



经全国中小学教材审定委员会 2004 年初审通过

普通高中课程标准实验教科书

(信息技术·必修)

信息技术基础

主编 陶增乐

浙江教育出版社

PUTONGGAOZHONG KECHENGBIAOZHUN SHIYANJI AOKESHU
XINXI JISHUJICHU

普通高中课程标准实验教科书

信息技术基础

陶增乐 主编

-
- 出版 浙江教育出版社
(杭州市天目山路 40 号 邮编 310013)
- 重印 浙江省出版总社
- 发行 浙江省新华书店集团有限公司
- 责任校对 雷 坚
- 责任印务 温劲风
- 印 刷 浙江新华数码印务有限公司
-

- 开 本 890 × 1240 1/16
- 印 张 9.25
- 字 数 300 000
- 版 次 2004 年 4 月第 1 版
- 印 次 2011 年 5 月第 37 次
- 书 号 ISBN 978-7-5338-5253-5
- 定 价 10.55 元
-

批准文号: 浙价教材批[2006] 2 号 举报电话: 12358

联系电话: 0571-85170300-80928

e-mail: zjjy@zjcb.com

网址: www.zjeph.com

如发现印、装质量问题,请与本公司联系。电话: 0571-85155604

编者的话

提高信息素养,为国家培养出新一代掌握基本信息技术的合格公民,是高中阶段文化教育的基本任务之一。

信息技术是不断创新发展变化的,正是这个原因,它是那么吸引人,那么有活力,它创造了许多对社会发展,对个人生活极有意义的东西。网络通信技术的发展,多媒体技术的进步,数据管理能力的提高,新处理系统、新算法的不断出现,使信息获取、加工和传播的能力以前所未有的速度提高,不断地改变着人们的生活方式。使人们在富于变化的信息社会里,生活得更幸福自如。

根据教育部制订的《信息技术课程标准》,对高中学生来说,要掌握的信息技术内容主要是指信息处理技术,即对信息的获取、加工、管理、表达和交流方面的技术。整个课程由:“信息技术基础”、“算法与程序设计”、“多媒体技术应用”、“网络技术应用”、“数据管理技术”、“人工智能初步”等六个模块组成。课程分为必修和选修两个层次。“信息技术基础”为必修模块,其余五个模块为选修,每位同学应至少完成一个选修模块的学习。

作为一本信息技术的基础性教科书,我们要展示的是一幅关于最基本的信息技术当前情况的鸟瞰式的写照以及它的历史背景的介绍。同学们应该得到的是一本能确切解释基本原理,概念通俗易懂,与实践活动紧密结合的教科书,而不应该是一种缩了水的应用软件操作使用手册——只包括那些被认为高中生可以接受的某些常用软件的操作使用技巧和某些专题的简化介绍。同学们掌握了这些基础知识和基本技能之后,能够伴随信息技术的进步而健康成长。我们鼓励大家通过探索超越本教科书的知识范围,使自己在成长的过程中有扎实的基础和较高的起点。

经验表明,只有积极动手通过实践才能理解和掌握信息技术的奥秘。对一些重要内容,采用“讨论交流”、“实践体验”和“综合探究”等活动方式学习是值得提倡的。同学们应从中体会到这是一门实践课程,而不是理论课程。

要让学生通过解决问题的活动潜移默化地学会解决问题的本领,而不是依靠“讲授”问题求解来获得这种本领,课堂教学中的问题设计的作用是相当关键的。

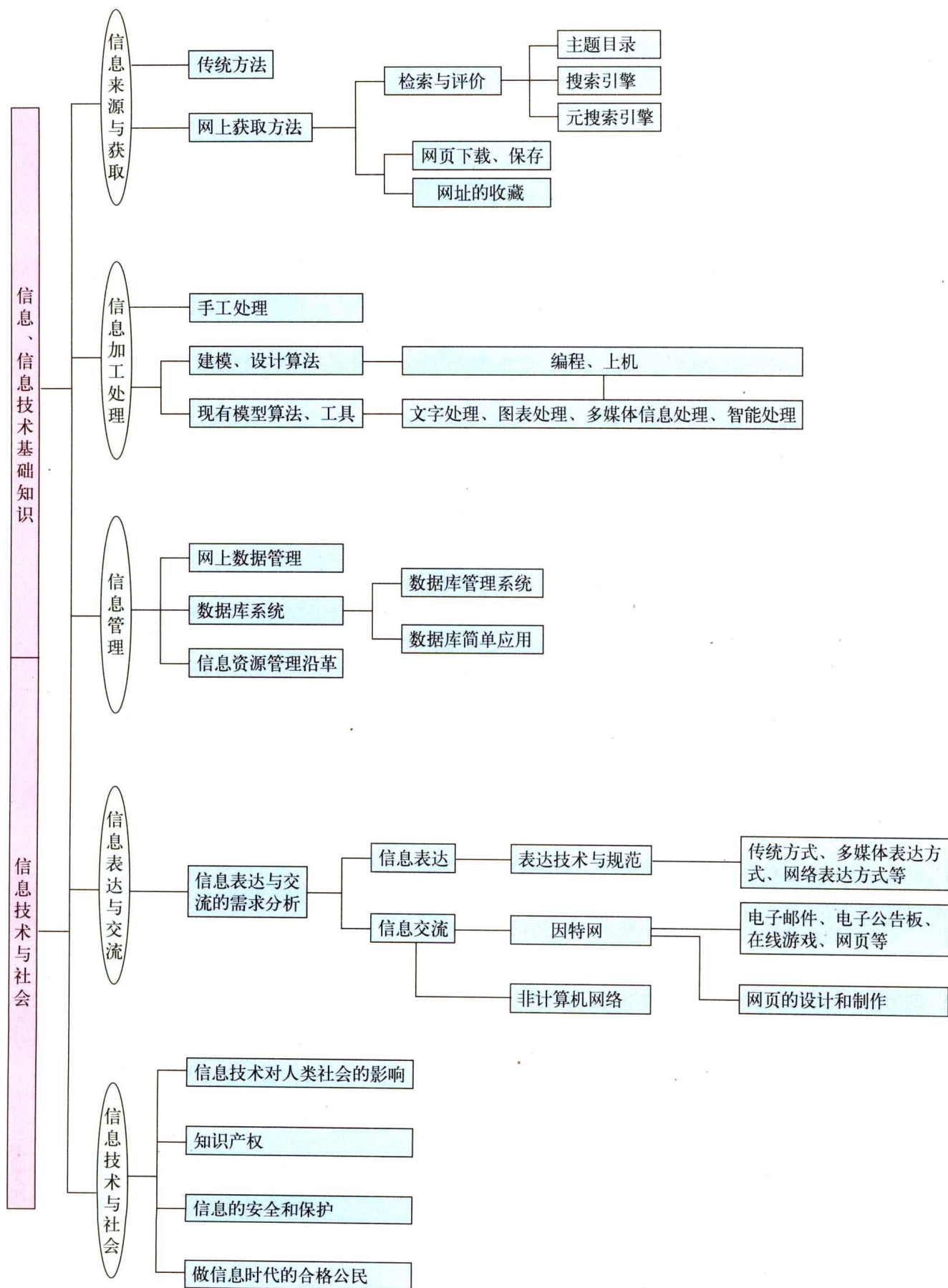
实践证明,过程性评价和结果性考查的有机结合有助于提高学习的主动性和积极性。虽然书中对每次实践探究活动都给出了评价要求,并在书末附录中提供了含有全部活动的学生学业成长记录表,但是我们建议带“*”号的实践活动的评价应尽可能地进行。

我们感谢用本书进行教学的教师,请你们把这本教科书作为一种重要的教学资源,而不要照本宣科地予以使用。我国幅员辽阔,东西南北、城市农村情况千差万别,本书所包含的全部内容比一个学期所能教的多,请你们根据当地,尤其是本校的实际情况,节选内容,创造性地组织教学。也可将某些题目的内容作为阅读材料,鼓励学生自主学习课堂上没有讲过的内容。

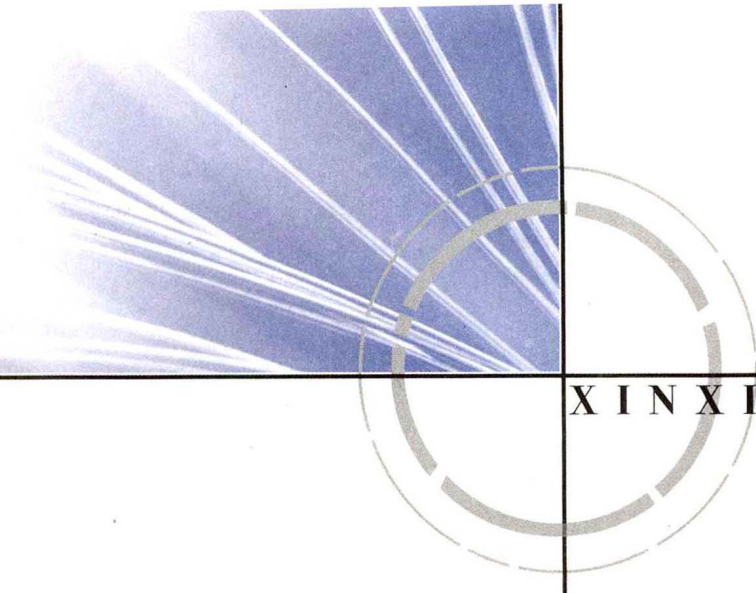
本教科书的内容架构图展示了“基础”模块的内容框架及其相互关系。书中的“知识链接”和“阅读材料”向大家提供了一些相应问题的补充知识。每章末所附的学习线路图,是作为该章的知识技能架构的概括,有助于学生对这一章的学习的总结回顾,经过相应的补充修改,即可形成自己头脑中关于本章内容的“大局观”。

经过实验区的试用,我们倾听了有关教研机构的领导和广大教师的意见,对原版内容中的疏误欠妥之处作了补正,在此特向指出问题、提供改正意见的老师表示一个同行的衷心谢意。

2007年5月



“信息技术基础”模块内容架构图



目 录

XINXI JISHU JICHU

第一章 信息与信息技术

1.1 信息及其特征	2
1.2 信息的编码	4
1.3 信息技术	12
本章小结	19

第二章 信息的来源与获取

2.1 信息获取的方法	21
2.2 因特网上信息的浏览与获取	26
2.3 网上资源检索	33
2.4 因特网信息资源评价	38
本章小结	43

第三章 信息的加工

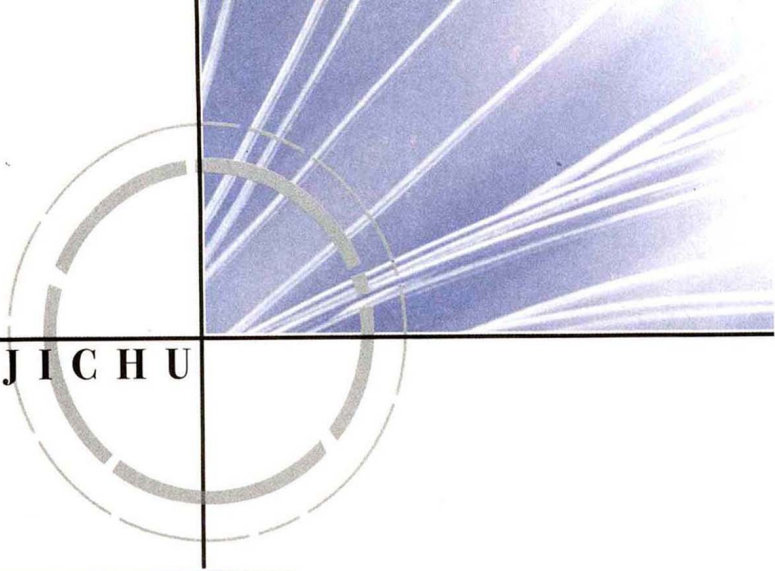
3.1 现代信息处理工具——计算机	45
3.2 字处理和表处理	48
3.3 多媒体信息处理	60
3.4 算法及其实现	66
3.5 智能处理	71
本章小结	76

第四章 信息的管理

4.1 信息资源管理及其沿革	78
4.2 数据库系统	84
本章小结	90

第五章 信息的表达与交流

5.1 信息表达	92
5.2 电子邮件	95
5.3 电子公告板、在线游戏	100
本章小结	106



XINXIJISHU JICHU

第六章 网页的设计与制作

6.1 网站和网页	108
6.2 网页制作	111
本章小结	124

第七章 信息技术与社会

7.1 信息技术对人类社会的影响	126
7.2 知识产权	129
7.3 信息的安全和保护	130
7.4 做信息时代的合格公民	136
本章小结	139

附录 学生学业成长记录表	140
参考书目	143

第一章 信息与信息技术



学习任务

1

能够描述信息的基本特征

2

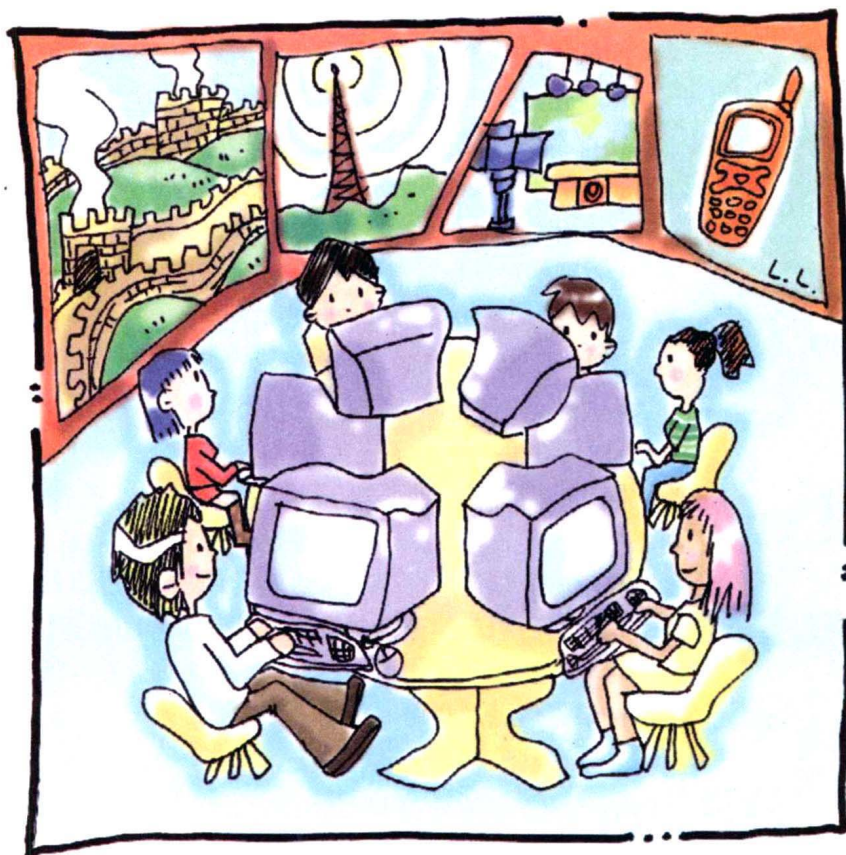
列举信息技术的应用实例

3

了解信息技术的历史和发展趋势

自古以来，人类的生存和发展与信息有着不解之缘，人们的生活一时一刻都离不开接收信息、传递信息、处理信息和利用信息。因此，人们对信息的采集、处理、传输、存储、表达和应用等方面的问题进行过许多卓有成效的研究。语言和文字是人类社会表达和传递信息的最基本的工具，造纸术和印刷术的发明使信息表示和存储方式产生了一次重大变化，文字成为记录、储存和传递信息的有效手段。电报、电话、广播和电视的发明，帮助人们拉近了信息传递的时空距离。

但是，把信息作为一门科学来研究，则是20世纪中后期的事。1948年，美国科学家申农(C.E.Shannon)发表了著名的论文《通信的数学理论》，奠定了信息论的理论基础。20世纪后半期，计算机技术、微电子技术、传感技术、激光技术、卫星通信和移动通信技术、广播电视技术、多媒体技术、新材料和新能源技术以及航空航天技术的飞速发展和应用，尤其是以计算机为主体的互联网技术的发展，它们相互结合，相互促进，将人类社会推入高度信息化时代，使信息的采集、处理、传输、存储、表达和应用的方式达到了前所未有的水平。学习信息技术，感受信息文化，增强信息意识，内化信息伦理，对信息时代的每个人都具有重要的意义。





1.1 信息及其特征

物质、能源和信息(information)是人类社会的三大要素。在瞬息万变的世界,如果没有信息的交换和处理,就不会有今天的人类文明。

1.1.1 信息无处不在

清晨,闹钟的铃声告诉你该起床的信息,书刊、报纸、广播、电视运载着大量的社会信息;刮风下雨、春华秋实表达了天气变化、季节更替的自然信息;喜怒哀乐则表现出人的情感活动信息。读书、看报、听课、做实验,都是为了获得有用的信息。聪明有为的单位领导会从各种数据报表、行情变化、员工报告及相应政策法规中获取有用的信息,做出正确的决策。医生从病人病历、各种检查结果和症状观察中获得疾病诊断所需的信息。科研人员从国内外科技文献、学术交流、科学实验中获得课题开发和研究所需要的信息。

每个人时时刻刻都在与周围交换信息,每个单位经常处于单位内外的信息交换过程之中。一个国家要发展就不能闭关自守,应该时刻掌握国内外政治、经济、军事和科学文化等方面的信息。

总之,我们一时一刻也离不开信息,信息无处不在。我们就生活在一个千变万化的信息世界里。

信息已经成为当今社会一个最基本、最重要的概念。尽管人们时时处处接触信息,用到信息,但是究竟什么是信息,却很难有一个十分确切而又简明的说法。人们从不同的角度给信息提出了各种定义和说法,但是至今还没有一个公认的定义。

信息是指数据(data)、信号、消息中所包含的意义。电视上有重大新闻的消息,报纸上有足球比赛结果的消息,这些消息都是信息。

信息是事物的运动状态和关于事物运动状态的描述。世界上的万事万物都在不停地运动、变化,万事万物里都有信息。信息既是世界上各种事物的特征和事物运动变化的反映,又是事物之间相互作用和联系的表示。

用符号传送的报道,其内容是接收者预先不知道的东西都被认为是信息。即信息是指对消息接受者来说是预先不知道的东西,具有“不确定性”。当获得信息之后,这种“不确定性”就可以减少或消除。信息量的大小可用消除“不确定性”的多少来表示。如广播天气预报时,收听者预先不知道明天是阴、雨还是晴,本次天气预报对急需了解明天天气状况的收听者来说就是很要紧的信息。天气预报越详细、越出乎收听者的预料,则信息量越大。假如广播时有外界干扰,收听受到影响,则广播的信息也受到损失。

人类自古以来就不断地通过感官摄取信息,通过头脑处理信息,通过科学研究和创造性思维产生新的信息,通过语言、文字、图画等交流信息,并根据所积累的信息去进一步认识世界和改造世界。因此,信息既是主观与客观相互联系作用的媒介,又是物质世界与精神世界相互作用、联系的桥梁。信息是我们人类的宝贵财富。



1.1.2 信息的特征

信息具有哪些特征呢?

第一,信息的表示、传播、储存必须依附于某种载体,载体就是承载信息的事物。语言、文字、声音、图像以及纸张、胶片、磁带(magnetic tape)、磁盘(magnetic disk)、光盘等,甚至人的大脑,都是信息的载体。例如,从电台听到的气象预报,信息是通过语言、声音和电磁波等信息载体传递的,从报纸上读到的新闻,信息是通过文字和纸张等信息载体来传递的。不存在没有载体的信息。

第二,信息是可以加工和处理的。经过加工、处理,特别是经过人的分析、综合和提炼,使信息具有更高的使用价值。例如,国家统计局部门每年都要统计反映国家经济特征的一些数据。这就需要在统计过程中对大量的信息进行处理,得到的结果既可以概括地表示当年国家的经济运行状况,也可以利用这些统计分析结果对下一年进行预测,制定计划和做出决策。另外,人们还可以利用各种信息技术,把信息从一种形态转换为另一种形态。例如,打电话时,要把发话人讲话的声音转换成电信号,通过电话线路传送,在受话端再把电信号转换成声音的形态。

第三,信息可以脱离它所反映的事物被存储、保存和传播。信息生成后,可以用各种载体来存储、保存和传播。这就是为什么我们通过书籍、录音、影像可以看见或听到以前发生的事物,而不需要把以前的真实事物保存下来的原因。正是信息的此项特征,可以使我们了解过去,甚至是自己出生之前发生的事情,也可以让我们了解距离遥远,甚至无法到达的地方的事情。

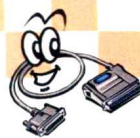
第四,信息是可以传递和共享的,信息可以被重复使用而不会像物质和能源那样产生损耗。比如,电视台每晚播放的新闻,有很多人同时在观看,而新闻节目主持人却不会因为播放这些信息而失去它们。信息的共享性特点与物质和能源相比,有很大的不同,物质和能源一旦被人占有,其他人就得不到了。例如,某位同学把一支钢笔送给了好友,他自己就失去了那支钢笔。

第五,信息具有时效性。人们总是要及时掌握最新的、有用的信息。比如,上月某天的气象资料对于安排明天的出行通常是没有用的。又如,对股票投资者来说,及时掌握股票市场的即时行情信息是至关重要的。为此,需要有能力识别信息,能够判定哪些是最新的信息,哪些是过期的信息。



练一练

- (1) 结合生活中的例子,说出信息具有哪些特征。
- (2) 举例说明信息和信息的载体。



阅读材料

信息与知识

英国哲学家弗兰西斯·培根提出过一个著名的论断：“知识就是力量。”它曾鼓舞许许多多青年人刻苦学习，努力掌握知识。那么，什么是知识，与信息又有什么关系呢？

《辞海》(1999版)对知识的解释是，“知识是人类认识的成果或结晶。依反映对象的深刻程度，可分为生活知识和科学知识；依反映层次的系统性，可分为经验知识和理论知识。经验知识是知识的初级形态，系统的科学理论是知识的高级形态。按具体的来源，知识可分为直接知识和间接知识。但是，从总体上说，人的一切知识(才能也属于知识范畴)都是在后天社会实践中获得和形成，是对现实的反映。社会实践是一切知识的基础和检验知识的标准。知识(精神的东西)借助于一定的语言形式或物化的某种劳动产品的形式，可以交流和传递给下一代，成为人类共同的知识财富”。

可见，知识是人们在社会实践中积累起来的，是人们对各种自然现象和社会现象的认识的总结，是系统化、规范化、结构化的信息。人们从实践中获取到信息后，经过大脑的加工，就可能转换为系统的知识。在这种意义上说，信息是知识的原料，信息经过加工、提炼，去粗取精、去伪存真，由此及彼、由表及里，抽象概括而成为一种知识。

由此可见，在知识的形成过程中，涉及许多信息的采集、处理、交流方面的问题。同样的信息能够加工出什么知识，对于不同的人可能是不同的，这与各人的实践经验和所掌握知识的深广度有关。譬如，同样一个沙尘暴消息，环境生态学家的认识和一个休闲旅游的人所形成的知识就不一样。又如，一件交通事故，交通民警和路上行人所得知识也不同。

1.2 信息的编码

信息本身是看不见摸不着的，但是它可以用一定的方式表现出来。通常人们把用来表示信息的符号组合叫做信息的代码。例如：由18位数字组成的我国公民身份证号码，前6位代码用来描述居民户籍所在的省、市、地区信息，接下来8位是该公民的出生年、月、日的信息，最后4位是序列号及校验码。又如我国的电话号码是由两组数字连接而成的，前一组数字表示地区代码，如上海是021，北京是010，杭州是0571等，接下来的一串数字则是本地固定电话的号码。

在信息科技中，特别是在计算机领域，“代码”两个字具有特指性，是指由“0”、“1”两个符号组成的数字代码。因为数字计算机只能识别和处理由“0”、“1”符号串组成的代码。所以，其他信息代码都要转换成这种由“0”、“1”符号串构成的代码，才能被计算机识别和处理。

客观世界的大量事物、概念的存在状态与变化方式都可用“0”、“1”两种符号的组合来表示。17世纪，德国数学家莱布尼兹提出了二进制记数系统。19世纪爱尔兰逻辑学家乔治·布尔创立了逻辑代数，实现了用数字方法来研究逻辑命题，把对逻辑命题的思考过程转化为对符号“0”、“1”的某种代数演算。20世纪40年代以后，在自动控制和电子技术中大量应用开关线路，迫切需要用数学工具来处理开关线路中日益复杂的逻辑问题，进一步推动了布尔代数的发展，使其内容日益丰富。电子计算机本身是由众多的高速电子开关组合而成的。著名科学家冯·诺依曼关于电子数



字计算机系统结构的经典性建议中,有一条是关于计算机内的信息,包括数据和程序都应采用二进制代码表示,这已成为业界共同遵守的标准。电子计算机将所有输入的信息(数据、程序等)都转化为机器能识别和处理的二进制数字代码。由于二进制代码中用到的只有“0”和“1”两个符号,从而可以方便地用电脉冲、电位、电路的状态、磁化的极性方向来表示。

在使用计算机进行信息处理时,首先要对信息进行编码,把问题转化成二进制代码的计算问题。要使计算机能够处理文字、声音、图像和视频等信息,采用正确的编码方法是首先要解决的问题之一。

1.2.1 二进制代码的特征

计算机采用二进制代码可以方便地存储、处理和传送信息。二进制计数系统的特点是:

- (1) 有两个基本数码: 0, 1。
- (2) 采用逢二进一的进位规则。
- (3) 每个数码在不同的数位上, 对应不同的权值。

例如, $(1101.01)_2 = 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2}$ 。

二进制数不便于书写和记忆,人们经常采用十六进制数来表示它们,因为两者之间的转换非常方便。每4位二进制数可以用1位十六进制数字代替,见表1.2.2。

表 1.2.2 进制制转换

十进制	二进制	十六进制
0	0000	0
1	0001	1
2	0010	2
3	0011	3
4	0100	4
5	0101	5
6	0110	6
7	0111	7
8	1000	8
9	1001	9
10	1010	A
11	1011	B
12	1100	C
13	1101	D
14	1110	E
15	1111	F

例如, 8 位二进制数 11010010B 可以转换为 2 位十六进制数 D2H。十六进制数转换为二进制数时, 每一位十六进制数对应 4 位二进制数。例如, 7FH 可以转换为 8 位二进制数 01111111B。

1.2.2 字符编码

计算机除了要处理数值类型的数据外,还要处理各种非数值类型的数据,例如英文字母和汉字。为了能让计算机存储、处理这些数据,需要为



知识链接

进位制标识

为了区别各种进位制的数码,通常用一个下标来表示该数的进位制(十进制数可以省略),也可以在该数的最后以字母来标识,见表 1.2.1。

表 1.2.1 进位制标识

进位制	二进制	十进制	十六进制
标识	B	D	H



每个字符(character)规定一个二进制形式的代码。当然, 这种0、1组合的编码, 是人为的, 可以有各种各样的编码方案, 但为了便于信息的交换, 必须采用标准化编码。目前, 国际上普遍采用的一种字符编码是ASCII码(American Standard Code for Information Interchange 美国信息交换标准码)。该编码使用7位二进制数, 由128个代码组成(码值范围为0~127), 见表1.2.3。ASCII码包含两部分: 94个图形字符码和34个控制字符码。图形字符包括52个大小写英文字母、10个数字符号、32个标点及其他常用符号, 它们的十进制代码值的范围从33到126。控制字符有34个, 包括10个传输控制符、6个版面调整符、4个设备控制符、4个信息分隔符和10个特殊控制符, 它们的十进制代码值的范围为0~32和127。

表 1.2.3 ASCII 字符编码表

代码	字符	代码	字符	代码	字符	代码	字符
0	NUL (null)	32	Space	64	@	96	`
1	SOH (start of heading)	33	!	65	A	97	a
2	STX (start of text)	34	"	66	B	98	b
3	ETX (end of text)	35	#	67	C	99	c
4	EOT (end of transmission)	36	\$	68	D	100	d
5	ENQ (enquiry)	37	%	69	E	101	e
6	ACK (acknowledge)	38	&	70	F	102	f
7	BEL (bell)	39	'	71	G	103	g
8	BS (backspace)	40	(	72	H	104	h
9	TAB (horizontal tab)	41	)	73	I	105	i
10	LF (NL line feed, new line)	42	*	74	J	106	j
11	VT (vertical tab)	43	+	75	K	107	k
12	FF (NP form feed, new page)	44	,	76	L	108	l
13	CR (carriage return)	45	-	77	M	109	m
14	SO (shift out)	46	.	78	N	110	n
15	SI (shift in)	47	/	79	O	111	o
16	DLE (data link escape)	48	0	80	P	112	p
17	DC1 (device control 1)	49	1	81	Q	113	q
18	DC2 (device control 2)	50	2	82	R	114	r
19	DC3 (device control 3)	51	3	83	S	115	s
20	DC4 (device control 4)	52	4	84	T	116	t
21	NAK (negative acknowledge)	53	5	85	U	117	u
22	SYN (synchronous idle)	54	6	86	V	118	v
23	ETB (end of trans. block)	55	7	87	W	119	w
24	CAN (cancel)	56	8	88	X	120	x
25	EM (end of medium)	57	9	89	Y	121	y
26	SUB (substitute)	58	:	90	Z	122	z
27	ESC (escape)	59	;	91	[	123	{
28	FS (file separator)	60	<	92	\	124	
29	GS (group separator)	61	=	93	]	125	}
30	RS (record separator)	62	>	94	^	126	~
31	US (unit separator)	63	?	95	_	127	DEL

例如, 字母“t”的ASCII码是116, 也就是1110100B。数字“3”的ASCII码是51, 也就是0110011B。符号“>”的ASCII码是62, 也就是0111110B。

计算机存储器在存放ASCII码时, 占用一个字节(byte)(二进制8位)



的右面 7 位，最左位用“0”填充。

了解字符在计算机内的表示方法后，很自然会产生一个问题：究竟怎样区分数值和字符呢？例如，内存中有一个字节的内容是 65，它究竟表示的是 65 这个数值，还是表示字母“A”，或是其他？如果面对一个孤立的字节，确实无法区分，但存放和使用这个数据的软件，会以其他方式保存这些数码属于什么类型的信息，指明这个数据是数值类型，或是字符类型，或是其他类型。

1.2.3 汉字编码

汉字与西文字符一样，也是一种字符，在计算机内同样是以二进制代码形式表示的。用计算机处理汉字信息有以下几个问题需要解决：

1. 怎样将汉字输入计算机？
2. 在计算机内部怎样处理汉字？
3. 在各计算机系统之间怎样交换汉字信息？
4. 计算机怎样实现汉字信息的输出(显示)？

图 1.2.1 表示了汉字输入、处理、输出编码之间的关系。

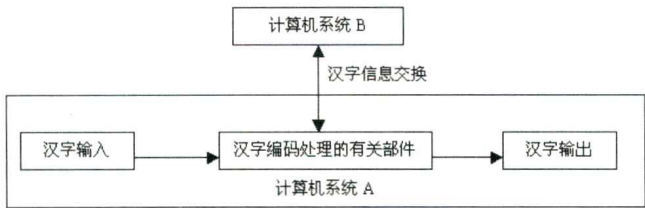


图 1.2.1 汉字输入、处理、输出编码关系图

实际上，上述问题的解决是通过相应的汉字编码方案来实现的。

1. 汉字输入

向计算机输入汉字的方法有两大类：一类是自动识别方式，包括汉字字形的自动识别和汉字语音的自动识别；另一类就是将汉字编码输入，这种用来将汉字输入计算机的编码称为输入码(又称外码)。

输入码通常是利用汉字的音、形或其他的相关特征信息，按照一定的规则，对指定的汉字集编制相应的代码。按照汉字的字音特征编码，称为音码(如智能拼音码)；按照汉字的形体结构特征编码，称为形码(如五笔字型码)；以字音为主，辅以字型特征的编码，称为音形码。反之，则称为形音码。汉字输入编码方案品种繁多，已经制成电脑软件可上机运行的方案约有几百种之多(见图 1.2.2)。

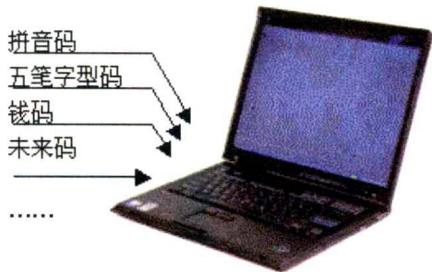
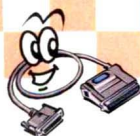


图 1.2.2 多种多样的汉字输入编码

2. 交换码

为了方便各种数字系统(不仅是计算机系统)之间的汉字信息的通信交换，早在 1981 年，国家就颁布了编号为 GB2312-80 的标准《信息交换用



汉字编码及字符集》，这种汉字信息交换用的代码又称为区位码。区位码分成 94 个区，每区包含 94 个位，构成一张 94×94 个单元的表格。每个字符占一个单元，因此每个字符都可以用其所处的“区号”(称为区码)和该区中的“位号”(称为位码)来表示，以此构成该字符的区位码。见图 1.2.3。

区 \ 位	01	02	03		92	93	94
01							
16	啊	阿	埃		包	褒	剥
17	薄	雹	保		秉	饼	炳
.....							
93							
94							

图 1.2.3 区位码表示图

在计算机中，用第一个字节标记区码，第二个字节标记位码。

3. 处理码

处理码是计算机内部用于信息处理的汉字代码，也称汉字机内码。

一个区位码占两个字节，每个字节最高位为“0”；英文字符的机内码是 7 位 ASCII 码，最高位也是“0”。为了在计算机内部能够区分是汉字编码还是 ASCII 码，将区位码的每个字节的最高位设置为“1”，并将区号和位号各增加一个适当的常数，构成汉字机内码。这样，汉字机内码的每个字节都大于 128，而每个西文字符的 ASCII 码值均小于 128，两者就不会混淆了。需要说明的是并不是说交换码一定与处理码不同，例如后来制订的标准 ISO/IEC10646 以及 Unicode 所用的交换码和处理码就是统一的。

4. 字形码

为了汉字的输出显示和打印，需要描述汉字的字形，汉字字形通常有两种表示方式：点阵方式和矢量方式。这种对汉字字形的编码，称为汉字的字形码。

图 1.2.4 表示了汉字输入码、处理码、交换码、字形码之间关系。

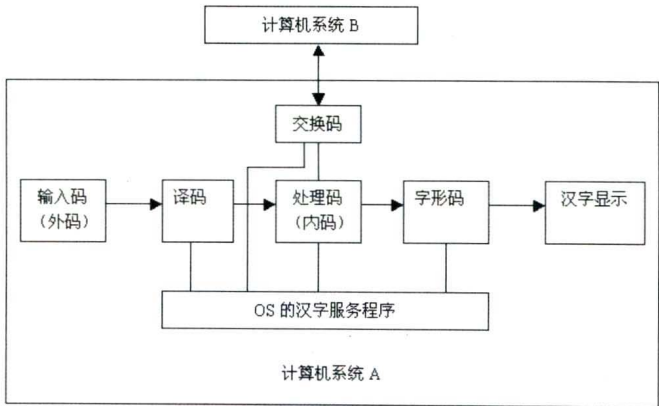


图 1.2.4 汉字输入码、处理码、交换码、字形码之间关系图

在图 1.2.5 中，用 16 进制形式显示了“PC 个人电脑”这几个字的内码，共 10 个字节。其中英文字母使用单字节的 ASCII 编码，汉字使用双字节的 GB 18030 - 2000 编码。

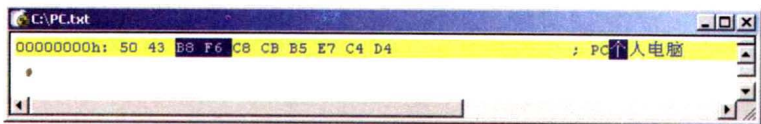


图1.2.5 字符的内码



实践体验

1. 实验名称：观察字符的内码。

2. 实验要求：了解计算机操作系统中字符的编码形式。

3. 实验方法提示：

(1) 使用文本编辑器(例如“记事本”工具软件)输入一些英文、汉字和符号，生成文本文件。

(2) 使用 UltraEdit 或 WinHex 工具软件打开文本文件，以 16 进制形式显示字符的内码。

(3) 试着自己编辑修改字符，观察内码的变化。

4. 结果呈现：与其他同学交流在本次实验中工具软件的使用方法，查看并记录下列字符的内码(以 16 进制形式)。

字符	科	普	知	识	A	B	C
内码							

5. 实验评价：

(1) 填写下表。

项 目	了解内码的形式和作用			
评价指标	完全理解(6分)	较为理解(5分)	基本理解(3分)	有待理解(2分)
评 价				

(2) 填写“学生学业成长记录表”(见附录)。

1.2.4 多媒体信息编码

在汽车刹车过程中，汽车不会立刻停止，只能逐渐地慢下来，直到停止。钟摆平滑地摆动，我们可以预测它的轨迹。打开水龙头，水流量也是逐步地增加。汽车的速度、钟摆的位置和水流量这些都是连续、平滑变化的量，通常称为模拟量。计算机如果要存储、处理它们，首先要将它们数字化，即将它们变成一系列二进制数据。

传感器(sensor)的作用是进行能量方式的转换，例如它可以把各种物理量的变化转换成电流或电压的变化形式。常见的传感器有话筒、温度传感器、光敏传感器、红外传感器、距离传感器等，见图 1.2.6。通过对它们获得的电流或电压波形进行取样和量化，变为数字形式的数字，计算机就可以存储和处理这些信息了。

1. 声音数字化

声音是人们用来表达和传递信息最方便、最常用的一种载体。从物理学上看，声音是一种波，声波通过空气的振动传递到人的耳膜，引起耳膜振动，由听觉神经传到大脑产生听觉效果。话筒以及相关电压放大电路把声波转换成电压的波形，这仍然是一种连续、平滑变化的模拟信号。

模拟信号怎样才能转换成数字信号呢？基本的办法是“采样”和“量化”，通过“采样”和“量化”可以实现模拟量的数字化，这个过程称为“模数转换”(A/D 转换)，承担转换任务的电路或芯片称为“模数转换器”(analog-to-digital converter, 简称为 ADC)。

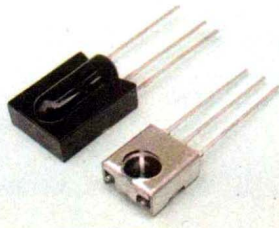


图 1.2.6 声音、光敏和红外传感器

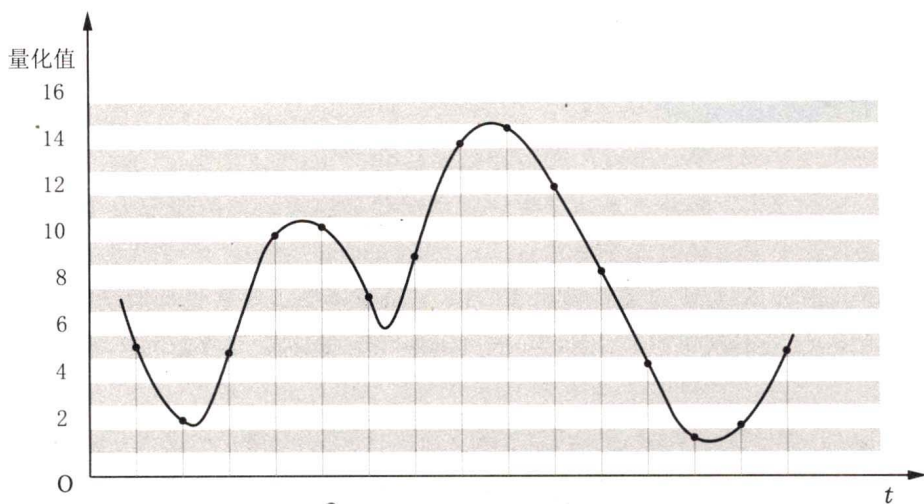
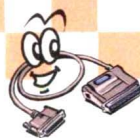


图 1.2.7 模拟量的采样示意

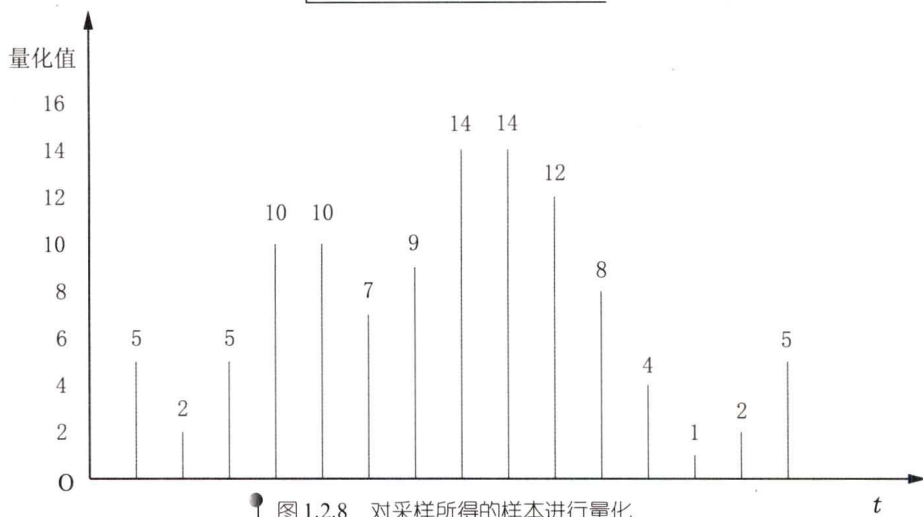
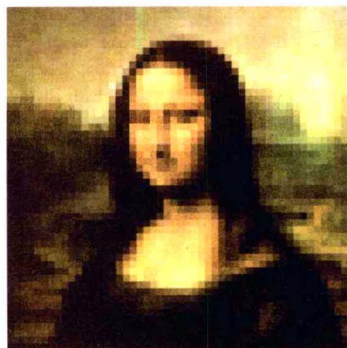
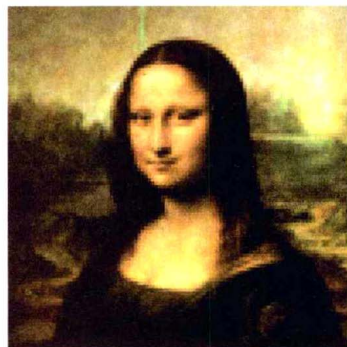


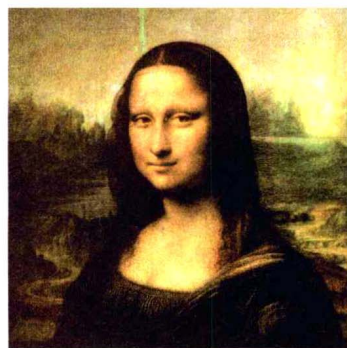
图 1.2.8 对采样所得的样本进行量化



50 × 50 像素



100 × 100 像素



800 × 800 像素

图 1.2.9 像素越多图像越清晰

采样就是按一定的频率，即每隔一小段时间，测得模拟信号的模拟量值，见图 1.2.7。例如，CD 采用的采样频率为 44.1kHz，即每秒钟要采样 44100 次。

采样时测得的模拟电压值，要进行分级量化。方法是按整个电压变化的最大幅度划分成几个区段，把落在某个区段的采样到的样本值归成一类，并给出相应的量化值，见图 1.2.8。

通过采样和量化，一个连续的波形变成了一系列二进制数字表示的数据。数字化声音的质量取决于采样频率和量化分级的细密程度。采样频率越高，量化的分辨率越高，所得数字化声音的保真程度也越好，但是它的数据量也会越大。

在播放声音时，计算机还要进行“数模转换(D/A 转换)”，即将数字化的声音数据转换成模拟声音信号，通过喇叭来播放。

2. 图像和视频数字化

扫描仪(scanner)、数码相机和连接在计算机上的摄像头等可以把图像和视频数字化，在操作系统和有关应用软件的作用下，图像和视频被转换成数据文件存入内存或磁盘。在一些视频处理卡的输入端口可以用来输入模拟视频信号(如录像机的视频输出)，由视频卡进行数字化处理。扫描仪和数码相机等图像设备大多采用 CCD(电荷耦合器件)或 CMOS(互补金属氧化物半导体)图像检光部件。

图像数字化的基本思想是把一幅图像看成由许许多多彩色或各种级别



灰度的点组成的, 这些点按纵横排列起来构成一幅画, 这些点称为像素(pixel)。每个像素有深浅不同的颜色, 像素越多, 排列越紧密, 图像就越清晰。每个像素的颜色都被数字化成一定的数值。在图1.2.9中, 3幅图像的像素分别为: 50×50 , 100×100 和 800×800 。

用扫描仪扫描文字稿, 一般只需要区分黑、白两种颜色。这时, 一个像素使用一位二进制数就能表示了。黑白照片一类的单色的灰度图像, 每个像素可以用一个字节来表示, 一字节可以表示256种不同的灰度。彩色图像中的每个像素, 可以使用3个字节来表示, 每个字节分别表示此像素中的红、绿、蓝成分; 每个字节用来存储该成分在256个等级中的某一量值。

视频是由连续的图像帧组成的。我国使用PAL制式每秒显示25帧。也有使用NTSC制式和SECAM制式的国家。

一幅画、一首歌或一段视频, 经过数字化后产生的数据量很大, 为了提高存储、处理和传输这些信息的效率, 很多种关于图像、声音、视频的压缩(compress)标准被制订出来。JPG是静态图像常用的压缩格式, MP3是音乐信息常用的压缩格式, VCD和DVD格式的影视内容是分别使用MPEG-1和MPEG-2压缩标准来压缩、存储数据的。



练 一 练

- (1) ASCII码包含_____个图形字符码和 _____个控制字符码。
- (2) 什么是汉字编码?
- (3) 为了提高信息的存储、处理和传输效率, 一般要对数字化了的多媒体信息进行_____处理。
- (4) 提高声音和图像数字化精度有哪些途径? 受到哪些条件的制约?



阅读材料

申农与信息论

1948年, 美国数学家申农发表了一篇著名的论文《通信的数学理论》, 建立了比较系统的信息理论基础, 即信息论。这门科学大大促进了通信技术和信息技术的发展, 申农的贡献是很大的。

第一, 他提出了信息的形式化。申农认为, 在通信中, 要排除信息的语义, 只考虑信息的形式因素。在通信中, 如果接收端能够把发送端发出的信息从形式上复制出来, 那么也就复制了信息的语义内容。例如, 有三个句子:

“信息是人类社会三大资源之一。”

“(a + b)²=a² + 2ab + b²。”

“Z!12#5\$7**XYZ”。

第一句是正常的句子, 第二句是代数公式, 第三句是含义不清的字符串。这三个句子有明显差别, 但它们都是信息。在通信时, 一旦把它们转变为电信号, 那么这种差别就没有了, 因为都是电信号。当接收端将电信号转换成原来形式时, 语义就恢复了。所以, 信息形式化后, 就有可能用数学进行描述。

第二, 他提出了如图所示的通信系统模型。信源给出要传输的信息; 编码器把信息转变为信号, 使之能在信道中传输; 解码器把信号恢复成信息传给信息的接受方, 即信宿。