准、商业机密、文件、图纸、管理规章、关键人员等,就像其他重要的商业资产那样,信息资产具有重要的价值,因而需要进行妥善保护。

1.3.1 信息安全的基本概念

美国联邦政府给出了信息安全的标准定义:“保护信息系统免受意外或故意的非授权泄漏、传递、修改或破坏”。另外,我国信息安全国家重点实验室对信息安全的定义是:“信息安全涉及信息的保密性(Confidentiality)、可用性(Availability)、完整性(Integrity)和可控性(Controllability)。保密性就是保证信息不泄漏给未经授权的人;可用性就是保证信息以及信息系统确实为授权使用者所用;完整性就是抵抗对手的主动攻击,防止信息被篡改;可控性就是对信息以及信息系统实施安全监控。综合起来说,就是要保障电子信息的有效性。”国际标准化组织(ISO)定义信息安全为:“为数据处理系统建立和采取的技术和管理的安全保护,保护计算机硬件、软件和数据不因偶然和恶意的原因而遭到破坏、更改和显露。”随着计算机应用范围的逐渐扩大以及信息内涵的不断丰富,信息安全涉及的领域内涵也越来越广。信息安全不仅是保证信息的机密性、完整性、可用性、可控性和可靠性,并且从主机的安全技术发展到网络体系结构的安全,从单一层次的安全发展到多层次的立体安全。目前,信息安全涉及的领域还包括黑客的攻防、网络安全管理、网络安全评估、网络犯罪取证等方面。因此在不会产生歧义时,常将计算机网络信息系统安全简称为信息安全。由于计算机网络具有联结形式多样性、终端发布不均匀性和网络开放性、互联性等特征,使得网络易受到黑客、恶意软件和其他不轨行为的攻击,所以网络信息的安全和保密性就是一个至关重要的问题。无论是在单机系统、局域网还是在广域网系统中,都存在着自然环境和人为等诸多因素的脆弱性和潜在威胁。因而计算机网络系统的安全措施应该是可以全方位地针对各种不同的威胁和脆弱性,才能确保网络信息的保密性、可用性、完整性和可控性。一切影响计算机网络安全的因素和保障计算机网络安全措施都是计算机网络安全的研究内容。在这里,我们可以这样来定义信息安全:信息安全是指信息在产生、传输、处理和储存过程中不被泄漏或破坏,确保信息的可用性、保密性、完整性和不可否认性,并保证信息系统的可靠性和可控性。

1.3.2 信息安全的基本内容

一、信息安全的属性

信息安全的基本属性有信息的完整性、可用性、机密性、可控性、可靠性和不可否认性。

(1) 完整性。完整性是指信息在存储、传输和提取的过程中保持不被修改延迟、不乱序和不丢失的特性。一般通过访问控制阻止篡改行为,通过信息摘要算法来检验信息是否被篡改。完整性是数据未经授权不能进行改变的特性,其目的是保证信息系统上的数据处于一种完整和未损的状态。

(2) 可用性。信息可用性指的是信息可被合法用户访问并能按要求顺序使用的特性,即在需要就可取用所需的信息。可用性是信息资源服务功能和性能可靠性的度量,是对信息系总体可靠性的要求。目前要保证系统和网络能提供正常的服务,除了备份和冗余配置之外,没有特别有效的方法。

(3) 机密性。信息机密性又称为信息保密性,是指信息不泄漏给非授权的个人和实体,或供其使用的特性。信息机密性针对信息被允许访问对象的多少而不同。所有人员都可以访问的信息为公用信息,需要限制访问的信息一般为敏感信息或秘密,秘密可以根据信息的重要性或保密要求分为不同的密级,如国家根据秘密泄露对国家经济、安全利益产生的影响(后果)不同,将国家秘密分为A(秘密级)、B(机密级)和C(绝密级)三个等级。秘密是不能泄漏给非授权用户、不被非法利用的,非授权用户就算得到信息也无法知晓信息的内容。机密性通常通过访问控制阻止非授权用户获得机密信息,通过加密技术阻止非授权用户获知信息内容。

(4) 可控性。信息可控性是指可以控制授权范围内的信息流向以及行为方式,对信息的传播及内容具有控制能力。为保证可控性,通常通过握手协议和认证对用户进行身份鉴别,通过访问控制列表等方法来控制用户的访问方式,通过日志记录对用户的所有活动进行监控、查询和审计。

(5) 可靠性。可靠性是指信息以用户认可的质量连续服务于用户的特性(包括信息的迅速、准确和连续地转移等),但也有人认为可靠性就是人们对信息系统而不是对信息本身的要求。

(6) 不可否认性。不可否认性是指能保证用户无法在事后否认曾对信息进行的生成、签发、接受等行为,是针对通讯各方面信息真实同一性的安全要求。一般用数字签名和公证机制来保证不可否认性。

二、信息安全的特征

信息安全具有整体的、动态的、无边界和发展的特征,是一种非传统安全。信息安全涉及多个领域,是一个系统工程,需要全社会的共同努力和承担责任及义务;信息安全不是静态的,它是相对和动态的,经历了从最初纯粹的物理安全问题到今天随着信息技术的发展和普及,以及产业基础、用户认识、投入产出而出现的动态的全方位的安全问题;信息安全已经是全球性的而非某个国家或地区特有的问题,尤其是网络高度的互动性、渗透性使得信息安全问题变得越来越难以控制,不可避免地影响到我们生活的方方面面。信息安全是过程的安全,它不是固定不变的,而是贯穿于整个信息技术的发展过程中,应在系统建设过程中同步考虑。

互联网的全球性、快捷性、共享性、全天候性决定了信息安全问题的新特征。信息基础设施本身的脆弱性和攻击手段的不断更新,使信息安全领域易攻难守。网上攻击无论距离还是速度都突破了传统安全的限制,具有多维、多点、多次实施隐蔽打击的能力。由于网络覆盖全球,因而助长了犯罪分子的破坏能力和有恃无恐的犯罪心理,给世界带来了更多的不稳定因素。各国的民族文化和道德价值观面临前所未有的冲击和颠覆,为此付出的经济成本和时间精力难以计算。信息安全问题日益严重必将给人类发展、国家管理和社会稳定带来巨大的危害。

信息安全属于非传统安全,非传统安全主要是相对于传统安全而言的。传统安全是指以军事安全为核心的安全;而将军事以外的对主权国家及人类整体生存与发展构成威胁的因素称为非传统安全。除军事以外的非传统安全问题主要包括:经济安全、金融安全、环境安全、信息安全、