



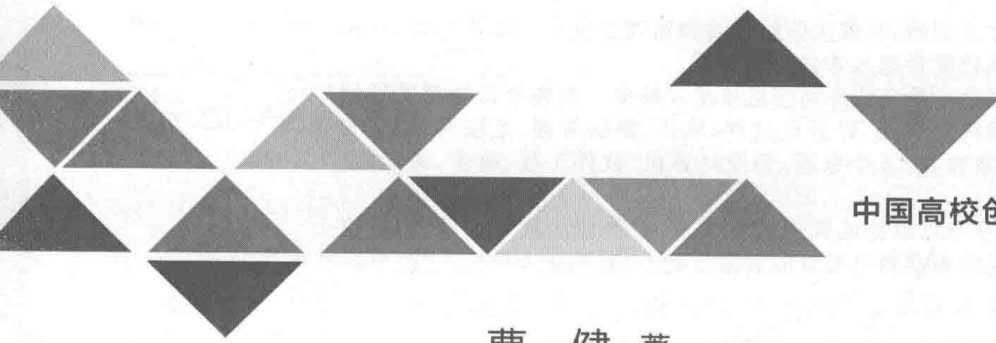
曹 健 著

IT文化

——揭开信息技术的面纱

清华大学出版社





中国高校创意创新创业教育系列丛书

曹 健 著

IT 文化

——揭开信息技术的面纱

清华大学出版社
北 京



内 容 简 介

信息技术已经融入了人们生产和生活的方方面面,其催生的新事物和新理念也令人眼花缭乱、应接不暇,熟练地掌握和运用它们,已经成为一个当代复合型人才必备的素质。

本书主要围绕着网络、数据、计算、智能、安全 5 条主线介绍信息技术发展史上的趣事以及相关领域的核心技术,详细内容分为引领时代的浪潮、编码的奥秘、计算机硬件、软件、数据管理、互联网、信息安全、人工智能、物联网、大数据、云计算、信息的素养共 12 个专题,涉及计算机、软件工程、电子、通信、自动化等多个专业的入门知识。

希望通过本书,使读者能够系统地了解信息学科取得的辉煌成就以及 IT 产业的技术底蕴,初步接触到前沿的、富有挑战性的热点问题,培养敏锐的观察能力和发散的思维能力,激发深入探究、主动创新的科研热情。

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

IT 文化:揭开信息技术的面纱/曹健著. —北京:清华大学出版社,2018
(中国高校创新创业教育系列丛书)
ISBN 978-7-302-49914-5

I. ①I… II. ①曹… III. ①信息技术—研究 IV. ①G202

中国版本图书馆 CIP 数据核字(2018)第 055451 号

责任编辑:谢琛 战晓雷

封面设计:常雪影

责任校对:李建庄

责任印制:丛怀宇

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

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

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 装 者:北京嘉实印刷有限公司

经 销:全国新华书店

开 本:210mm×235mm

印 张:26.5

字 数:638 千字

版 次:2018 年 7 月第 1 版

印 次:2018 年 7 月第 1 次印刷

定 价:79.00 元

产品编号:078328-01





前言

从大学阶段学习信息技术开始,我感觉到社会大众对这个领域的看法趋向两个极端:一部分人觉得信息技术非常高、大、上,进入这个圈子的人要么是收入丰厚的白领,要么是智商超群的精英;另一部分人觉得信息技术特别难、苦、累,从事相关行业的人要么是满面沧桑的码农,要么是呕心沥血的极客。为什么两种看法差别如此之大,简直是“天渊之别”?我想这很可能是由于信息技术自身的特点造成的:它不仅与各个学科交织在一起,而且和人们的生活息息相关;人们很容易以偏概全,拿自己接触到的部分当作整体,将个例当作普遍;更何况,信息产业就是经济形势的晴雨表,每两三年就可能让你坐上一趟“过山车”,体会一下什么叫“冰火两重天”……

我自毕业之后就打算写一本介绍信息技术的书,主要也是被两方面的需求所驱动:一方面是觉得有必要给身边的亲戚朋友科普一下,澄清大家对信息技术的误解,以免大家见面都问我电脑修得熟不熟练,漏洞补得瓷不瓷实,网络疏通得拿不拿手;另一方面是在北京工商大学的教学改革中,我尝试开设了一门面向全校本科生的通识选修课——“信息技术前沿”,很需要一本科普性的教学参考书,虽然我翻阅了上百本信息技术类的书籍,但目前还没找到非常合适的书。

从事科研和教学的人不难发现,跟同一领域的专家或者学生交流比较容易。因为大家有相同的知识背景,“行家一出手,就知有没有”,专业词汇和公式一写,不用太多解释,就已“了然”。而“隔行如隔山”,一旦要面向亲朋好友这样的普通大众,就没这么省事儿了,就算你从头至尾全是中文表述,他们还觉得你说得像外语。同理,来上我这门课的很多学生都是文、商、艺术类专业的,即使他们都上过“计算机文化”之类的基础课程,也不能指望他们在完成自己繁重的学业之后,再去苦读信息技术方面的专业书籍。

可以想象,通识课程与专业课程的培养目标、教学内容和授课方式也是大不一样的。我开这门课的初衷就是希望通过课堂上教师的讲述以及多媒体材料的展示,让同学们了解到信息技术如何应用到生产生活的方方面面,透过令人眼花缭乱的各种新理念认识到它们的本质,并初步具备一个当代大学生必需的信息素养。我从一开始就把这门课定位在两个字上——“易”和“趣”。

“易”主要指的是“容易”。长久以来,信息技术类的课程一直以内容深奥、公式繁杂、门槛较高著称。这让非专业人士学习起来无从下手、举步维艰,以致大家“不明觉厉”、敬

而远之……我在课程的准备过程中将相关内容重新整理,粗分为12个专题,采用深入浅出的方式,结合生活中的实际例子来描述信息技术的基本原理、方法和应用,即使难度较大的地方也只需要具备中学的基础就能理解。

“趣”指的就是“有趣”。老师尽可能采用举例子、打比方、播视频、放图片的形式完成课堂教学,并讲述相关大牛的趣事、信息产业的规律、知名企业的境况,这不仅有利于吸引同学们的注意力和活跃课堂氛围,而且对于学生扩展知识面、培养发散思维、提升就业竞争力都有很大的帮助。

这门课程教过几轮之后,我开始着手整理这本科普书籍,一边写作一边继续丰富我的教学材料,因此本书的主要框架和课堂教学安排基本一致。当然,没有了授课时间这样的限制,书中就可以在有些地方有意思的地方铺展开来,并通过“扩展阅读”和“附录”的形式保证每个主题内容的相对完整。至于正文中间的小字内容和旁注,有兴趣的读者可以简单看一下,直接跳过也不会影响对主体内容的理解。

有必要说明一点,本书与课程的目标一致,着重于从纷繁复杂、体量庞大的信息技术领域里捋出清晰的脉络。因此有很多技术细节被我刻意忽略,以免导致信息冗杂而使读者产生认知负担。所以建议大家把这本书看成了了解信息技术的入门读物,搭起一个简易的“大框架”之后,再去深入了解每一个小领域的技术细节就会变得轻松一些。

我在写作过程中参考了与信息技术相关的大量书籍,详细书目参见每一章最后列出的参考文献。可以说,没有前辈们的贡献,就没有本书的出版,也希望本书能成为这些参考文献的前期读物。在此,向有关作者表示由衷的感谢。

书中一部分图片是作者创作的,一部分引用的图片明确标注了来源,其余图片大都来自百度图片。书中引用的图片如有版权问题,请原作者告知作者,谢谢!

由于信息技术领域仍处于快速发展和不断更迭的状态,加之作者水平有限,书中难免存在一些不足之处,敬请读者批评指正。如果读者希望在信息技术科普或教学方面与作者进行交流,请将邮件发送到 caojian9527@sina.com。

曹 健

2018年4月

于北京工商大学



目 录

●第1章 引领时代的浪潮	1
1.1 信息活动的意义	2
1.2 信息科学与技术	4
1.3 信息技术的革命	6
1.3.1 语言的突破	7
1.3.2 文字的诞生	9
1.3.3 印刷术的出现	12
1.3.4 电磁波的应用	13
1.3.5 计算机的发明	15
1.4 信息时代的到来	17
参考文献	19
●第2章 编码的奥秘	20
2.1 计数有学问	21
2.1.1 掰指头数数	21
2.1.2 画线做记录	23
2.2 简约而不简单	24
2.2.1 伟大的发明	25
2.2.2 统一的框架	27
2.3 用电来计数	29
2.3.1 原始的模拟	29
2.3.2 开关的电路	30
2.4 神奇的二进制	32
2.4.1 运算法则	32
2.4.2 编码字符	34
2.4.3 度量信息	37

扩展阅读	39
布莱叶盲文	39
莫尔斯电码	40
参考文献	42

●第3章 计算机硬件 43

3.1 元器件的进化	44
3.1.1 自动控制的开关	44
3.1.2 从继电器到晶体管	47
3.1.3 集成出来的奇迹	49
3.2 理论指导实践	51
3.2.1 抽象的模型	51
3.2.2 系统的结构	54
3.3 应用无处不在	56
3.3.1 超越科学计算	57
3.3.2 走进千家万户	59
3.3.3 颠覆传统行业	61
3.4 样式千姿百态	64
扩展阅读	67
计算机的那些“祖先”	67
“硅谷”名字的由来	69
参考文献	71

●第4章 软件..... 72

4.1 分门别类话软件	73
4.2 强大的操作系统	76
4.2.1 分身有术	76
4.2.2 管理得法	82
4.3 PC时代的主角	87
4.3.1 打造产业链	87
4.3.2 安迪-比尔定律	89
4.4 软件工程的困境	92
4.4.1 软件危机	92

4.4.2 质量保证	95
扩展阅读	98
屡试不爽的“三板斧”	98
业务扩展的背后推力	100
参考文献	102
●第5章 数据管理	103
5.1 决策的依据	104
5.1.1 多算胜少算	104
5.1.2 数据的文化	107
5.2 科学的方法	110
5.2.1 站在前人的肩上	110
5.2.2 事实胜过雄辩	113
5.2.3 提高质量的法宝	115
5.3 管理的技术	118
5.3.1 数据的批量处理	118
5.3.2 数据库的基本思想	122
5.3.3 挖掘数据中的金矿	126
扩展阅读	128
IBM——数据处理的先行者	128
Oracle——平淡无奇的传奇	129
参考文献	131
●第6章 互联网	132
6.1 互联网的结构	133
6.1.1 连接的开始	134
6.1.2 规则的统一	136
6.1.3 分组的传递	139
6.2 互联网的服务	143
6.2.1 电子邮件和文件传输	143
6.2.2 万维网——互联网的“灵魂”	145
6.2.3 域名和域名服务	149
6.2.4 服务提供的方式	152

6.3 互联网的进化	155
6.3.1 互联网 1.0 时代	156
6.3.2 互联网 2.0 时代	158
6.3.3 互联网 3.0 时代	162
扩展阅读	165
思科的经营绝招	165
互联网产业的规则	167
参考文献	168
●第 7 章 信息安全	169
7.1 密码学基础	170
7.1.1 古典密码的技术	170
7.1.2 现代密码的思想	174
7.1.3 计算上的安全	178
7.2 完整性技术	181
7.3 身份认证	184
7.3.1 用户所知道的	185
7.3.2 用户所拥有的	187
7.3.3 用户生物特征	190
7.4 网络攻防	192
7.4.1 谁在攻击网络	193
7.4.2 计算机病毒	196
7.4.3 常见黑客技术	199
7.4.4 网络安全防御	201
扩展阅读	203
“一代名机”Enigma	203
“世界头号黑客”凯文·米特尼克	207
参考文献	209
●第 8 章 人工智能	210
8.1 智能的前世今生	210
8.1.1 古代的机械智能	211
8.1.2 人工智能的起源	212

8.1.3 人工智能的诞生	214
8.2 思维的探索研究	215
8.2.1 状态搜索与博弈	216
8.2.2 知识表示与推理	219
8.2.3 神经网络与学习	223
8.3 机器人的广泛应用	227
8.3.1 传统机器人	227
8.3.2 半机器人	229
8.3.3 智能机器人	230
8.4 智能时代的隐忧	231
8.4.1 “机器上岗”之利	232
8.4.2 “失业危机”之痛	233
8.4.3 “奇点临近”之惧	236
扩展阅读	238
谷歌公司的智能搜索	238
自然语言处理难在何处?	241
参考文献	243

●第9章 物联网 244

9.1 网络协议的变革	245
9.2 独一无二的标识	248
9.2.1 一维条形码	248
9.2.2 二维条形码	250
9.2.3 射频识别	252
9.3 时时刻刻的定位	257
9.3.1 卫星定位	258
9.3.2 基站定位	260
9.3.3 网络定位	262
9.4 全面感知的技术	264
9.4.1 感官的延伸	264
9.4.2 智能的终端	266
9.5 美好生活的憧憬	268
9.5.1 构建智能的家居	269

9.5.2 点亮智慧的地球	273
扩展阅读	277
生活在智慧城市的一天	277
物理世界与信息世界的大融合	278
参考文献	280
●第10章 大数据	281
10.1 大数据的产生	283
10.1.1 大数据的来源	283
10.1.2 大数据的特征	286
10.2 大数据的思维	289
10.2.1 全体数据的威力	290
10.2.2 观其大略的意识	295
10.2.3 预测未来的能力	299
10.3 大数据的应用	307
10.3.1 互联网与大数据	308
10.3.2 物联网与大数据	311
10.4 大数据的挑战	314
10.4.1 没有遗忘的时代	315
10.4.2 全息可见的困境	318
扩展阅读	321
情报整理——多维度的数据分析	321
缉毒新招——大数据让你无所遁形	322
参考文献	324
●第11章 云计算	325
11.1 云计算的先驱	326
11.2 云计算的概念	328
11.2.1 构建数据中心	329
11.2.2 租用网络资源	331
11.2.3 提供云端服务	333
11.2.4 云计算的本质	335
11.3 云计算的实现	338

11.3.1 核心技术问题	339
11.3.2 全社会的关注	342
11.3.3 经济上的考量	343
11.4 云计算的应用	344
11.4.1 人工智能的飞速发展	345
11.4.2 信息产业的重新布局	347
11.4.3 生产生活的全面影响	350
扩展阅读	354
“并行”不是那么简单	354
节能省电各出奇招	355
参考文献	357
●第12章 信息的素养	358
12.1 思维方式的转变	359
12.1.1 机械思维的贡献	359
12.1.2 不确定性的颠覆	361
12.1.3 新时代的方法论	363
12.2 信息时代的生活	366
12.2.1 拓展带宽的依据	366
12.2.2 学习观念的转变	369
12.2.3 数据主义的困局	375
12.3 信息时代的创业	378
12.3.1 挖掘“第一桶金”	379
12.3.2 站在“风口浪尖”	381
12.3.3 创新“商业模式”	384
参考文献	389
●附录A 计算机的外部设备	390
A.1 海量存储器	390
A.2 输入设备	392
A.3 输出设备	396

●附录 B IT 产业的那些定律	398
B.1 摩尔定律	399
B.2 安迪-比尔定律	399
B.3 70 - 20 - 10 定律	400
B.4 诺威格定律	402
B.5 基因决定定律	403
●附录 C 信息时代的科学基础	405
C.1 信息论	405
C.2 控制论	406
C.3 系统论	407
●后记	409

第1章

引领时代的浪潮

最高深的技术是那些令人无法察觉的技术,这些技术不停地把它们编织进日常生活,直到你无从发现为止。

——马克·韦泽(普适计算之父)

“信息”这个词很早就出现在人们的生活中,唐代杜牧有诗云:“塞外音书无信息,道傍车马起尘埃。”鱼玄机曾在诗中写道:“春来秋去相思在,秋去春来信息稀。”宋朝词人柳永更是伤怀:“自别后,幽怨与闲愁,成堆积。鳞鸿阻,无信息。梦魂断,难寻觅。”中国古典名著《水浒传》中也有一段描写:“宋江大喜,说道:‘只有贤弟去得快,旬日便知信息。’”时至今日,“信息”一词在日常用语中出现得更加频繁。我们会经常说:“这人真有水平,短短一句话,字字珠玑,信息量很大啊!”或者埋怨同伴:“你啰里巴唆地讲了半天,没啥有用的信息……”但信息究竟是什么?如何度量信息的多与少?我们这个时代为什么被称为“信息时代”?仔细想一想,这些在我们的头脑中还都是一个个不小的问号呢。

同样,“信息技术”也是很早就在人类文明史上占据了重要的地位。且不说人类的语言、文字、珠算都属于信息技术的范畴,我们引以为傲的“四大发明”中就有两项——造纸术和印刷术——也是实打实的信息技术。“诗圣”杜甫的《春望》中,有一句耳熟能详的描述:“烽火连三月,家书抵万金。”这里面的“烽火”和“家书”都是古代最常用的信息技术。到了现代,就更不用说了,你手中的电话、桌上的电脑、客厅里的电视、教室里的投影、街道上的广播、太空中的卫星,无一不是信息技术的产物。尤其是它的英文缩写——IT (Information Technology)热得发烫,几乎到了无人不晓的地步。IT产业不仅是当今全球经济形势的晴雨表,也是世界经济前五大行业中发展速度最快的,没有之一!那么,问题来了:信息技术可以分为哪些门类?它们在人们的工作生活中有哪些具体应用?将来会变成什么样子?只有搞清楚这些问题,我们才算拥有现代社会所需的基本见识,才能初步具备一个复合型人才的信息素养。

下面,就让我们做好准备,带着一颗好奇之心去领略IT世界的风光。这个世界是如

世界经济的
前五大行业为金融、信息技术、医疗和制药、能源以及日用消费品。其中,只有IT一个行业可以以持续翻番的速度进步。

此的瑰丽雄奇、宽广辽阔,相信你总能找到令自己心驰神往的景色,获取一份终生难忘的阅历。

1.1 信息活动的意义

这次实验的研究成果被写成论文,发表在1954年的《加拿大心理学研究》(Canadian Journal of Psychology)。

在20世纪50年代初,一群心理学家在加拿大的麦克吉尔大学做了一个实验。这个实验在当时看来似乎有些莫名其妙,但意义深远,后人称之为“感觉剥夺实验”。如图1.1所示,实验方让志愿者长时间待在缺乏刺激(没有图形知觉,限制触觉和听觉)的环境中,来测试“感觉剥夺”对人体的影响。具体地说,就是整个实验在隔音室里进行(用空气调节器的单调嗡嗡声代替人的听觉的背景噪声),让志愿者戴上特制的半透明的塑料眼镜(可以透进光亮但看不见周围物体),手和臂上都套有纸板做的手套和套袖,然后静静地躺在舒适的帆布床上。

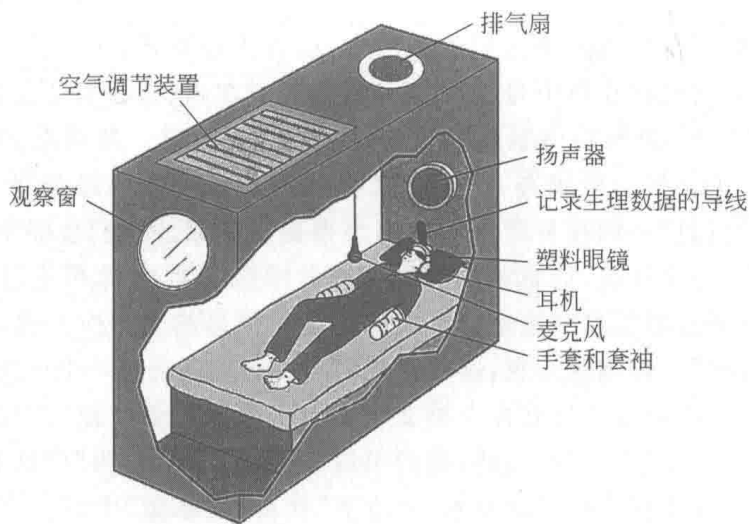


图 1.1 “感觉剥夺”实验示意图

实验支付的报酬是每天20美元,而当时大学生打工一小时大约只能挣50美分,这让很多大学生跃跃欲试,认为利用这个机会可以好好睡一觉,或者考虑论文、课程计划。但结果却令很多人大跌眼镜,没过几天,这20多个志愿者就纷纷退出。有材料记录了他们当时的状况:

他们感到非常难受,根本不能进行清晰的思考,哪怕是在很短的时间内注意力都无法集中,思维活动似乎总是“跳来跳去”。更为可怕的是,50%的人出现了幻觉,包括视幻觉、听幻觉和触幻觉,例如,出现光的闪烁,似乎听到狗叫声、打字

声、滴水声等,感到有冰冷的钢板压在前额和面颊,或感到有人从身体下面把床垫抽走。在过后的几天里,被试注意力涣散,不能进行清晰的思考,智力测试的成绩不理想。通过对脑电波的分析,证明被试的全部活动严重失调,有时被试甚至出现了“白日做梦”的现象。

通过这个实验不难看出,人的身心要想保持在正常的状态,就需要不断地从外界获得刺激,也就是和周边环境进行信息交流。一旦感觉被剥夺了,人会产生难以忍受的痛苦,各种心理功能将受到不同程度的损伤。所以说,信息是人类生存必不可少的基本条件,也是人类生存的基本需求。

其实,所有的生物每天都在进行信息活动,如寻找食物、迁移、辨别敌害、辨认配偶等等,这是生物与生俱来的决定存亡的能力。当然,它可能只是很简单的本能,如当微生物来到酸碱度不适宜的环境中就会逃走;也可能需要高等的能力,如一个人需要从橱柜下面倒数第二个抽屉里取出复习参考书。

感觉剥夺实验企图屏蔽人们和周围环境的所有信息交流,这显然是过于极端了。如果把条件放宽一些,允许人们不断从自然界获得刺激,但是禁止和其他人进行信息交换,也就是说屏蔽人际沟通,对人们的影响是不是就无关痛痒了呢?心理学家的研究同样会否定这个论断——仅仅是人类之间的信息活动,即人际沟通,也是不可或缺的!

很多人在小时候曾经玩过这种游戏:孩子们会先选出一个“受害者”——也许是因为他犯了一个实际的或想象中的错误,也许纯粹是为了好玩。不久之后,惩罚就以沉默的形式开始了。没有人跟他或她说话,也没有人会回应这个倒霉鬼所说的话或做的事。如果你曾经被如此教训过,你可能会体验到一连串的情绪。刚开始你会觉得(至少会表现出)无所谓,但过了一阵子,这种被当成透明人的压力开始萌生。如果这个游戏持续得足够久,你会发现自己要么已经退缩到某种令人沮丧的状态中,要么已经对别人产生敌意。敌意一方面是为了表达自己的愤怒,另一方面是想要得到别人的回应。

成年人也和孩子们一样,会把沉默以对当作一个有力的工具,用来表达他的不愉快。我们都能直觉地感受到人际沟通的重要,缺少与他人的接触对任何人来说都是一种非常残酷的惩罚和无比痛苦的经历。那些自己选择或者不得不成为独居者的事例不断地证实着这一点。W. 卡尔·杰克逊(W. Carl Jackson),一位独自航行 55 天、横越大西洋的探险家,描述了大多数独居者的普遍心情:

我发现在第二个月出现的寂寞感使我感到很痛苦。我一直以为自己是一个自给自足的人,但是此刻我终于明白,没有旁人做伴的生活是没有意义的。我开始有了想要跟别人——一个真实的、鲜活的、有气息的人说话的强烈需求。

除了带来情感上的痛苦,若一个人被剥夺陪伴甚至会危及生命本身。腓特烈二世

(Fredrick II, 亦称弗里德里希二世), 公元 1196—1250 年的德国(神圣罗马帝国)皇帝, 可能是第一个以系统化方法证实这个论点的人。一位中世纪的历史学家描述了腓特烈二世独特且残忍的实验:

他命令保姆和护士喂养婴儿, 帮他们洗澡, 但是不准对他们说话, 因为他想要知道, 在没有人跟婴儿接触之前, 婴儿开口会先说哪一种语言——是最古老的希伯来语、希腊语还是拉丁语, 又或者是他们亲生父母说的语言? 最后, 他徒劳无功, 因为所有的婴儿都死了。缺乏养育者的拥抱、慈爱的脸孔、深情的言语, 他们根本无法存活。

1.2 信息科学与技术

我们一般会想当然地认为: 生物体进行信息活动是轻而易举、习以为常的事情, 这种能力的获得也是与生俱来、不足为奇的。但是, 我们静下心来仔细思考一下其中的机理, 尤其是自己动手制造一个物件去模仿实验的时候, 才发现其中大有学问, 往往完成一个简单的动作就需要一个异常复杂的系统工程。

就拿“在黑夜里摸索前行”来说吧, 让我们仔细回忆一下自己是怎么做到的: 当我们接触到某个物体的时候, 体表的感受器(如手部的皮肤)就产生了相应的触觉, 如冷、热、尖锐、粗糙等等; 感受器对这些信息进行编码, 然后通过传入神经传导到神经中枢(脑和脊髓), 由神经中枢中相应的功能区域对信息做进一步的加工处理; 当大脑识别出碰到的究竟是什么物体之后, 再将反馈信息通过传出神经传导到效应器(例如手或脚); 最后由效应器对反馈信息进行解码, 并进行相应的动作, 例如前进、躲避或将其挪开, 如图 1.2 所示。

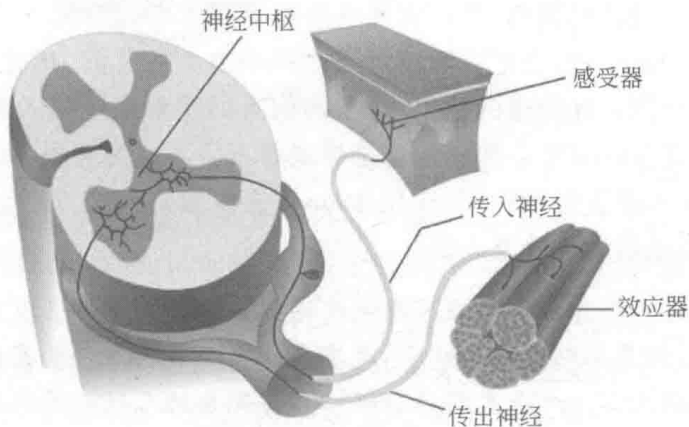


图 1.2 人体的信息器官示意图