

国家自然科学基金项目研究成果

# 信息距离与信息

王浣尘 著

DIT and Information



科学出版社

[www.sciencep.com](http://www.sciencep.com)

国家自然科学基金项目研究成果

# 信息距离与 信息

王浣尘 著

DIT and  
Information

 科学出版社  
[www.sciencep.com](http://www.sciencep.com)

## 内 容 简 介

本书提出信息距离(DIT)的概念,将信息与信息距离在概念上加以区分,从而进一步探索并提出信息与知识定量测度的思路和方法体系,并以应用案例加以说明;提出了完全信息量、实现度、目标状态信息量、信息转移代价、信息背景框架(BF),以及知识背景框架(KBF)、知识量、知识密度、知识关联度、知识价值观等概念、定量测度思路、计算方法及其部分的应用举例,有利于知识、管理与控制等方面在运作中的实际应用。本书还指出了平均信息距离、单位信息代价与 Shannon 信息熵三者的等效性;同时探讨了信息量的守恒性问题。

本书章节安排采用“B-A-B”结构,即前后各 3 章较为侧重于应用性,中间共 7 章较为侧重于理论性。读者可以根据各自的兴趣和需要有所选择地阅读。

本书适用于高等院校师生以及各有关行业的研究工作者和实践工作者参阅。

### 图书在版编目(CIP)数据

信息距离与信息/王浣尘著. —北京:科学出版社,2006

ISBN 7-03-017109-8

I. 信… II. 王… III. 信息学-研究 IV. G201

中国版本图书馆 CIP 数据核字(2006)第 030153 号

责任编辑:陈 亮/文稿编辑:王玉敏/责任校对:张 琪

责任印制:张克忠/封面设计:耕者设计工作室

科学出版社 出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

中国科学院印刷厂印刷

科学出版社发行 各地新华书店经销

\*

2006 年 6 月第 一 版 开本:B5(720×1000)

2006 年 6 月第一次印刷 印张:11 1/2

印数:1—2 500 字数:216 000

定价:30.00 元

(如有印装质量问题,我社负责调换〈科印〉)

# 前 言

在 20 世纪 50 年代，作者刚刚开始学习控制论与信息论的时候，似懂非懂。后来正所谓“学而时习之，不亦说乎”，不断地反复学习，兴趣与日俱增，不过对于信息的定量测度问题则总是存在着困惑。信息与知识的运作、管理方便性的直接定量以及控制操作的傻瓜化倾向等，都不断地呼唤着信息与知识的定量测度。1996 年经合组织（Organization for Economic Co-operation and Development，简称 OECD）关于知识经济的专题报告指出：“知识经济本质上最大的困难乃是知识自身的定量及其定价问题。”这指明了现有的 Shannon 信息定量方法对知识定量应用之困难，这也就更加激起了作者进一步探索信息与知识定量测度的积极性。

在 20 世纪 90 年代后期，作者承担的多项国家自然科学基金资助的研究项目都同信息与知识定量测度有关。通过“枚系统经济学及其有关可持续发展的基础研究”（79870052）项目资助研究，明确了人类财富的根本源泉只有两条，一是大自然的恩赐，二是信息的有效增殖；并探索了信息与知识定量可能的多种思路，同时也认同了 Shannon 信息定量方法的伟大贡献及其对知识应用之不足。通过“支持宏观经济决策的人机结合综合集成体系研究”（79990580）重大项目资助研究，涉及学习、创新、共识、管理、综合集成等信息与知识的运作；通过“网格化管理的理论及其应用前景研究”（70450001）项目与“电子政务管理理论与方法的基础研究”（70533030）重点项目以及其他有关信息化项目等的资助，探索信息与知识定量测度的方法及其应用的可能性，逐步有所推进。对信息与知识定量测度多种可能的思路加以比较选择与多次反复，螺旋式推进，目前选择并

认定了一种思路,提出信息距离(DIT)的概念,将信息与信息距离在概念上加以区分,从而进一步探索并提出信息与知识定量测度的思路和方法体系。

在取得点滴进展的基础上,作者于2004年以“信息距离与管理”为主题在国内外分别发表了论文,得到学术界的首肯与鼓励。在上述论文发表和本书撰写的前后,很多前辈、同行和博士生对于“DIT理论”都很热情地参与切磋、研讨、建议和鼓励,作者得益匪浅,深为感激。科学出版社及陈亮编辑的热情支持和鼎力相助,使得本书得以及早出版,特此表示由衷的感谢!

经过这一轮的研究实践,深感在科学技术领域中探索以冀有所细微的推进而能有所自主创新,决非易事。按照科学家钱学森的观点,科学技术上通哲学,下达实践,中分基础科学、技术科学与工程技术,总共涉及5个层次。作者据此体会,从抽象程度来分,一门科学技术将会涉及哲学、理论、方法、技术与实践5个层级,细分之,可有概念、理念、哲理、理论、方法论、方法、技术、工具、措施与实务10个层级。因此,如果冀望能够有所创新甚或有所点滴推进,并能获得成功的话,那就必须实现多个层级的贯通,必须理论与实践互动。理论必须满足科学性的要求,实践必须满足实用性的要求。科学性要求有三:定性上合逻辑,定量上可计算,运作上可重复;实用性要求有一:实践中可应用。不过,这些还只是满足了必要性,还未必在事实上能够获得成功。如果要在事实上真正获得成功,还需要满足更多的条件才有可能构成充分性。这些条件至少有二:学术界能认可,社会上愿接受。满足如此众多的要求和条件,实乃困难重重。凡是在研究或实践的道路上品尝过种种甜酸苦辣的人们或许都会有同感吧!

在研究或实践的道路上需要披荆斩棘,决非单枪匹马所能奏效。有鉴于此,这在宏观上需要有一个切磋、研讨、批评而又较为宽容的环境,而在微观上则需要具有谦虚而又勇敢的素质,继承前人的成果,依托巨人的肩膀,允能有所前进。作者有意参与信息与知识定量测度的探索行列经受磨难,至今抛砖引玉,毕竟一叶障目、井底观天,错误与不妥之处在所难免,恳望不吝赐教指正!

# 目 录

## 前言

### 第1章

|                          |    |
|--------------------------|----|
| 绪论 .....                 | 1  |
| 1.1 信息定量进一步表述的迫切需要 ..... | 1  |
| 1.2 信息概念 .....           | 4  |
| 1.3 信息定量的现有若干表述 .....    | 9  |
| 1.4 关于信息定量测度的思考 .....    | 14 |

### 第2章

|                                    |    |
|------------------------------------|----|
| 信息定量进一步探索的困惑与思路 .....              | 16 |
| 2.1 信息定量进一步探索的涉及方面与现有测度方法的思路 ..... | 16 |
| 2.2 Shannon 信息测度在进一步探索中的困惑 .....   | 17 |
| 2.3 信息定量进一步探索中的关键点 .....           | 19 |
| 2.4 信息定量进一步探索思路的基本特征 .....         | 21 |
| 2.5 信息定量进一步探索思路的框架 .....           | 21 |

### 第3章

|                   |    |
|-------------------|----|
| 信息距离的概念 .....     | 23 |
| 3.1 信息状态的概念 ..... | 23 |

|     |                     |    |
|-----|---------------------|----|
| 3.2 | 信息状态转移的概念           | 25 |
| 3.3 | 信息状态转移的一种定量表述——信息距离 | 26 |
| 3.4 | 举例讨论                | 28 |

## 第4章

|  |           |    |
|--|-----------|----|
|  | 信息距离的计算方法 | 30 |
|--|-----------|----|

|     |               |    |
|-----|---------------|----|
| 4.1 | 全网信息结构与信息距离   | 30 |
| 4.2 | 单向环信息结构与信息距离  | 34 |
| 4.3 | 双向环信息结构与信息距离  | 37 |
| 4.4 | 特殊形式信息结构与信息距离 | 40 |
| 4.5 | 组合形式信息结构与信息距离 | 42 |

## 第5章

|  |              |    |
|--|--------------|----|
|  | 信息距离的组合性与可加性 | 45 |
|--|--------------|----|

|     |                           |    |
|-----|---------------------------|----|
| 5.1 | 串联形式的信息结构 (总的“N选1问题”)     | 45 |
| 5.2 | 并联形式的信息结构 (总的“N选多问题”)     | 48 |
| 5.3 | 并联形式的信息结构 (多路径总的“N选1问题”)  | 52 |
| 5.4 | 串联形式的信息结构 (反复路径总的“N选1问题”) | 54 |

## 第6章

|  |         |    |
|--|---------|----|
|  | 加权的信息距离 | 56 |
|--|---------|----|

|     |             |    |
|-----|-------------|----|
| 6.1 | 加权信息距离的一般含义 | 56 |
| 6.2 | 时间加权与信息距离   | 57 |
| 6.3 | 公共照明的日常管理举例 | 59 |

## 第7章

|  |                |    |
|--|----------------|----|
|  | 信息状态转移概率与信息实现度 | 62 |
|--|----------------|----|

|     |               |    |
|-----|---------------|----|
| 7.1 | 概率的含义和定义      | 62 |
| 7.2 | 实现度           | 65 |
| 7.3 | 状态概率与实现度      | 68 |
| 7.4 | 状态转移概率与实现度    | 70 |
| 7.5 | 实现度转移关系中的若干特性 | 71 |
| 7.6 | 实现度转移方程的迭代计算  | 73 |
| 7.7 | 实现度转移方程的稳定性   | 77 |

|               |                                                |
|---------------|------------------------------------------------|
| <b>第 8 章</b>  |                                                |
|               | <b>状态转移概率与状态概率在实现度上的等效性</b> ..... 80           |
| 8.1           | 等效性概念及其表述 ..... 80                             |
| 8.2           | 状态转移概率与状态概率等效性的一般表述 ..... 81                   |
| 8.3           | 状态转移概率同状态概率在实现度等效时的一般定量关系 ..... 82             |
| 8.4           | 状态转移概率同状态概率在实现度恒等效时的定量关系——与原先实现度无关的情况 ..... 82 |
| 8.5           | 状态转移概率同状态概率在实现度守恒等效时的定量关系——实现度守恒的情况 ..... 85   |
| <b>第 9 章</b>  |                                                |
|               | <b>完全信息量与目标状态信息量</b> ..... 88                  |
| 9.1           | 完全信息量 ..... 88                                 |
| 9.2           | 目标状态信息量 ..... 90                               |
| 9.3           | 信息状态转移的信息代价 ..... 92                           |
| 9.4           | 平均信息距离、单位信息代价与 Shannon 信息熵 ..... 95            |
| 9.5           | 完全信息量与信息不变量 (信息量守恒) ..... 97                   |
| 9.6           | 创新在信息量上的反映 (信息量可能不守恒) ..... 98                 |
| 9.7           | 完全信息量的层次性 ..... 101                            |
| 9.8           | 信息传递与信息状态背景框架 ..... 104                        |
| <b>第 10 章</b> |                                                |
|               | <b>知识定量测度的探索</b> ..... 111                     |
| 10.1          | 知识与信息 ..... 111                                |
| 10.2          | 知识的背景框架、关联度与关联强度 ..... 113                     |
| 10.3          | 知识的定量测度 ..... 118                              |
| 10.4          | 知识的密度 ..... 119                                |
| 10.5          | 知识的价值观 ..... 123                               |
| 10.6          | 小结 ..... 124                                   |
| <b>第 11 章</b> |                                                |
|               | <b>信息距离与信息在知识运作方面应用的研讨</b> ..... 126           |
| 11.1          | 语句的知识量 ..... 126                               |

|      |                                         |     |
|------|-----------------------------------------|-----|
| 11.2 | 语句复制的信息距离·····                          | 129 |
| 11.3 | 语句撰写的信息距离·····                          | 130 |
| 11.4 | 寻找语句所需的信息距离·····                        | 131 |
| 11.5 | DIT 理论在优化算法定量评价中的应用——以有序表的查找为例<br>····· | 132 |
| 11.6 | 知识创新与学习·····                            | 136 |

## 第 12 章

|      |                           |     |
|------|---------------------------|-----|
|      | 信息距离与信息在管理运作方面应用的研讨 ····· | 144 |
| 12.1 | 系统协调的信息距离·····            | 144 |
| 12.2 | 医疗服务的信息距离·····            | 146 |
| 12.3 | 密码运作的信息距离·····            | 148 |
| 12.4 | 金融理财电话服务的信息距离·····        | 153 |
| 12.5 | 网格化管理的信息距离·····           | 155 |
| 12.6 | 电子政务中的信息距离分析与建议·····      | 157 |

## 第 13 章

|      |                           |     |
|------|---------------------------|-----|
|      | 信息距离与信息在控制运作方面应用的研讨 ····· | 160 |
| 13.1 | 电子邮件系统登录过程的信息距离·····      | 160 |
| 13.2 | 在表格中光标移动控制的信息距离·····      | 164 |
| 13.3 | 烧煮过程旋进式控制的信息距离·····       | 166 |
| 13.4 | 代码之间的信息距离·····            | 169 |
| 13.5 | 文字之间的信息距离·····            | 171 |
| 13.6 | 照相机操作的信息距离·····           | 173 |

|           |     |
|-----------|-----|
| 参考文献····· | 175 |
|-----------|-----|

# 第 1 章

## 绪 论

### 1.1 信息定量进一步表述的迫切需要

涉及信息的定量表述，其研究探索已经超过半个多世纪了，可是实际上还存在着很多问题需要认真解决。

#### 1.1.1 时代的呼唤

人类发展至今，对客观世界的认识不断深化，对客观世界的运作和扰动也越来越强大。到了 1992 年，人类经过漫长的发展历程，在认识上产生了一个质的飞跃，其标志是 1992 年在巴西里约热内卢召开的联合国环境与发展大会上，确立了全球可持续发展（Sustainable Development）的观念和战略。该大会有 183 个国家代表团和 70 个国际组织代表参加，102 位国家元首或政府首脑讲话，签字通过了一系列纲领性文件，诸如《里约热内卢环境与发展宣言》、《21 世纪议程》、《关于森林问题的原则声明》、《气候变化框架性公约》以及《生物多样性公约》等。人类的行动从自发到自觉，从此又开启了新的一页<sup>[1,2]</sup>。

可持续发展目前在国际上比较一致的定义是：“既能满足当代人的需求，又不损害后代人满足其需求之能力的一种发展。”可持续发展的基本精神可用“1145”加以概括，即一个本，一个中心，四句话，五位一体：以人为本；以经济发展为中心；经济要讲究效率，社会要强调公平，代际要重视兼顾，人与自然要力争和谐；努力实现人口、经济、社会、资源、环境五位一体的协同。显然，

可持续发展战略的贯彻实施的确有着很高的难度。中国当前特别强调科学发展观,全面、协调、可持续,五个统筹,高度重视自主创新,以及建设和谐社会等,正以此引领着全国上下努力实践,奋勇前进。

人类社会历经原始工具时代、农业时代、工业时代,逐步发展到信息时代、知识经济时代;从原始经济起,有农业、工业、服务业、信息业、创意产业等不同的产业群。特别有意思的是,从长江上游到下游入海口正巧体现出人类社会所经历的四个发展阶段特征,即原始社会、农业社会、工业社会到知识社会,被称为人类文明发展长河的“长江模型”<sup>[3]</sup>。上述社会发展历程正反映着可持续发展的精神,其相应产业结构的提升和循环经济的强化正符合可持续发展的要求。

按照“枚系统经济学”<sup>[2]</sup>的观点,物质存在及其运动形式具有三要素,即物料(material)、能量(energy)和信息(information)。人类的活动离不开物料、能量、信息三要素,人类财富的获得当然也离不开“物、能、息(MEI)”三要素。经过分析可以得出一个重要结论,即人类财富的根本源泉:一是大自然的恩赐;另一是信息的有效增殖,即复制、创新和涌现。这里,复制是指广义复制,创新是指发现发明创造的统称,涌现是指系统中的“ $1+1>2$ ”的新质涌现,其中,复杂性系统与系统集成系统中将会有更多的涌现。随着人类活动的不断增强,人类作为一个整体已经意识到,必须“善待大自然的恩赐,崇尚信息的有效增殖”。大自然的恩赐是有一定限度的,因而除了大力推进可持续发展和循环经济之外,人类为了获得更多、更大的财富,必须在宏观上和微观上同时大力推进信息化。

目前,信息和知识在人类活动中的重要性已经日益彰显。Gartner 于 2005 年在关于“信息爆炸”的预测中指出:“企业信息管理(Enterprise Information Management,简称 EIM)的需求正在递增,因为企业需要进行流程简化、提升营运效率及灵敏度,力求在竞争对手中脱颖而出。Gartner 预料到 2015 年,企业将要处理比 2005 年多 30 倍的数据量,因此单看企业是否具备强劲的数据管理、处理元数据(Metadata)及分析信息能力,便立见高下。”海量的信息与知识、海量的分析与处理,与日俱增<sup>[28]</sup>。在产业结构提升过程中,关于信息、知识、创新、智慧、智能等,不无时延地波及消费、生活等领域,从而波及人类社会活动的所有领域,并已经得到全社会的高度重视,甚至可以说,已经形成了时代一股热烈的强大的潮流。但是,有关的定量问题还没有得到满意的解决。在 1996 年,经合组织(Organization for Economic Co-operation and Development,简称 OECD)在有关知识经济的专题报告中明确地指出,“知识经济本质上最大的困难乃是知识自身的定量及其定价问题”<sup>[4]</sup>。

“知识如何定量?”的确是一个难题。目前,用字数、语句数或页数来衡量,那么,废话连篇怎么办?言简意赅怎么看?一言千金又怎么说?

“学习和创新怎样定量?”又是一个更大的难题。目前,在学术研究领域往往用 R&D (research and development) 的经费来间接衡量,或者用时间来衡量。同样一篇论文,“探索创新”出来也许得经历几十年的废寝忘食,“研究学习”可能需要十年寒窗,而“死记硬背”、“囫囵吞枣”则可以立等见效,至于“拷贝复印”只不过弹指一挥间而已。其中,难度的高低又怎么考虑?效率的高低怎么体现?

再有,知识的编辑、整理和排序,知识的搜索、收集和整合,知识的分析、综合和摘要,等等,其运作之繁简又如何给以直接定量呢?

### 1.1.2 现实的需要

1) 在现代化的社会中,人们越来越离不开机电设备和电子器具

这些设备和器具的设计和操作越来越复杂,而人们所需要的恰恰是思路简洁与操作方便 (easy & convenience), 这就是所谓的“傻瓜化”倾向。“傻瓜化”有没有客观一些的标准?符合什么标准可以算是“傻瓜机”?简便不简便,如何定量?很难。目前,一般的办法就是用打分的办法,比如采用 100 分制,或者 1、3、5、7、9 的等级分制。是不是偏主观了呢?能不能直接定量?能不能找到一种测度,像千克、千米等一样,可以对“简便”给出较为客观的直接定量?那就更难了。

2) 在现代社会中,服务业越来越发达

顾客希望能够很方便地得到服务。有的服务热线电话,用户将电话打进去,先强求用户听上一段音乐,再一次地按数字键,进了一层又一层,一不小心还得从头来。方便不方便,如何定量?而直接定量又如何操作呢?

3) 在现代网络社会中,涌现出大量的各式各样的网页,尤其是电子政务的各级网站

用户需要搜寻各式各样的资料,寻求各式各样的服务,有时操作非常繁琐,进了一个端口,不行,退出,再寻找另外的端口,不胜其烦。电脑屏幕上的操作也有同样的情况。诸如此类的操作方便与否,需要定量,也更需要客观一些的直接定量。

### 1.1.3 工程与巨型工程管理的需要

在现代管理领域中,特别是涉及到工程管理,尤其是钱学森提到的“巨型工程”<sup>[26]</sup>,需要强调“以人为本”、“以服务为本”。在现代社会中,有的工程项目规模越来越大,内容越来越多,涉及面越来越宽,要求越来越高,对社会的影响也越来越深远。如何能够管理好这样的“巨型工程”,必须科学地管理,除了宏观上必须深思熟虑之外,在微观上必须一丝不苟,正所谓“细节决定成败”。比

如,美国的阿波罗登月,中国的长江三峡工程,中国的神舟飞船,等等,都是如此。其中,一般要涉及到三种“人”,即用户、管理者和相对的高层领导<sup>[5]</sup>。这三个层次和角度必须统一考虑,一环扣一环。各个层次、各个岗位,都必须实现三种基本要求:①服务到位;②管理高效;③监管有效。

这样,在实现基本要求的基础上,经过优化,用户得到服务需要方便,管理者运作需要简便,领导层监管需要简捷。对这些要求的定量,有难度;较为客观一些的直接定量,其难度就更高。

#### 1.1.4 学术上的需要

众所周知,一门学科得以自立于科学之林,必须具有逻辑性、可重复实现性和可定量表述性。系统科学与系统工程历来强调定性与定量相结合。定性与定量两者缺一不可。两者缺其一,其相应的“科学”就是一门不完整的科学。

目前,物料的测度可以公斤(kg)计;能量的测度可以焦耳(J)计;几何距离的测度可以公尺(m)计;时间的测度可以秒(s)计;信息的测度一般可以比特(bit)计<sup>[2]</sup>。但是,涉及到知识的学习和创新、管理运作的方便、设备器具操作的简便,如果缺少直接定量的测度和方法,总是一种缺陷和遗憾。这种缺陷和遗憾必须予以破解。

## 1.2 信息概念

信息(information)在中文中又有被称为“资讯”的。

诸多文献都提到信息(information)作为一个科学词汇没有得到公认统一的定义<sup>[6~19]</sup>。

在 Webster 字典<sup>[19]</sup>中,关于 information 的解释为:“n (14c) 1: the communication or reception of knowledge or intelligence; 2 a (1) : knowledge obtained from investigation, study, or instruction; (2): INTELLIGENCE, NEWS; (3): FACTS, DATA ; b: the attribute inherent in and communicated by one of two or more alternative sequences or arrangements of something (as nucleotides in DNA or binary digits in a computer program) that produce specific effects; c (1): a signal or character (as in a communication system or computer) representating data; (2): something (as a message, experimental data, or a picture) which justifies change in a construct (as a plan or theory) that represents physical or mental experience or another construct; d: a quantitative measure of the content of information; specify: a numerical quantity that measures the uncertainty in the out-

come of an experiment to be performed...”;

而关于 information science 的解释为: “n (1960): the collection, classification, storage, retrieval, and dissemination of recorded knowledge treated both as a pure and as an applied science”;

至于对 information theory 的解释为: “n (1950): a theory that deals statistically with information, the measurement of its content in terms of its distinguishing essential characteristics or by the number of alternatives from which it makes a choice possible, and with the efficiency of processes of communication between humans and machines (as in telecommunication or in computing machines)”。

从该字典的解释可以看出,在英文里作为一般词汇,“信息”(information)的出现是比较早的;而作为科学词汇“信息科学”(information science)或“信息论”(information theory)的出现则是20世纪中后期的事情了。

在中文里,作为一般词汇的“信息”一词,在中国久已有之;而作为科学词汇含义的同样是在“信息论”出现之后的事情。《辞海》<sup>[6]</sup>所载:“(1)音讯;消息。李中《暮春怀古人》诗:‘梦断美人沉信息,目穿长路倚楼台。’(2)通信系统传输和处理的对象,泛指消息和信号的具体内容和意义。通常须通过处理和分析来提取。信息和物质、能量被称为构成系统的三大要素。信息的量值与信息的随机性有关,如在接收端无法预估消息或信号中所蕴含的内容或意义,即预估的可能性越小,信息量就越大。”其中,第一点为“信息”在中文中所理解的一般含义;第二点为“信息”在通信学科中所理解的一般含义。

《系统科学大辞典》<sup>[7]</sup>中由钟义信执笔的“信息”词条,对信息的科学含义做了较为全面的概括:

“一类研究对象的总称,在概念上与材料(狭义的物质)及能量相并列。”

“历史上,信息概念存在许多不同的表述。例如,控制论的创始人 N. Wiener 指出‘信息就是信息,既不是物质也不是能量’(1948);信息论的奠基者 C. E. Shannon 认为‘信息是用来消除随机不定性的东西’(1948);……。众多的信息定义令人眼花缭乱,扑朔迷离。”

“深入的研究表明,按照约束条件的不同,信息的概念形成了自己的层次体系:没有任何约束条件的信息概念是最广义的信息概念,即本体论层次的信息概念,它被定义为‘事物运动的状态和(状态变化的)方式’;如果进一步引入新的约束条件,限定主体必须有观察力、理解力和价值判断力,就会得到认识论层次的全信息,它被定义为‘主体所感知所表述的事物运动状态/方式的形式(称为语法信息)、含义(称为语义信息)和效用(称为语用信息)’;引入的条件越多,相应的信息定义的层次越低,适用的范围越窄。……。例如, Wiener……

的定义属于本体论层次, Shannon 的定义则属于认识论层次, 等等。”

“按照本体论和认识论层次信息的定义不难理解, 信息是一切知识的‘原材料’, 而知识则是人类头脑对信息进行加工的产物, 是一类新的具有抽象概括品格和具有相对普适相对稳定特性的信息。在这个意义上可以说: 任何知识都是信息, 然而并非所有信息都是知识。知识是信息的一个特殊的子集。类似地, 情报是一类具有特定价值功能的信息, 因此是信息的一个特殊的子集。”

“利用全信息, 人类头脑不仅可以从中加工出知识, 而且可以加工出‘如何使用这些知识’的知识。也就是说, 人类可以从全信息中加工出知识和使用这些知识来达到某种目的的能力。这种能力就是所谓的‘智慧’或‘智能’。”

“……; 只有综合利用物质、能量和信息资源才可以制造‘聪明的’生产工具, 它既不需要人力的驱动, 也可以在很大程度上不需要人的驾驭, 从而可以帮助人类完成许多复杂的劳动和工作, 促进人类从自然力的束缚下获得越来越深刻的解放。这是研究信息问题的意义之所在。”

2005 年, 徐飞与高隆昌在一本专著<sup>[8]</sup>中强调: “信息的本原性定义: 信息是脱离了载体的属性”; 并且说明“信息的初始概念最早于 1922 年由卡松提出, ……虽然提出了‘信息’概念, 仍与属性一样, 无法使用。但在 1928 年, 哈特莱强调了‘信息是脱离了载体的属性’, 或说‘信息是脱离了载体的一个新的事物’。”又进一步强调: “……, 属性必不能脱离载体, 一旦脱离就不再叫属性。……它就是‘信息’, 是个新事物。较之相应的原属性, 它可能失真, 可能被处理、加工、改变。这就是它的定义的含义。”“……正是哈特莱(1928 年) 给予的升华性定义。它既提高了后来人的信息思维质量, 又为后来的信息技术发展提供了思想基础。”并且还指出香农(Shannon) 正是在这一思想基础上于 1948 年奠定了信息的数理基础的工作。

综上所述, 作者建议选用信息的定义为: 信息是事物运动状态之表征。简而言之, 信息是事物状态之表征。

这个定义需要加以简单的说明:

(1) 事物, 泛指世界上一切的“物料、能量、事、人、自然信息……”, 表明信息所依据的客观性。对此更为泛义的理解是, 事物可以泛指为“物质”。

(2) 运动, 泛指事物所有的“存在、变化、演化”, 表明信息所涉及的可以是静态的, 也可以是动态的。如此理解, 是同当今对“运动”公认的理解相一致的, 即运动是“物质的存在形式及其固有属性。它包括宇宙中所发生的一切变化和过程, 从简单的位置变动到复杂的思维活动。运动同物质不可分离。运动和物质一样, 是不可创造和不可消灭的。它只能由一种形态转化为另一种形态。运动是绝对的, 静止是相对的。静止是物质运动的特殊形式”<sup>[6]</sup>。

(3) 状态, 特指事物运动所有属性中所需关注的属性。

(4) 表征，是对特定“标的”（状态）的一种具有某种“代表性”的表述。在这里具体地说，“表征”的含义所指的是，信息是状态（所关注属性）的反映、映射或复制，而不是事物运动状态（所关注属性）的本身。

(5) 从状态（所关注属性）转化为信息的过程由于种种内在的、外在的原因或扰动，因此，既可能逼真，也可能失真，可能经受过某种“故意的、主动的或被动的”处理、加工、变换、扭曲、扰动等。由信息反映状态（所关注属性）：可能很正确，很精确；也可能失真，可能不正确，不精确，从而只能保证某种程度的准确；甚至可能有“真”与“假”之分。

(6) 信息，一旦不再等同于原来所特指事物运动状态（所关注属性）的本身，当然也更不等于原来所特指事物的本身，那么，信息本身也就得以相对独立地存在，并转化成为新一类事物中的一员。因而，从概念上来说，就可以有“信息之信息”，即“多重信息”的出现。

(7) 信息必须选择某一事物作为“载体”。

这样的一些说明，可以利用一个示意图近似地表述，如图 1-1 所示。

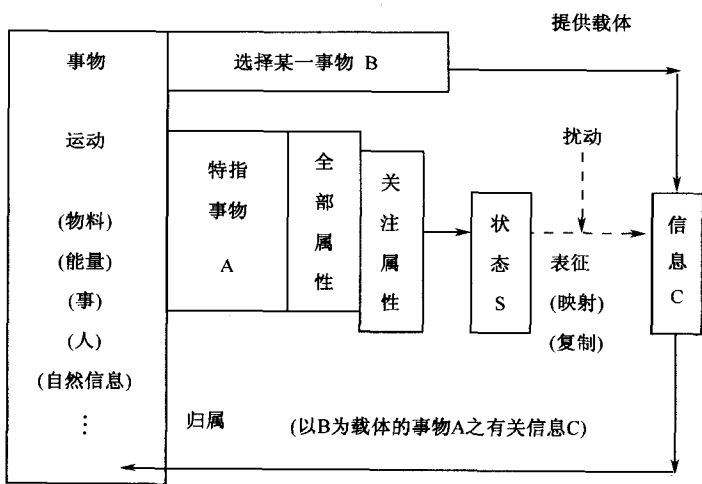


图 1-1 事物、属性、状态、信息等之间关系近似表述的示意图

对于图 1-1，还需要做进一步的说明：

(1) 以 B 为载体事物 A 之有关信息 C，已经相对独立于所特指的事物 A，从而“B+C”形成为一个新事物，并归属于一般性的事物的大类之中。

(2) “B+C”这个新事物中，“载体 B”原本就属于广义的事物的大类之中，而“信息 C”则是新增加的一个新事物。这个“信息 C”在一定的条件下又可以生成一个新“状态 S2”，从而又可以转化生成成为一个新“信息 C2”。这个“信

息 C2”就是“信息 C 的信息 C2”。这就是所谓的“信息之信息”、“2 次信息”、或“多重信息”的现象，如图 1-2 所示。

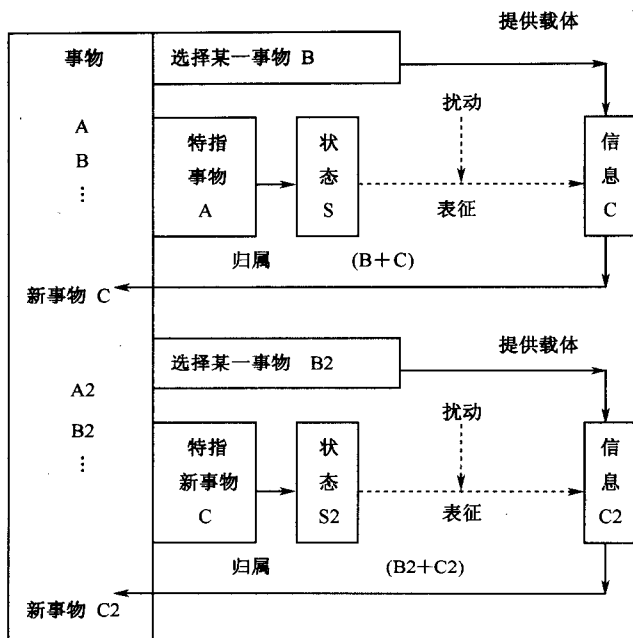


图 1-2 “信息之信息”的示意图

(3) 物料、能量、事、人等作为信息的载体，一般是容易被理解的。至于某一个“信息 1”能不能作为另一个“信息 2”的载体呢？其答案也是肯定的。比如，在通信技术中，一种低频信号（对应于信息 2）可以利用高频信号（对应于信息 1）作为载体，可有调幅、调频、调相等多种不同的调制制式；其中高频信号本身就带有一种信息，可以用来作为“信号水平自动调整”的依据。又如，在保密通信技术中，一篇公开的文章（对应于信息 1）作为载体，而一份保密的情报（对应于信息 2）可以通过一定的规则寄生于上述载体（对应于信息 1）之上。

(4) “人”作为广义事物运动之中的一部分，是一个“特指的事物”。这是一个非常复杂的对象。从其属性角度来分析，可以有自然人、生物人、物理人、化学人、个人、家庭人、家族人、社区人、单位人、社会人、文化人、经济人、政治人、军人等；从其身上可以从各种角度提取出各式各样的属性、状态、信息；从其提取信息之后所依托的载体来看，可以有照片、录像、拟人机器人、数字人、虚拟人、简历等；同时，“人”还可以作为其他各式各样信息的载体，比如，