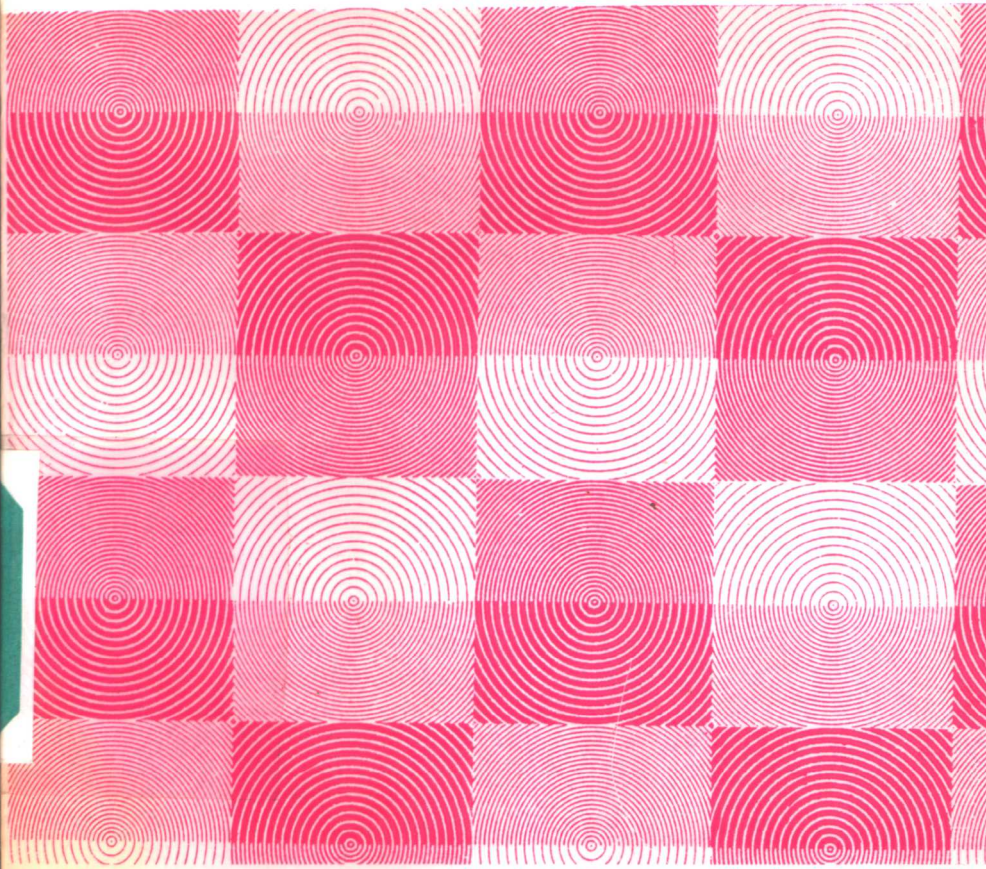


信息学概论

The Fundamentals of Information Science

邹志仁 主编



南京大学出版社

信 息 学 概 论

沈固朝 吕 斌 孙建军 著
朱庆华 施 云 邹志仁

南京大学出版社

信息学概论

邹志仁 主编

南京大学出版社出版发行
(南京大学校内 邮政编码:210093)

南京豪利电脑公司激光照排
扬中印刷厂印刷

*

开本 850×1168 1/32 印张 12.375 字数 322 千
1996 年 9 月第 1 版 1998 年 10 月第 2 次印刷

印数 2001-4000

ISBN 7-305-02987-4/G·453

定价:16.00 元

前 言

自申农创立信息论至今已近半个世纪。40多年来，信息技术获得了飞速的发展并广泛应用于科学研究、经济、生产、管理、文化、社会等各个方面，信息理论的研究对象、领域不断拓宽和深入并向一些相关学科渗透，经过众多科学家、学者的研究和推动，信息理论已发展成为一门大学科——信息科学，这是一个包含多个分支学科的信息学科群。

在当代社会，信息成为社会、经济发展的重要支柱，重要的生产力要素。经济信息化和社会信息化已成为当今社会、经济发展的世界性潮流。

1992年，国家科委将“科技情报”改为“科技信息”、“国家科委情报司”改为“国家科委信息司”，紧接着国家教委将大学本科专业目录中的“情报学”专业改为“信息学”专业，许多高等学校的图书馆学情报学系纷纷更名为信息管理系。自此，“信息”一词成为我国报纸、期刊、大众传播媒介中使用频率最高的词汇之一。这些引发了我国学术界关于“情报”与“信息”、“情报学”与“信息学”异同或关系的争鸣，推动了信息学科的研究热潮。

著者在多年研究和教学实践的基础上，撰写了本书，意图以宽阔的视野、从更一般的意义上论述“信息”的各个方面，构筑“信息学”的理论框架，以对信息学的学科建设有所贡献，并适应高等学校专业教学和广大信息工作者的需要。由于著者学术水平所限，书中疏漏和不妥之处在所难免，欢迎专家、学者和广大读者批评、指正。

本书第一章由邹志仁撰写，第二、七章由孙建军撰写，第三、九章由吕斌撰写，第四、六章由沈固朝撰写，第五章由施云撰写，第八章由朱庆华撰写，全书由邹志仁修改、定稿。

著 者

1995. 11.

目 录

1 绪论	(1)
1.1 信息概述	(1)
1.2 信息科学的形成和发展	(15)
2 信息度量	(27)
2.1 信息的传输	(27)
2.2 信息量	(31)
2.3 信息熵	(36)
2.4 互信息量和条件互信息量	(40)
2.5 平均互信息量	(42)
2.6 信息定量方法	(47)
3 信息交流	(49)
3.1 信息交流与社会	(49)
3.2 信息交流的模式和要素	(63)
3.3 信息交流的类型	(76)
3.4 信息交流的效率	(83)
4 信息源	(94)
4.1 文献信息源概述	(95)
4.2 主要的文献信息源	(111)

4.3	非文献信息源	(137)
5	信息整序	(152)
5.1	信息整序的理论基础	(153)
5.2	信息整序的基本原理	(163)
5.3	信息整序方法	(169)
5.4	计算机信息整序方法	(195)
6	信息服务	(204)
6.1	信息服务概述	(206)
6.2	传统信息服务	(212)
6.3	电子信息服务	(234)
7	信息用户	(260)
7.1	用户信息需求	(260)
7.2	用户信息心理过程	(264)
7.3	用户信息需求的行为规律	(273)
7.4	信息用户的分类与培训	(277)
7.5	信息用户研究的基本方法	(278)
8	信息系统	(290)
8.1	信息系统概述	(290)
8.2	计算机化的信息系统	(299)
8.3	信息系统的开发	(307)
8.4	信息网络	(312)
8.5	国家信息系统	(323)
9	信息产业与信息经济	(337)
9.1	信息产业概述	(337)

9.2	信息产业的地位、作用和特征	(352)
9.3	信息产业的测度和信息经济规模	(360)
9.4	信息技术与信息产业	(370)
	参考文献	(380)

1 绪 论

1.1 信息概述

1.1.1 信息的含义

信息广泛存在于自然界、生物界和人类社会（包括人制造的机器）中。物质、能量、信息是构成客观世界的三个基本要素。信息无处不在，无人不用。信息是一个十分普通而复杂的概念。人们对信息有着许多不同的认识和理解——从各自不同的角度来认识和理解。

信息是什么？

N. 维纳在 1948 年发表的《控制论——动物与机器中的通讯与控制问题》中提出：“信息既不是物质，又不是能量，信息就是信息。”

从产生信息的客体来定义信息：从哲学的角度认为，信息是物质的一种普遍属性，本质属性。不同的事物具有不同的本质、特征、运动规律。事物的特征通过一定的媒介或传递形式（如声波、电磁波、图象、文字、符号）使其它事物感知。这些能被其它事物感知的表征该事物特征的信号的内容即为该事物向其它事物传递的信息。所以，信息是事物本质、特征、运动规律的反映。不同的事物有不同的本质、特征、运动规律，人们就是通过事物发出的信息来认识该事物，区别于其它事物。

从信息的发送、传输、接收的过程中，客体和接收（认识）主体之间的相互作用来定义信息：控制论的创始人 N. 维纳在《控制论和社会》一书中，对信息的定义是：“信息是我们适应外部世界并且使这种适应为外部世界所感到的过程中，同外部世界进行

交换的内容的名称。”维纳又说，信息是“生物以及具有自动控制系统的机械系统，通过感觉器官和外界交换的一切内容”。例如，自动门在获得了某种信号后自动关闭（或开启）。

从接收信息的主体来定义信息：申农在1948年发表的《通讯的数学理论》中指出，一个系统所接收的“信息是能够用来消除不确定性的东西”。与此相近的提法有：“传递的消息中使概率发生变化的东西”。认识主体关于客体事物的概率发生了变化，不确定性减少，即获得了信息。例如，每晚电视“新闻联播”之后的“天气预报”，使人们对次日的天气状况有了了解，减少了原来对次日天气状况的不确定性，即获得了信息。

“信息”的定义之所以呈现多样化，主要有三方面的原因。第一，信息本身的复杂性。它是一个多元化、多层次、多功能的混合物。第二，信息科学是一门新兴学科，是一门“大”学科，它有许多分支学科，它的内涵与外延不很确切，而且随着社会、经济和科学技术的发展处于不断发展之中。第三，人们出于不同的研究目的或使用目的，从不同的角度出发，对“信息”作出不同的理解与解释。前两个原因是根本的，第三个原因是前两个原因的派生。

1.1.2 信息的性质

我们暂且不去深究“信息”的确切定义，而来探讨“信息”的性质。我们将“信息”的基本性质归纳如下。

(1) 普遍性

世界是物质的，物质是普遍存在的，物质都处于运动之中。凡有物质及其运动的存在，就有信息的产生。物质及其运动的普遍性，决定了信息的普遍存在性，即使处于第二位的精神（意识、思维），它的活动也会产生信息。信息既存在于没有生命的无机界（自然界、机器），也存在于有机界（生物界、生物体内、人类社会）。

(2) 依附性 (载体不可分性)

信息源于物质及其运动,并以运动的物质为其存在的条件。信息必须依附于物质载体而存在,即需要物质承担者。任何信息都是以某种物质的特定的运动形式表现出来。信息必须有物质承担者,即载体。但信息的载体是可以多种多样的,如磁介质、电磁波、纸张、实物等;其表现方式也可以多种多样,如语言、文字、符号、代码等。信息不可能不依附于任何物质载体(媒介)而存在和交流,即信息具有载体不可分性。

(3) 可度量性

如前所述,信息是物质的普遍属性,信息是无机系统或有机系统通过感觉器官和外界交换的一切内容,信息能够减少或消除系统的不确定性。所以,信息是可以度量的。信息的度量,如同物质和能量的度量一样,关键在于对度量原理的认识和度量方法以及相应的度量标准的确定。目前,信息学界一般以系统不确定性的变化程度来度量信息。

(4) 可识别性

信息是物质的一种普遍属性,是事物本质、特征和运动规律的反映。信息在宇宙中是普遍存在的。人类可以通过自己的感觉器官(或借助于各种仪器设备)来感知它、接受它,进而识别它。人类就是通过感知、识别事物的信息来认识事物。所以,信息的可识别性,是人类能够认识客观世界的基础。

信息是客观存在的,但人类感知(即使是借助仪器设备)、识别信息的能力总是有限的,这就是所谓信息的不完全性。正是由于信息的不完全性,使得人类在每一特定的历史阶段对于客观世界的认识总是有限的、不完全的。

由于种种原因,人类在感知、接收、识别(或借助于仪器)事物信息的过程中会产生错误,客观信息的错误识别导致人们对于客观事物认识的错误,判断和决策的错误。这就是所谓信息的可伪性。

信息的不完全性、可伪性，派生于信息的可识别性。随着科学技术，特别是信息技术的高度发展，人类感知、接收、识别信息的能力不断增强，信息的不完全性、可伪性将不断减少，但不可能完全消除。

(5) 可处理性

信息是可以加工处理的。人的感觉器官能够把接收到的电磁波（光波）、声波、压力、温度等形式的各种物质信息，一律编译成生物电流的脉冲信号，通过神经纤维传给大脑。大脑随即对信息加工处理，进行一系列意识活动。人类对于客观世界的认识，就是通过对信息的加工处理而获得的。

人类为了更好地对信息进行加工处理，创造了一系列的仪器、技术和方法。人们运用这些仪器、技术、方法对信息进行转换、编码、压缩和有序化。

信息是可以存储的。人脑细胞中的大分子DNA（脱氧核糖核酸）是储存信息的具体物质基础。大脑储存信息的过程就是人的记忆过程。人类对于信息的存储，主要的还是通过对信息载体的存储来实现。图书、期刊、资料、档案、磁带、磁盘、光盘等存储了大量信息。信息通过存储可以实现信息的累积。人类对于信息进行存储的目的，是为了利用。为了日后能方便、有效地利用信息，人类创造了一整套对信息加工处理的技术和方法，以实现信息的有序存储。

信息是可以转换的。转换有两方面的含义。一是同种意义的信息可以用不同的形态、方式来表达，并可相互转换，如文字、语言、代码、电磁波可以用来表达同一内容信息。二是同种意义的信息可以依附于不同的物质载体，如纸张、胶片、磁带（盘）、光盘，它们之间也可以相互转换。

信息的可存储性、可累积性、可转换性，派生于信息的可处理性。

(6) 可传递性

信息的产生与信息的传递是联系在一起的,是不可分割的。信息是事物本质、特征和运动规律的反映。信息在信源和信宿之间通过一定的信道(媒介或载体)传递,为信宿(人或仪器设备)所感知和接收。这种传递包括信息在时间上的传递和在空间上的传递。在多数情况下,信息的传递依赖于信息载体的传递,并伴随有对信息的处理、转换和存储。

信息既不是物质,也不是能量,但必须指出,信息的传递离不开物质和能量。在信息的传递过程中必定有一定的物质及其运动的传递或变换,必定有能量的传递或能量形式的变换。例如,文献的传递,电报、电话、电视、广播等都离不开物质和能量的传递和变换。

(7) 可共享性

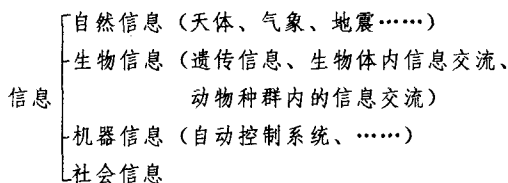
一般的物质、能量资源为所有者拥有,在交换(使用)过程中实现了所有(使用)权的转移,转让方失去,受让方获得,这种交换和转移遵循一定的原则(如等价交换原则)。而信息却具有可共享性。同一内容的信息可以在同一时间或不同时间为两个或两个以上的信宿(用户)获得、使用。信息在传递、交换的过程中,受让(接收)方获得了信息,而转让(发送)方并没有失去信息。可共享性是信息区别于物质和能量的重要特征。当然,不同的信息,共享的范围是不同的,而且信息的共享往往是有条件的,在特定的条件下,对于特定的信息并不实行共享。但是,这些并不否定信息具有可共享性。

1.1.3 信息的结构和类型

信息广泛存在于自然界、生物界和人类社会。信息是多种多样、多方面、多层次的,具有自己的结构和成分。从不同的角度可以对信息总体进行不同的分类。

根据产生信息的客体的性质来划分,信息可分为自然信息、生物信息、机器信息和(人类)社会信息。如表 1-1 所示。

表 1-1

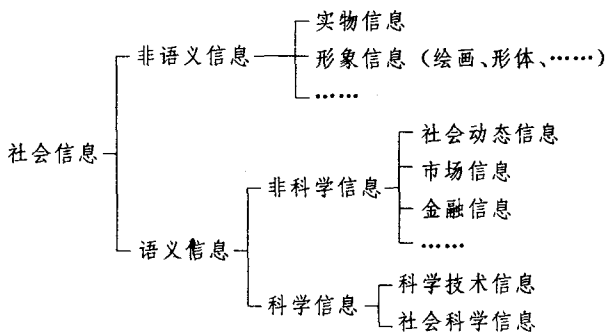


本书主要以社会信息为对象。下面对社会信息的结构和类型作进一步的讨论。

(1) 按信息的性质分

社会信息按其性质可分为语义信息和非语义信息，并可再作进一步的划分，如表 1-2 所示。

表 1-2



语义信息，是指用语言、文字或语义性质的符号（如交通标志、旗语）荷载（表述）的信息。

非语义信息，是指用非语义性质的声音、颜色或实物荷载的信息。非语义信息也是人可以感知，并且对其含义有所认识和理解的信息。因此，就能将其转换成相应的语义信息。就是说，能用相近似的语言、文字表述出来，其含义基本保持不变。

非科学信息，是对社会历史发展过程中的具体现象、事件、事实的描述性信息，不经过逻辑思维加工，直接来源于社会实践，如

社会消息、市场行情、各类社会统计指标等。

科学信息，指作为传递交流对象的科学知识，反映自然界、生物界、社会和思维运动的本质和规律，来源于科学实践，是人们逻辑思维的产物。

(2) 按信息的内容分

按信息的内容范畴可分为：科学信息、技术信息、经济信息、法律信息、政务信息、金融信息、市场信息、营销信息、文化信息、旅游信息、……。

(3) 按信息的表达形式分

文字信息：指用文字（符号、代码）来表达其内容的信息，如图书、期刊、专利、科技报告、书信等。

声像信息：指通过声频或视频信号荷载和传递的信息，如通过电视、广播、录音、录像、电话、讲演、谈话所传递的信息。

实物信息：指通过实物荷载、传递的信息，如样品、样机等。

机读信息：指用只有通过机器（主要是电子计算机）才能读取的方式来表达的信息，如用磁带、磁盘、光盘存储的信息。

1.1.4 信息的作用

没有物质就没有我们生活的世界；没有能量，也没有我们生活的世界；但若只有物质和能量而没有信息，世界将是混浊、无序、无组织、无生命力的。信息、物质与能量，是构成客观世界的三大要素，缺一不可。信息对于人类社会的基本作用是增强世界的有序性。信息对于人类社会的生存与发展具有十分重要的作用，主要表现在以下四个方面。

(1) 信息是人类社会生存的条件

人类的活动并不是孤立的个体活动，通常表现为以个人活动为基础的社会性。人类活动的社会性赖以形成、维持和发展的根本保证就是人与人之间能够进行有效的信息交流。人类的任何形式的组织——具有一定目的的有组织的人类活动群体，都必定存

在着一定的信息交流方式以及按此方式交流的信息流。没有人與人之间的有效的信息交流，就不可能构成人类社会。所以，信息是人类社会的粘合剂。

社会是各种各样组织（系统、机构、社团）的有序集合，即人类社会具有有序性、组织性。这种有序和组织是基于信息交流之上的。组织内部的有效的信息交流越多，其组织的有序程度就越高，反之亦然。组织的形成、完善和运行，离不开对组织的管理。从根本上讲，对任何系统（组织）的管理过程都是一种信息管理过程，任何管理系统都可看作是一个包括信息输入、处理、测量、控制、输出、反馈多个组成部分（子系统）的信息管理系统。所以，信息和信息交流是社会组织存在的条件，是管理的基础。

从人的社会行为与行为目的之间的关系来看，信息是人们行为（行动）的依据，组织决策的依据。毫无例外，人（组织）的社会行为都是在他作出一定的决策后进行的。所谓决策，是指人们在改造世界、管理社会（组织）和自身活动中，为达到一定目的的选择活动。在当代社会，由于人类活动的社会性日益增强，人们的行为往往受到多方面的制约和影响，同时他的行为也将影响和涉及到社会的许多方面，所以，决策的重要性空前突出。全面、准确、及时的信息，最大限度地消除人们面临的不确定性，是作出切合实际、可以达到预期目的的正确决策的基础。特别是那些投入大、涉及面广、影响深远的重大决策，更依赖于信息的全面、准确、及时的获得，依赖于现代信息系统的支持。

N. 维纳在他的《人当作人来使用》一书中指出：“要有效地生活，就要有足够的信息。所以，正象通讯和控制属于人的社会生活那样，它们也是人们的内部生活的要素。”

所以，信息是人类社会生存的条件、管理的基础、行为（决策）的依据。

（2）信息是人类认识世界的媒介

信息是物质的普遍属性。信息是事物表象、本质、特征和运

动规律的反映。信息是客观事物与认识主体的中介。信息对于作为认识主体的人类的基本作用，就是减少、消除人们认识上的不确定性。不同的事物有不同的信息，同一事物在不同的运动状态下也会产生不同的信息。人类就是在接收、感知信息的基础上区别事物的差异，从而认识事物的。如果没有信息作为认识主体的人和客观事物的中介，人和客观事物之间的关系只能是物与物的关系，而不是认识与被认识的关系。所以，信息的存在是人类发挥认识能力的必要条件。从进化论的角度看，信息促进了人的认识能力的提高，信息促进了作为物种的人的进化。

信息是人认识客观事物的媒介，是思维的素材。人在接收、感知客观事物信息的基础上，由大脑进行思维活动——即对信息进行加工处理，从而认识事物的本质、特征和运动规律，指导自己的行动。思维与行动的结果，又会产生新的信息。它又成为自己的或他人的信息材料，再经过自己的或他人的思维功能的变换、加工，并导致某种行动，于是又会产生出更新的信息。

(3) 信息是重要的、活跃的生产力要素

决定社会发展水平，推动社会进步的决定性因素是社会生产力水平和它的发展。经典理论表明，社会生产力主要包含劳动者、劳动工具和劳动对象三个基本要素。在这三大要素中，劳动工具是最积极、最活跃的因素。正是劳动工具的不断进步推动着社会生产力的发展。

从历史发展过程中可以看出，信息技术的进步有力地促进了劳动工具的进步，从而极大地推动了社会生产力的提高。

在原始社会和农业-手工业社会，一方面劳动者依靠自身的体能器官或使用简单的人力、畜力工具向劳动对象施加作用，另一方面，劳动者又只能凭自身的感觉器官去获得劳动对象的各种信息。这两方面的情况，决定了当时社会生产力水平的低下。18世纪以来，劳动工具有了重大的进步。一方面由于蒸汽机的发明与电的发现及应用，出现了新型动力工具，使人的行动器官和体力