论 文

客观信息的模型和度量研究

许建峰^{①*}, 汤俊^②, 马雪峰^①, 徐斌^①, 沈艳丽^①, 乔永杰^①^① 中国电子科学研究院, 北京 100041^② 清华大学电子工程系, 北京 100084

* 通信作者. E-mail: xjfcetc@163.com

收稿日期: 2014-07-25; 接受日期: 2014-10-16; 网络出版日期: 2015-02-11

摘要 当今时代, 信息的重要性毋庸置疑. 但是, 人们对于信息的认识依然众说纷纭, 对于信息度量的研究同样莫衷一是. 本文基于前人的丰富研究成果, 总结了在信息的性质及度量方面已经取得的重要进展和存在的主要不足, 探讨了信息的本质内涵和数学表达等基本问题. 基于对信息客观性的认识, 提出了信息的定义和六元组模型, 讨论了信息的基本性质, 建立了信息的广阔性、细致性、持续性、丰富性、包容性、延迟性、遍及性、真实性和适配性 9 类度量的定义和数学表达, 形成了客观信息论的基础理论框架, 能够系统全面地支持信息与信息系统的分析研究.

关键词 信息 信息熵 客观信息论 六元组模型 度量

1 引言

计算机和通信技术的迅速发展使人类进入信息化时代, 无处不在的信息系统使人类的生产生活方式发生了巨大变革. 但是, 人们对于信息的认识依然众说纷纭, 对于信息度量的研究同样莫衷一是. 究其本质, 不在于人们没有倾注大量的兴趣和精力, 而在于众多研究者对信息研究的对象、目的和方法缺乏提纲挈领和一以贯之的共识. 首先, 关于信息研究的对象, 由于人们心目中的信息既有属于客观范畴, 也有属于主观范畴^[1], 因而很多研究没有区分客观存在的信息与人们主观意识活动之间的关系, 从而使问题变得极为复杂; 其次, 关于信息研究的目的, 很多研究满足于学术意义上的探索和推演, 很大程度上忽视了客观世界中已经或即将深刻改变人类生产生活方式的众多信息系统, 忽视了信息研究对信息系统研究、设计和开发的实际指导作用; 再次, 关于信息研究的方法, 很多研究局限于继承 Shannon 针对通信系统而运用的传统概率统计方法, 忽视了运用更为广泛的数学理论全面系统地建立信息科学理论基础的可行途径.

鉴于此, 本文分析并借鉴了前人的丰富研究成果, 提出了客观信息论的基本框架, 具体包括:

一是明确将信息的本质界定于客观范畴, 既符合信息作为客观世界三大组成要素之一的根本定位, 也适应信息系统能够且只能处理客观实在信息的运作要求;

二是提出了信息的数学表达和六元组模型, 为其后的各项研究提供了统一明晰和便利可行的数学理论基点;

三是基于六元组模型和映射的数学理论, 阐述证明了信息的客观性、还原性、传递性、关联性和组合性等基本性质, 进一步体现了信息及其模型定义的合理性;

引用格式: 许建峰, 汤俊, 马雪峰, 等. 客观信息的模型和度量研究. 中国科学: 信息科学, 2015, 45: 336–353, doi: 10.1360/N112014-00040

四是系统全面地建立了信息的 9 类度量体系, 并从基本模型出发定义了各项度量的数学表达, 初步探讨了各项度量的基本性质, 为定量研究分析信息运转和信息系统应用提供了理论和方法基础.

2 几个关键问题

建立信息的基本理论体系, 必须以信息的本质内涵为源, 从数学模型着手, 形成完整的度量体系和一致的表达方式. 对此, 前人已经做了大量富有成效的研究, 需要分析总结.

2.1 关于信息的本质内涵

信息作为科技术语最早出现在 Hartley^[2] 于 1928 年发表的《信息传输》中. 一般认为, Shannon^[3] 1948 年发表的《通信的数学理论》为现代信息论奠定了理论基础. 其中 Shannon 将信息描述为通信过程中需要传输的服从一定概率分布的随机编码. 由此人类在历史上第一次清晰地认识和把握到通信技术的本质, 促进了通信系统蓬勃发展, 为沟通交流带来了极大便利. 随着信息技术和信息系统的发展, 信息的存在和应用远远超出通信系统范畴, 对于信息的认识日趋丰富和深刻, 以至于“围绕信息定义所出现的流行说法已在百种以上”^[4], 分别将信息类比作消息、信号、数据、情报、知识等概念, 由于各有侧重, 难免执于一词, “所以直到现在, 学术界关于‘信息’的本质远未达成共识”^[5].

众多研究者中, 钟义信^[4] 对信息概念作了最为充分的研究. 他建立的“全信息理论”, 以“事物运动的状态和状态变化方式的自我表述/自我显示”作为本体论层次的信息定义, 同时认为只有在感知了它的形式, 理解了它的含义, 判明了它的价值, 才算真正掌握了该事物的认识论层次信息. 将仅仅计及其中形式因素的部分称为“语法信息”, 将计及其中含义因素的部分称为“语义信息”, 将计及其中效用因素的部分称为“语用信息”, 将同时考虑事物运动状态及其变化方式的外在形式、内在含义和效用价值的认识论层次信息称为“全信息”. “全信息”理论关于信息的语义和语用研究完成了西方信息科学家没有完成的任务^[6], 却也因为求全而面临不少困难. 首先, 未将客观存在且机理可能相对简单的“本体论层次信息”作为研究重点; 其次, 主要关注的“认识论层次信息”与人的主观思维密切相关, 其规律和规则非常复杂, 需要逾越至今仍未很好解决的“语义鸿沟”; 再次, 由于力求面面俱到地将识别论、通信论、控制论、决策论、优化论和智能论等统一在一个完整的理论体系中, 尚有很多错综复杂的问题需要澄清.

Burgin^[7] 基于本体论和价值论两组原则, 建立了一般信息论, 将信息定义为引起变化的能力. 同时引入系统 R 的信息逻辑系统 $IF(R)$ 概念, 指出信息具有相对性, 即相对于系统 R 的信息是只能够引起信息逻辑系统 $IF(R)$ 变化的能力. 一般信息论基于公理化基础建立完整信息理论体系的思想十分可取, 但其主要问题在于: 一是将信息描述为引起变化的能力, 在概念上能量也属于这个范畴, 与人们通常的认识存在很大差距; 二是系统 R 及其信息逻辑系统 $IF(R)$ 既包括客观系统, 也包括主观系统, 仍然难以区分信息与人们的主观认识之间的差别.

Fleissner 和 Hofkirchner^[8] 更提出了统一信息论的概念, 基于对信息从语法到语义和语用状态演化过程的分析, 将信息描述为系统的结构、状态或者行为变化的产物. 统一信息论虽然试图运用一致的概念包容众多各不相同的信息内涵, 但其基本分析方法类似于钟义信的全信息理论, 而理论深度和系统性则远逊于后者. 因此, 甚至连原创者本人也提出了“统一信息理论是否可行”的疑问, 指出了构建统一理论体系的主要障碍^[9].

2.2 关于信息的数学表达

Shannon 在文献 [3] 中虽然没有直接点明信息的数学定义, 但实质上已经体现其考虑的信息就是“一组发生概率分别为 $p_1, p_2, \dots, p_n$ 的编码”, 而熵公式

$$H(X) = H(p_1, p_2, \dots, p_n) = -k \sum_{i=1}^n p_i \log p_i \quad (1)$$

就是反映信息不确定性的度量, 进而才得以确立了通讯系统设计的基本准则.

张继国等^[10] 阐述了信息点的概念, 给出了信息点的数学表达:

$$R = (N, c, \nu), \quad (2)$$

其中, N 表示信息, c 表示信息特征的名称, ν 表示 N 关于 c 所取的度量. 基于这样的有序三元组模型, 形成了信息场、信息群、信息势等研究成果.

张维明^[11] 提出了信息系统的数学模型:

$$S = \langle O, A, V, \rho \rangle, \quad (3)$$

其中, O 为对象的有限集合, A 为对象属性的有限集合, V 为每个属性定义域 V_a 的和集, 即 $V = \cup_{a \in A} V_a$, ρ 为对象属性值函数. 在此基础上, 形成了最大信息系统、最简信息系统、理想信息系统、不确定性信息系统以及信息系统的连接等研究成果.

这些成果, 都尝试运用数学表达或建模的方法研究信息. 其中 Shannon 提出的就是信息的数学表达, 只不过局限于通信系统关注的统计信息, 虽然成果卓著, 也在其他领域得到应用, 但与一般信息的数学内涵相去甚远. 信息点的数学表达与一般信息的概念最为接近, 但其组成参量既与信息的实际定义不相关联, 又未能在进一步的研究中发挥充分作用. 信息系统的数学模型虽不直接针对信息, 却与信息本身的要素构成极为相近, 并且其中的各类参量在后续研究中得到了运用, 这种方法值得借鉴.

笔者认为, 迄今为止信息数学表达研究中最为重要的成果是 Burgin^[12] 在一般信息论中建立的三元组数学模型:

$$\langle C, I, R \rangle, \quad (4)$$

其中, C 表示载体, I 表示信息, R 表示信息系统. 基于此模型运用范畴理论方法, 通过 C, I, R 以及信息逻辑系统 $\text{IF}(R)$ 之间的一系列算子, Burgin 研究了信息发送、信息传输、信息接收和信息存储时 C, I 和 R 三者之间的关系^[12]. 这种方法将信息与其载体和所属系统视为一个完整的体系, 为运用各种数学方法研究信息的本质和运作打开了便利之门, 应该成为信息研究一以贯之的基本方法. 但此三元组模型与一般信息论对信息的定义不具有严格的对应关系; 且其中信息 I 本是一个研究对象, 却并未得到进一步分解; 信息系统 R 应该与信息相对独立, 却成为与 I 密不可分的一个元素; 载体 C 应该是信息的一个要素, 却没有纳入 I 之中; 信息逻辑系统 $\text{IF}(R)$ 是一般信息论定义信息的重要元素, 却没有在三元组中体现. 因此, 一般信息论三元组模型并未充分反映信息所含要素的基本结构, 基于范畴理论的信息算子研究实际上也未与三元组模型充分相关.

2.3 关于信息的度量

Hartley^[2] 最早提出信息传输中度量信源信息量的方法. 其大意是从 S 个彼此不同的符号中取出 n 个符号组成一个字, 如果各个符号出现的概率相同, 且选取是随机的, 就可以得到 S^n 个不同的字.

所有这些字包含的信息量为

$$H = n \log S. \quad (5)$$

Hartley 建立纯粹形式化信息测度方法虽然过于简单, 却对 Shannon 有直接影响. 熵公式 (1) 显然是式 (5) 的演进和拓展. 它作为信息度量, 宣告了信息论的诞生. 其后 Wiener^[13], Ashby^[14], De Luca 和 Termini^[15] 等, 都运用各自的熵公式描述信息度量. 后续的大量研究分别引入了 Hartley 熵和 Shannon 熵的扩展^[16]、相对熵^[17]、累积剩余熵^[18~21]、联合熵^[22,23]、条件熵^[24~26]、互信息^[27~32]、交叉熵^[33~38]、模糊熵^[15,39]、最大熵原理^[40,41]和最小交叉熵原理^[42,43]等概念, 取得了一系列深化研究的成果. 钟义信^[4]则用一般信息函数统一了运用熵公式表征信息度量的方法. 需要强调的是, 这些关于熵的研究主要基于概率统计数学方法.

1945 年, Rao^[44]提出统计流形, 用 Fisher 信息矩阵定义流形上的黎曼 (Riemann) 度量, 并采用测地线距离来度量概率分布函数之间的差异, 该距离也称为 Fisher 信息距离, 与 Shannon 信息熵形成对比, 成为反映信息度量的另一种方式. 通过概率统计与几何方法的有效结合, 开启了信息几何理论的研究^[45]. 其后, Efron^[46]给出了统计流形上的曲率, Chentsov^[47]在统计流形上引入一族仿射联络, Amari 和 Nagaoka^[48~50]提出了对偶仿射联络概念, 通过引入势函数定义散度的概念, 成为一种较弱的距离函数. 近年来, 信息几何在信息理论、系统理论、神经网络、统计推断、通信编码、物理学、医学成像等各个领域逐步得到广泛应用^[51~62]. 信息几何是研究信息度量的有力工具, 被誉为“第二代信息论”^[63]. 但由于发展时间较短, 其理论和应用研究仍处于起步阶段. 并且在数学上几何空间只是测度空间或拓扑空间的一种形式, 信息的各种度量未必完全能够通过其在几何空间的投影来反映, 需要引入更为广泛的数学工具.

Burgin^[64]的一般信息论也研究了信息度量问题. 根据作用对象和改变程度, 分别定义了内部、中间和外部等 3 类结构型信息度量, 又根据不同的获取方式分别定义了抽象、现实和实验等 3 类构造型信息度量. 这些定义未给出具体数学表达方式, 因而只具有概念研究的意义.

显然, 信息的作用不仅限于通信领域, 对于信息的度量也绝不仅限于消除不确定性或产生变化所需的信息量. 因此, 人们还对其他信息度量进行了很多有益的探索.

与信息量较为接近的度量是信息的丰富性. Daft 和 Lengel^[65]在 1986 年首先提出信息丰富度的概念, 又被称为媒介丰富性理论, 将关注焦点从通信拓展到媒介, 从单向信息传输拓展到双向信息交互, 其研究思路与 Shannon 信息论全然不同, 但以信息丰富度衡量媒介价值和能力的方式方法对于建立并运用信息的度量体系具有非常重要的启发.

信息的细致性显然也是需要关注的一个度量. 梁吉业等^[66]针对完备和不完备信息系统定义了信息熵、粗糙熵和知识粒度, 并给出了信息粒度的公理化定义. 其信息粒度实际上就是对信息细化的不同层次的平均度量. 滕书华等^[67]继而研究了熵与信息粒度的关系, 指出信息系统中信息粒度的大小反映了知识对论域分类的聚集度, 即知识对论域分类越粗糙, 信息粒度越大, 反之则信息粒度越小, 使信息细致性度量的含义更加明晰.

信息的价值在很大程度上在于其真实性. 因此, 关于真实性的度量已经受到人们的高度重视, 尤其是在军事领域. Gaziano 和 McGrath^[68]研究了可信度的度量概念, Alberts 等^[69]和兰德公司^[70]分别从军事应用需求研究了信息的正确性或可信性的度量定义及估算方法, Niven^[71]则研究了媒介信息和新闻信息的可信度, 这里的正确性及可信度实际上都是信息与其所反映的客观事物真实状态之间的差异程度, 应该说是信息最为重要的一种度量.

此外, 信息的及时性、普及性以及对用户需求的满足程度等也都是衡量信息价值不可或缺的度量

指标, Alberts 等^[69]、兰德公司^[70] 等对这些度量都分别进行了或深或浅的研究, 为我们深入全面地构建信息的度量体系提供了有益借鉴。

上述研究工作的主要不足是缺乏系统性: 一是没有从信息的客观存在性出发研究其度量; 二是绝大多数研究都只关注于某一种度量方法, 缺乏全面考量; 三是大多数信息度量定义没有给出相应的数学表达, 不便于后续运用; 四是多数信息度量的应用仅限于特定领域和场景, 不具有普遍性。

3 信息的定义及性质

定义信息可能是一个十分艰难的哲学问题。事实上, 关于信息究竟属于客观还是主观范畴就一直是学界争议的焦点。一种具有代表性的观点是“信息表现为客观事物的一种状态和人们主观上对这种状态的定义规则”, 力求澄清信息在客观和主观两个范畴的交互关系及其本质规律, 目前已有众多研究成果, 但仍面临客观与主观问题相互融合而带来的不少疑难之处。本文无意统一所有歧见, 仅仅着眼于化繁为简, 对信息的客观和主观属性进行一定程度的解耦, 并结合信息系统的运作要求, 立足于客观范畴提出信息的定义、数学表达和基本性质。

3.1 信息的定义

基于普遍认识, 我们所处的世界由主观世界和客观世界构成。其中, 主观世界是指人的意识、观念世界, 是人们认识和把握整个世界的精神活动以及心理活动的总和; 客观世界是指物质的, 可以感知的世界, 是人的意识活动以外的一切物质及其运动的总和。

控制论的创始人 Wiener^[13] 在历史上第一个将信息提高到一种研究对象看待, 指出了“信息就是信息, 不是物质也不是能量”的著名论断。后人把这个观点简称为世界的物质、能量和信息三元论。笔者认为, 这里的“世界”实际上是指客观世界, 即物质、能量和信息是客观世界的三大构成要素^[72]。其中, 物质是指不依赖于人的主观意识而存在的客观实在, 能量是物质运动的能力, 信息是以物质或能量为媒介, 对客观世界和主观世界中事物及其运动状态的客观反映。简而言之, 物质是本源的存在, 能量是运动能力的存在, 信息是事物联系的存在。物理上的各类信息系统能够且只能够获取、传输、处理并应用各种客观存在的信息。因此, “三元论”认识满足信息系统运作对象的客观性要求。这些观点完全基于信息的客观实在性, 我们称之为客观信息论。由此可得信息的定义。

定义 1 信息是客观世界和主观世界中事物及其运动状态的客观反映。

这里的“客观反映”, 进一步展开就是“在客观世界中的反映”。因而, 无论客观世界或主观世界中的事物及其运动状态, 其客观反映都属于客观世界, 定义 1 所指的信息也完全属于客观世界范畴。据此定义, 可以进一步分解并引出如下具体概念。

定义 2(信息本体) 信息反映了客观世界和主观世界中“事物及其运动状态”, 这里的事物集合称为信息本体, 简称本体。

定义 3(信息状态发生时间) 信息所反映的应该是本体在特定时间集合上的运动状态, 这里的特定时间集合称为信息状态发生时间, 简称状态发生时间。

定义 4(信息状态集合) 信息反映了本体在状态发生时间上的“运动状态”, 这里的“运动状态”集合称为信息状态集合, 简称状态集合。

定义 5(信息载体) 信息只有依附客观世界中某些媒介集合才能得以反映, 这里的媒介集合称为信息载体, 简称载体。

定义 6(信息反映时间) 信息在载体中存在于特定的时间集合之上, 这里的时间集合称为信息反映时间, 简称反映时间.

定义 7(信息反映集合) 信息本身也以一系列状态存在于载体的反映时间之上, 这一系列状态称为信息反映集合, 简称反映集合.

3.2 信息的数学表达

我们定义 O 表示客观世界集合, S 表示主观世界集合, T 为时间集合, O , S 和 T 中的元素可以根据论域的特定要求进行合适的划分. 这样, 自然能够得到信息的数学定义.

定义 1' 设本体 $o \in 2^{O \cup S}$, 状态发生时间 $T_h \in 2^T$, o 在 T_h 上的状态集合 $f(o, T_h)$, 载体 $c \in 2^O$, 反映时间 $T_m \in 2^T$ 和 c 在 T_m 上的反映集合 $g(c, T_m)$ 均为非空集, 信息 I 便是 $f(o, T_h)$ 到 $g(c, T_m)$ 的满映射, 即

$$I : f(o, T_h) \rightarrow g(c, T_m), \quad (6)$$

或

$$I(f(o, T_h)) = g(c, T_m), \quad (7)$$

所有信息 I 的集合称为信息空间, 记为 $\mathcal{I}$, 是客观世界的三大构成要素之一.

需要说明的是, 定义中的映射并不仅限于单值映射, 也可以是多值映射. 为简化叙述, 下文一律将单值映射和多值映射统称为映射. 对于信息, 不仅映射 I 很重要, 本体 o 、状态发生时间 T_h 、状态集合 $f(o, T_h)$ 、载体 c 、反映时间 T_m 以及反映集合 $g(c, T_m)$ 也都具有非常重要的意义.

定义 8 称 $\langle o, T_h, f, c, T_m, g \rangle$ 为信息 I 的六元组模型, 也记为 $I = \langle o, T_h, f, c, T_m, g \rangle$.

3.3 信息的基本性质

根据六元组模型以及映射的数学理论, 可以进一步得到信息的一些基本性质.

(1) 客观性

按照定义 1', 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 是 $f(o, T_h)$ 到 $g(c, T_m)$ 的满映射, 因此信息 I 只有通过 $g(c, T_m)$ 才能得以体现. 而 $g(c, T_m)$ 是客观世界中的载体 c 在时间 T_m 上的状态集合, 所以信息 I 只有通过客观世界才能得以体现. 这是信息感知的约束条件.

(2) 还原性

信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 是 $f(o, T_h)$ 到 $g(c, T_m)$ 的满映射, 所以自然存在 I 的逆映射:

$$I^{-1} : g(c, T_m) \rightarrow f(o, T_h), \quad (8)$$

或

$$I^{-1}(g(c, T_m)) = f(o, T_h), \quad (9)$$

即根据 $g(c, T_m)$ 和 I^{-1} 能够还原到 o 在 T_h 上的状态 $f(o, T_h)$. 这是信息应用的根本前提.

(3) 传递性

信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 是 $f(o, T_h)$ 到 $g(c, T_m)$ 的满映射. 当然不排除存在客观世界中的集合 c' , 时间集合 T'_m 以及 c' 在 T'_m 上所有状态的集合 $g'(c', T'_m)$, 并形成 $g(c, T_m)$ 到 $g'(c', T'_m)$ 的满映射. 根据定义这时映射

$$I' : g(c, T_m) \rightarrow g'(c', T'_m), \quad (10)$$

也是信息, 并且

$$I'(g(c, T_m)) = I'(I(f(o, T_h))), \quad (11)$$

实际上是 $f(o, T_h)$ 通过 $g(c, T_m)$ 到 $g'(c', T'_m)$ 的复合映射.

可见, 复合映射 $I'(I(f(o, T_h)))$ 实现了信息由 o 到 c 再到 c' , 由 T_h 到 T_m 再到 T'_m , 由 $f(o, T_h)$ 到 $g(c, T_m)$ 再到 $g'(c', T'_m)$ 的传递. 这是信息传递的基本模式.

(4) 关联性

对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$, o 与 c , T_h 与 T_m , $f(o, T_h)$ 与 $g(c, T_m)$ 都成对出现, 特别是 $f(o, T_h)$ 与 $g(c, T_m)$, I 作为 $f(o, T_h)$ 到 $g(c, T_m)$ 的满映射, 建立了 o 和 c 这样一对事物所处状态之间的特定联系. 并且由于传递性, 信息还能将更多的事物联系在一起. 这是信息存在的重要意义.

(5) 组合性

信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 中 o, T_h, f, c, T_m, g 是作用不同的各种集合, 可以分解或组合为若干新的集合, 所以信息具有组合性质. 即

定义 9(子信息) 对信息 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$ 和 $I = \langle o, T_h, f, c, T_m, g \rangle$, 若

$$o' \subseteq o, T'_h \subseteq T_h, f' \subseteq f, c' \subseteq c, T'_m \subseteq T_m, g' \subseteq g, \quad (12)$$

并且

$$I'(f'(o', T'_h)) = I(f(o', T'_h)), \quad (13)$$

则称 I' 为 I 的子信息, 记为

$$I' \subseteq I, \quad (14)$$

同时在

$$o' \subset o, T'_h \subset T_h, f' \subset f, c' \subset c, T'_m \subset T_m, g' \subset g \quad (15)$$

中至少有一式成立时, 称 I' 为 I 的真子信息.

定义 10(组合信息) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 及其两个真子信息 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$ 和 $I'' = \langle o'', T''_h, f'', c'', T''_m, g'' \rangle$, 若

$$o = o' \cup o'', T_h = T'_h \cup T''_h, f = f' \cup f'', c = c' \cup c'', T_m = T'_m \cup T''_m, g = g' \cup g'', \quad (16)$$

且对任何 $o_\lambda \in o, T_{h\lambda} \in T_h, f_\lambda \in f, c_\lambda \in c, T_{m\lambda} \in T_m$, 有

$$I(f(o_\lambda, T_{h\lambda})) = I'(f'(o_\lambda, T_{h\lambda})) \text{ 或 } I''(f''(o_\lambda, T_{h\lambda})), \quad (17)$$

则称 I 是 I' 和 I'' 的组合, 记为

$$I = I' \cup I''. \quad (18)$$

4 信息的度量

建立信息的度量体系, 首先应该基于以下原则:

- (1) 溯源性原则 —— 从信息的定义模型出发, 形成各种度量的具体定义和数学表达;
- (2) 完整性原则 —— 从信息的实际内涵出发, 形成涵盖与其价值密切相关的完整度量体系;

(3) 普适性原则 —— 从信息的广泛类型出发, 形成普遍适用于信息获取、信息传输、信息处理、信息应用及其组合所成的各类信息系统, 而不仅限于某一特定领域的度量定义;

(4) 实用性原则 —— 从信息的应用要求出发, 形成具有实用性和可操作性, 能够指导信息系统分析研究的度量体系.

由此, 基于六元组模型的信息度量体系包含以下 9 类具体定义以及根据集合的测度、势和距离等性质很容易得到的相关基本命题.

4.1 信息的广阔性度量

信息是对本体状态的客观反映, 其价值与本体的覆盖范围直接相关. 一般而言, 在其他条件相同的情况下, 本体的覆盖范围越广阔, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有广阔性度量, 简称为广阔度. 其数学表达是:

定义 11 (信息的广阔度) 设 O 和 S 分别是客观世界和主观世界集合, $(O \cup S, 2^{O \cup S}, \mu)$ 是一个测度空间, μ 为可测集 $2^{O \cup S}$ 上的某种测度, 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 关于测度 μ 的广阔度 $\text{scope}_\mu(I)$ 就是 o 在可测集 $2^{O \cup S}$ 上的测度 $\mu(o)$, 即

$$\text{scope}_\mu(I) = \mu(o). \quad (19)$$

命题 1 (整体广阔度大于部分广阔度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 是 I 的子信息, 则

$$\text{scope}_\mu(I') \leq \text{scope}_\mu(I). \quad (20)$$

证明 因为 I' 是 I 的子信息, 根据式 (12),

$$o' \subseteq o. \quad (21)$$

基于测度的单调性, 故

$$\mu(o') \leq \mu(o), \quad (22)$$

所以式 (20) 成立.

4.2 信息的细致性度量

信息的价值也与其反映本体的细致程度直接相关, 直观上这种细致程度可以理解为本体能够分解成颗粒的粗细程度. 一般而言, 在其他条件相同的情况下, 本体的颗粒越细, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有细致性度量, 简称为细致度. 其数学表达是

定义 12 (原子信息) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 是 I 的真子信息, 且不存在任何 I 的其他真子信息 $I'' = \langle o'', T''_h, f'', c'', T''_m, g'' \rangle$, 使得

$$I'' \subset I', \quad (23)$$

则称 I' 为 I 的原子信息.

定义 13 (信息的细致度) 设 O 和 S 分别是客观世界和主观世界集合, $(O \cup S, 2^{O \cup S}, \mu)$ 是一个测度空间, μ 为可测集 $2^{O \cup S}$ 上的某种测度, 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 中所有原子信息的集合记为

$A = \{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{\lambda \in \Lambda}$, Λ 是指标集, 则信息 I 关于测度 μ 的细致度 $\text{granularity}_\mu(I)$ 就是 A 所有信息本体测度的最大值, 即

$$\text{granularity}_\mu(I) = \max_{\lambda \in \Lambda} \{\mu(o_\lambda)\}. \quad (24)$$

命题 2 (整体细致度粗于部分细致度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 的所有原子信息都是 I 的原子信息, 则

$$\text{granularity}_\mu(I') \leq \text{granularity}_\mu(I). \quad (25)$$

证明 设 $A = \{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{\lambda \in \Lambda}$ 和 $A' = \{I_{\lambda'} = \langle o_{\lambda'}, T_{h\lambda'}, f_{\lambda'}, c_{\lambda'}, T_{m\lambda'}, g_{\lambda'} \rangle\}_{\lambda' \in \Lambda'}$ 分别为信息 I 和 I' 的原子信息集合. 因为 I' 的所有原子信息都是 I 的原子信息, 所以 $A' \subseteq A$.

进而可知

$$\max_{\lambda' \in \Lambda'} \{\mu(o_{\lambda'})\} \leq \max_{\lambda \in \Lambda} \{\mu(o_\lambda)\}, \quad (26)$$

从而式 (25) 成立.

4.3 信息的持续性度量

信息的价值也与其反映本体的持续程度相关, 直观上这种持续程度体现在发生时间分布的密集性及其跨度上. 一般而言, 在其他条件相同的情况下, 发生时间分布越密集, 跨度越大, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有持续性度量, 简称为持续度. 其数学表达是

定义 14 (信息的持续度) 设 T 为时间集合, $(T, 2^T, \tau)$ 是一个测度空间, τ 为可测集 2^T 上的某种测度, 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 关于测度 τ 的持续度 $\text{sustainability}_\tau(I)$ 就是 T_h 在可测集 2^T 上的测度 $\tau(T_h)$, 即

$$\text{sustainability}_\tau(I) = \tau(T_h). \quad (27)$$

命题 3 (整体持续度大于部分持续度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 是 I 的子信息, 则

$$\text{sustainability}_\tau(I') \leq \text{sustainability}_\tau(I). \quad (28)$$

证明 因为 I' 是 I 的子信息, 根据式 (12),

$$T'_h \subseteq T_h. \quad (29)$$

基于测度的单调性, 故

$$\tau(T'_h) \leq \tau(T_h), \quad (30)$$

从而式 (28) 成立.

4.4 信息的丰富性度量

信息的价值还与其反映本体状态的丰富程度直接相关, 这种丰富程度实际上体现于反映函数内容的众寡. 一般而言, 在其他条件相同的情况下, 反映函数所含内容越丰富, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有丰富性度量, 简称为丰富度. 其数学表达是:

定义 15 (信息的丰富度) 设 $\mathcal{F}$ 是本体 o 在时间 T_h 上所有状态函数的集合, $(\mathcal{F}, 2^{\mathcal{F}}, \rho)$ 是一个测度空间, ρ 为可测集 $2^{\mathcal{F}}$ 上的某种测度, 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 关于测度 ρ 的丰富度 $\text{richness}_{\rho}(I)$ 就是 $f(o, T_h)$ 在可测集 $2^{\mathcal{F}}$ 上的测度 $\rho(f(o, T_h))$, 即

$$\text{richness}_{\rho}(I) = \rho(f(o, T_h)). \quad (31)$$

命题 4 (整体丰富度大于部分丰富度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 是 I 的子信息, 则

$$\text{richness}_{\rho}(I') \leq \text{richness}_{\rho}(I). \quad (32)$$

证明 因为 I' 是 I 的子信息, 根据式 (12),

$$f' \subseteq f. \quad (33)$$

基于测度的单调性, 故

$$\rho(f') \leq \rho(f), \quad (34)$$

从而式 (32) 成立.

4.5 信息的包容性度量

信息的价值还与其内在的包容性密切相关, 这种包容性体现在对载体容量的需求. 一般而言, 在其他条件相同的情况下, 载体的容量需求越小, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有包容性度量, 简称为容积度. 其数学表达是:

定义 16 (信息的容积度) 设 O 表示客观世界集合, $(O, 2^O, \sigma)$ 是一个测度空间, σ 为可测集 2^O 上的某种测度, 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 关于测度 σ 的容积度 $\text{volume}_{\sigma}(I)$ 就是 o 在可测集 2^O 上的测度 $\sigma(c)$, 即

$$\text{volume}_{\sigma}(I) = \sigma(c). \quad (35)$$

命题 5 (整体容积度大于部分容积度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 是 I 的子信息, 则

$$\text{volume}_{\sigma}(I') \leq \text{volume}_{\sigma}(I). \quad (36)$$

证明 因为 I' 是 I 的子信息, 根据式 (12),

$$c' \subseteq c. \quad (37)$$

基于测度的单调性, 故

$$\sigma(c') \leq \sigma(c), \quad (38)$$

从而式 (36) 成立.

Shannon 信息熵实际上就是通信系统传送离散消息所需要的信息容积度指标. 事实上, 在通信过程中, 信息的发送者、通信系统和接收者相继成为信息的载体, 发送者和接收者之间正是通过通信系统这个载体才得以实现信息的传递.

对于文献 [3] 中述及需要传送的离散编码序列 $\{x_i\}(i = 1, 2, \dots, n)$, 信息 I 的本体 o 可以是客观世界或主观世界中的任一事物, 发生时间 T_h 就是 o 取值 x_i 的具体时刻, 状态集合 $f(o, T_h)$ 是一组随机事件的集合, 通过 o 取值 x_i 的概率得以表达, 即

$$P(o = x_i) = p_i, i = 1, 2, \dots, n, \quad (39)$$

且 $\sum_{i=1}^n p_i = 1$. 载体 c 就是通信系统, 反映时间 T_m 就是通信传输的时间, 反映集合 $g(c, T_m)$ 就是传输的具体内容. 我们定义 c 的测度 $\sigma(c)$ 为传输内容的二进制编码位数, 显然它符合测度定义的要求.

直观上, 若 o 取值 x_i , 可以简单地通过传送序号 i 使接收者了解 o 的最终取值. 因为 i 可取的最大值为 n , 所以这时

$$\text{volume}_\sigma(I) = \log_2 n. \quad (40)$$

但是对于不同的 i , 载体 o 并不是以等概率取值 x_i , 所以根据 Shannon 在文献 [3] 中得到的熵公式, 理论上可以对传输内容进行更加合理的编码设计, 使其所需二进制位数为 $-\sum_{i=1}^n p_i \log_2 p_i$. 此时

$$\text{volume}_\sigma(I) = -\sum_{i=1}^n p_i \log_2 p_i, \quad (41)$$

式 (41) 右端总不大于式 (40) 右端, 且在 n 很大而概率 p_i 之间相差也很大时, 两者之间的差距会很大. 所以 Shannon 信息熵是很多情况下通信系统信息容积度指标的最优值.

4.6 信息的延迟性度量

信息的价值当然也与其时效密切相关, 这种时效体现在反映时间对状态发生时间的延迟上. 一般而言, 在其他条件相同的情况下, 反映时间对状态发生时间的延迟越小, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有延迟性度量, 简称为延迟度. 其数学表达是

定义 17 (信息的延迟度) 信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 中所有原子信息的集合记为 $A = \{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{\lambda \in \Lambda}$, Λ 是指标集, 则信息 I 的延迟度 $\text{delay}(I)$ 就是 A 中所有原子信息的反映时间下确界与发生时间上确界之差的最大值, 即

$$\text{delay}(I) = \max_{\lambda \in \Lambda} \{\inf T_{m\lambda} - \sup T_{h\lambda}\}, \quad (42)$$

其中, $\inf T_{m\lambda}$ 和 $\sup T_{h\lambda}$ 分别表示集合 A 中 $T_{m\lambda}$ 和 $T_{h\lambda}$ 的下确界和上确界. 若存在某个 $T_{h\lambda}$ 的上确界为 $+\infty$, 则

$$\inf T_{m\lambda} - \sup T_{h\lambda} = 0. \quad (43)$$

信息延迟度刻画了载体对本体状态反映的快慢. 从其定义可知, 延迟度具有正负性. 特别地, 当

$$\inf T_{m\lambda} < \sup T_{h\lambda}, \text{delay}(I) < 0, \quad (44)$$

这表示载体在本体状态发生时间 T_h 之前, 对未来相关状态的预测.

命题 6 (整体延迟度大于部分延迟度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 和 $I' = \langle o', T'_h, f', c', T'_m, g' \rangle$, 若 I' 的所有原子信息都是 I 的原子信息, 则

$$\text{delay}(I') \leq \text{delay}(I). \quad (45)$$

证明 设

$$A = \{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{\lambda \in \Lambda}$$

和

$$A' = \{I_{\lambda'} = \langle o_{\lambda'}, T_{h\lambda'}, f_{\lambda'}, c_{\lambda'}, T_{m\lambda'}, g_{\lambda'} \rangle\}_{\lambda' \in \Lambda'}$$

分别为信息 I 和 I' 的原子信息集合. 因为 I' 的所有原子信息都是 I 的原子信息, 所以 $A' \subseteq A$.

进而可知

$$\max_{\lambda' \in \Lambda'} \{\inf T_{m\lambda'} - \sup T_{h\lambda'}\} \leq \max_{\lambda \in \Lambda} \{\inf T_{m\lambda} - \sup T_{h\lambda}\}, \quad (46)$$

从而式 (45) 成立.

4.7 信息的遍及性度量

信息的价值在很多情形与其普及程度密切相关, 这种普及程度体现在信息载体的分布范围上. 对某些信息, 其载体在一定范围内分布越广泛, 价值越高; 而对另外一些信息, 其载体分布范围越狭窄, 价值反而越高. 因此, 不管何种情形, 都需要考虑信息的遍及性度量, 简称为遍及度. 其数学表达是:

定义 18 (同义信息) 设信息 $I_1 = \langle o_1, T_{h1}, f_1, c_1, T_{m1}, g_1 \rangle$ 和 $I_2 = \langle o_2, T_{h2}, f_2, c_2, T_{m2}, g_2 \rangle$ 是信息 $I = \langle o, T_h, f, c, T_m, g \rangle$ 的两个子信息, 若

$$I^{-1}(g_1(c_1, T_{m1})) = I^{-1}(g_2(c_2, T_{m2})), \quad (47)$$

则称 I_1 与 I_2 是 I 中的两个同义信息.

定义 19 (信息的遍及度) 设信息 $I_0 = \langle o_0, T_{h0}, f_0, c_0, T_{m0}, g_0 \rangle$ 是 $I = \langle o, T_h, f, c, T_m, g \rangle$ 的子信息, 则信息 I_0 在 I 中的遍及度 $\text{coverage}_{I_0}(I)$ 是 I 中所有与 I_0 同义的子信息载体集合的势与 c 的势之比. 即

$$\text{coverage}_{I_0}(I) = \frac{|\cup_{\lambda \in \Lambda} c_\lambda|}{|c|}, \quad (48)$$

其中, $\{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{\lambda \in \Lambda}$ (Λ 是指标集) 是与信息 I_0 同义的子信息集合.

命题 7 (部分遍及度大于整体遍及度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$, $I_0 = \langle o_0, T_{h0}, f_0, c_0, T_{m0}, g_0 \rangle$ 和 $I_1 = \langle o_1, T_{h1}, f_1, c_1, T_{m1}, g_1 \rangle$, 若

$$I_0 \subseteq I_1 \subseteq I, \quad (49)$$

则

$$\text{coverage}_{I_1}(I) \leq \text{coverage}_{I_0}(I). \quad (50)$$

证明 设

$$A_0 = \{I_{\lambda_0} = \langle o_{\lambda_0}, T_{h\lambda_0}, f_{\lambda_0}, c_{\lambda_0}, T_{m\lambda_0}, g_{\lambda_0} \rangle\}_{\lambda_0 \in \Lambda_0}$$

和

$$A_1 = \{I_{\lambda_1} = \langle o_{\lambda_1}, T_{h\lambda_1}, f_{\lambda_1}, c_{\lambda_1}, T_{m\lambda_1}, g_{\lambda_1} \rangle\}_{\lambda_1 \in \Lambda_1}$$

分别为所有与信息 I_0 和 I_1 同义的子信息集合. 对于任意 $I_{\lambda_1} \in A_1$, 根据式 (47),

$$I^{-1}(g_{\lambda_1}(c_{\lambda_1}, T_{m\lambda_1})) = I^{-1}(g_1(c_1, T_{m1})) = f_1(o_1, T_{h1}). \quad (51)$$

根据式 (12) 和式 (51),

$$f_0(o_0, T_{h0}) \subseteq f_1(o_1, T_{h1}), \quad (52)$$

$$I(f_0(o_0, T_{h0})) \subseteq g_{\lambda_1}(c_{\lambda_1}, T_{m\lambda_1}). \quad (53)$$

令

$$I(f_0(o_0, T_{h0})) = g_{\lambda_0}(c_{\lambda_0}, T_{m\lambda_0}), \quad (54)$$

则

$$\langle o_0, T_{h0}, f_0, c_0, T_{m0}, g_0 \rangle \in A_0, \quad (55)$$

并且

$$g_{\lambda_0}(c_{\lambda_0}, T_{m\lambda_0}) \subseteq g_{\lambda_1}(c_{\lambda_1}, T_{m\lambda_1}), \quad (56)$$

可见对于 A_1 中不同的 I_{λ_1} , 总存在 A_0 中不同的 I_{λ_0} 与之对应, 所以

$$|A_1| \leq |A_0|. \quad (57)$$

因此,

$$|\cup_{\lambda_1 \in \Lambda_1} c_{\lambda_1}| \leq |\cup_{\lambda_0 \in \Lambda_0} c_{\lambda_0}|, \quad (58)$$

从而式 (50) 成立.

4.8 信息的真实性度量

信息作为本体状态与载体状态之间的映射, 本无真假可言. 因为根据还原性, 理论上总存在其逆映射能够还原信息本体在发生时间的实际状态. 但实际上由于映射过程的复杂性, 加之于认识过程中的主观因素, 很多情况下根本无法准确求得信息的逆映射, 只能通过尽可能的推演处理, 才能近似地了解本体在发生时间的实际状态. 即

定义 20 (语义映射和语义状态) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$, 设有映射 J 使

$$J(g(c, T_m)) = \tilde{f}(\tilde{o}, \tilde{T}_h), \quad (59)$$

其中 $\tilde{o} \in 2^{O \cup S}$, $\tilde{T}_h \in 2^T$, $\tilde{f}(\tilde{o}, \tilde{T}_h)$ 为 $\tilde{o}$ 在 $\tilde{T}_h$ 上的某一状态集合, 则称 J 为 I 的一个语义映射, $\tilde{f}(\tilde{o}, \tilde{T}_h)$ 为 I 关于 J 的语义状态.

信息的价值在很大程度上取决于能否求得合适的语义映射, 使其语义状态与本体实际状态尽可能接近. 一般而言, 在其他条件相同的情况下, 信息所反映的语义状态与实际状态的差异越小, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有与语义映射相关的真实性度量, 简称为真实度. 其数学表达是

定义 21 (信息的真实度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$, 设其状态集合 $f(o, T_h)$ 和反映集合 $g(c, T_m)$ 关于语义映射 J 的语义状态 $\tilde{f}(\tilde{o}, \tilde{T}_h)$ 都是距离空间 $\langle \mathcal{F}, d \rangle$ 中的元素, 其中 d 是 $\mathcal{F}$ 上的距离, 则称信息 I 在距离空间 $\langle \mathcal{F}, d \rangle$ 中关于语义 J 的真实度 $\text{validity}_J(I)$ 就是 $f(o, T_h)$ 与 $\tilde{f}(\tilde{o}, \tilde{T}_h)$ 的距离, 即

$$\text{validity}_J(I) = d(f, \tilde{f}). \quad (60)$$

4.9 信息的适配性度量

信息的价值最终必定取决于其对用户需求的满足程度, 这种满足程度体现于信息的各项组成元素对用户需求的适配性. 一般而言, 适配性越强, 信息的价值越高; 反之, 信息的价值越低. 因此, 信息具有适配性度量, 简称为适配度. 其数学表达是

定义 22 (信息的适配度) 对信息 $I = \langle o, T_h, f, c, T_m, g \rangle$, 设 o, T_h, f, c, T_m, g 分别是集合 $\mathcal{P}_o, \mathcal{P}_{T_h}, \mathcal{P}_f, \mathcal{P}_c, \mathcal{P}_{T_m}, \mathcal{P}_g$ 中的元素且 I 是距离空间 $\langle (\mathcal{P}_o, \mathcal{P}_{T_h}, \mathcal{P}_f, \mathcal{P}_c, \mathcal{P}_{T_m}, \mathcal{P}_g), d \rangle$ 中的元素. 对六元组集合

$$\{I_\lambda = \langle o_\lambda, T_{h\lambda}, f_\lambda, c_\lambda, T_{m\lambda}, g_\lambda \rangle\}_{(o_\lambda \in \mathcal{P}_o, T_{h\lambda} \in \mathcal{P}_{T_h}, f_\lambda \in \mathcal{P}_f, c_\lambda \in \mathcal{P}_c, T_{m\lambda} \in \mathcal{P}_{T_m}, g_\lambda \in \mathcal{P}_g)}, \quad (61)$$

$\lambda \in \Lambda, \Lambda$ 是指标集, 其中的目标六元组 $I_{\lambda_0} = \langle o_{\lambda_0}, T_{h\lambda_0}, f_{\lambda_0}, c_{\lambda_0}, T_{m\lambda_0}, g_{\lambda_0} \rangle$, 信息 I 关于目标信息 I_{λ_0} 的适配度定义为

$$\text{suitability}_{I_{\lambda_0}}(I) = d(I, I_{\lambda_0}). \quad (62)$$

上述 9 类度量源自信息六元组模型, 结合了信息的内在特点和外在需求, 每个度量都通过特定要素的测度、势和距离等量化形式得以表达, 同时每个要素都在若干度量中得到体现, 整个度量体系较为全面, 度量之间具有独立性. 虽然为适用于一般情形, 测度、势和距离的定义较为抽象, 但对具体信息系统, 这些定义能够直接对应于明确浅显的数学公式, 因而可以在实际应用领域得到有效运用. 对于信息的真实度和适配度, 我们没有得到部分与整体之间的指标关系, 说明无论是真实性还是适配性, 都不会因为信息的众寡而变化, 这既符合人们的常识, 也符合信息系统的运作规律.

5 结论

世界万物只有通过其客观反映也就是信息, 才能表现存在及运动状态. 事物之间只有通过信息才能交互联系. 正因如此, 信息得以和物质、能量等价齐观, 在客观世界中鼎足三分有其一. 对于信息的重要性人们已有充分认识, 各方面研究也取得累累硕果. 特别是 Shannon 揭示了信息传输的根本要义, 为人类迈向信息社会作出了杰出贡献; 钟义信^[73]提出了全信息理论, 为信息与人类认知和智能的深度交互打开了理论之门; Burgin 基于公理化方法建立了一般信息论的模型和算子体系, 为运用各种数学工具研究信息的本质和运作提供了示范. 基于前人成果, 本文探索了信息的本质内涵, 较之于 Shannon, 将信息内涵从通信传输扩充到信息运转的所有范畴; 较之于钟义信, 将信息本质从主客观全维视角收敛到客观本源; 较之于 Burgin, 将局限于“引起变化”的信息拓展到“客观存在”的信息, 进而提出了客观信息的定义, 建立了信息的六元组模型和 9 类度量定义, 形成了客观信息论的基础理论框架.

客观信息论的数学基础是集合论、测度论和拓扑学, 虽然较为抽象, 但与人们日常广泛运用的统计、计算方法之间具有直接明了的对应关系, 因而完全适用于当前备受关注的大数据、流媒体等信息价值分析、信息系统运行机理和研究以及信息系统设计和评估等信息化进程中普遍需要面对的各项具体工作, 其完备性和实用性也将在后续大量应用中得到检验和修正, 因为任何一种理论都只有在实践中才能体现其价值并得到完善. 同时, 本文提出的一些概念还需要从很多方向结合实际不断深化研究. 比如信息依赖于载体的反映集合得以体现, 实际应用中如何鉴别载体反映集合的异同, 其具体准则和方法等都将是今后深化研究的重要方向. 另一方面, 信息是客观世界与主观世界相互作用的重要桥梁, 因而信息在主观世界对客观世界的认知过程中能够发挥多大作用, 这种作用应该如何度量等也应是今后结合物理、生物、心理、哲学等众多学科知识进行深化研究的重大课题.

致谢 本文撰写过程中, 作者曾经与中国北京邮电大学钟义信教授深入讨论, 并得到了钟教授的认真指导. 中国电子科学研究院朱德成研究员对本文提出了宝贵意见. 李欣欣和方芳两位女士为本文撰写提供了很多帮助. 在此, 一并向他们致以衷心的感谢!

参考文献

- 1 Zaliwski A S. Information-is it subjective or objective? Triple C: communication, capitalism & critique. *Open Access J Global Sust Inform Soc*, 2011, 9: 77–92
- 2 Hartley R V L. Transmission of information. *Bell Syst Tech J*, 1928, 7: 535–536
- 3 Shannon C E. The mathematical theory of communication. *Bell Syst Tech J*, 1948, 27: 379–423
- 4 Zhong Y X. Principles of Information Science. 3rd ed. Beijing: Beijing University of Posts and Telecommunications Press, 2002. 111–116 [钟义信. 信息科学原理. 第三版. 北京: 北京邮电大学出版社, 2002. 111–116]
- 5 Xiao F. Informationalism: From Society View to World View. Beijing: China Social Sciences Press, 2010. 268 [肖峰. 信息主义: 从社会观到世界观. 北京: 中国社会科学出版社, 2010. 268]
- 6 Yan X S. Information science: its past, present and future. *Inform*, 2011, 2: 510–527
- 7 Burgin M. Theory of Information: Fundamentality, Diversity and Unification. Singapore: World Scientific Publishing Co Pty Ltd., 2010
- 8 Fleissner P, Hofkirchner W. Emergent information: towards a unified information theory. *BioSyst*, 1996, 38: 243–248
- 9 Capurro R, Fleissner P, Hofkirchner W. Is a unified theory of information feasible? A dialogue. In: *Proceedings of the 2nd International Conference on the Foundations of Information Science*, Amsterdam, 1999. 9–30
- 10 Zhang J G, Singh V P. Information Entropy—Theory and Application. Beijing: China Water Power Press, 2012. 41–42 [张继国, Singh V P. 信息熵——理论与应用. 北京: 中国水利水电出版社, 2012. 41–42]
- 11 Zhang W M. Information System Principles and Engineering. 3rd ed. Beijing: Electronics Industry Press, 2009. 53–54 [张维明. 信息系统原理与工程. 第三版. 北京: 电子工业出版社, 2009. 53–54]
- 12 Burgin M. Information operators in categorical information spaces. *Inform*, 2010, 1: 119–152
- 13 Wiener N. Cybernetics: or the Control and Communication in the Animal and the Machine. 2nd ed. Paris: MIT Press, 1961
- 14 Ashby W R. An Introduction to Cybernetics. London: Chapman & Hall Ltd., 1956
- 15 De Luca A, Termini S. A definition of nonprobabilistic entropy in the setting of fuzzy sets theory. *Inform Control*, 1972, 20: 301–312
- 16 Rényi A. On the dimension and entropy of probability distributions. *Acta Math Acad Sci Hung*, 1959, 10: 193–215
- 17 Kullback S. Information Theory and Statistics. New York: Courier Dover Publications, 1997
- 18 Asadi M, Zohrevand Y. On the dynamic cumulative residual entropy. *J Stat Plan Infer*, 2007, 137: 1931–1941
- 19 Kumar V, Taneja H C. Some characterization results on generalized cumulative residual entropy measure. *Stat Probabil Lett*, 2011, 81: 1072–1077
- 20 Rao M, Chen Y, Vemuri B C, et al. Cumulative residual entropy: a new measure of information. *IEEE Trans Inform Theory*, 2004, 50: 1220–1228
- 21 Zografos K, Nadarajah S. Survival exponential entropies. *IEEE Trans Inform Theory*, 2005, 51: 1239–1246
- 22 Abbas A E. Entropy methods for joint distributions in decision analysis. *IEEE Trans Eng Manage*, 2006, 53: 146–159
- 23 Madiman M, Tetali P. Information inequalities for joint distributions, with interpretations and applications. *IEEE Trans Inform Theory*, 2010, 56: 2699–2713
- 24 Gupta G, Chaturvedi A. Conditional entropy based user selection for multiuser MIMO systems. *IEEE Commun Lett*, 2013, 17: 1628–1631
- 25 Ho S W, Verdú S. On the interplay between conditional entropy and error probability. *IEEE Trans Inform Theory*, 2010, 56: 5930–5942
- 26 Teixeira A, Matos A, Antunes L. Conditional Rényi entropies. *IEEE Trans Inform Theory*, 2012, 58: 4273–4277
- 27 Hu Q, Pan W, An S, et al. An efficient gene selection technique for cancer recognition based on neighborhood mutual information. *Int J Mach Learn Cyber*, 2010, 1: 63–74
- 28 Rusek F, Lozano A, Jindal N. Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels. *IEEE Trans Inform Theory*, 2012, 58: 331–340

- 29 Song S, Zhang Q. Mutual information of multipath channels with imperfect channel information. *IEEE Trans Commun*, 2009, 57: 1523–1531
- 30 Viola P, Wells III W M. Alignment by maximization of mutual information. *Int J Comput Vision*, 1997, 24: 137–154
- 31 Wang L J. An improved multiple fuzzy NNC system based on mutual information and fuzzy integral. *Int J Mach Learn Cyber*, 2011, 2: 25–36
- 32 Zhu Y, Kam P Y, Xin Y. On the mutual information distribution of MIMO Rician fading channels. *IEEE Trans Commun*, 2009, 57: 1453–1462
- 33 An S, Yang S, Ho S L, et al. An improved cross-entropy method applied to inverse problems. *IEEE Trans Magn*, 2012, 48: 327–330
- 34 Balakrishnan N, Rad A H, Arghami N R. Testing exponentiality based on Kullback-Leibler information with progressively type-II censored data. *IEEE Trans Reliab*, 2007, 56: 301–307
- 35 De Boer P T, Kroese D P, Mannor S, et al. A tutorial on the cross-entropy method. *Ann Oper Res*, 2005, 134: 19–67
- 36 Carvalho L M, González-Fernández R A, Leite da Silva A M, et al. Simplified cross-entropy based approach for generating capacity reliability assessment. *IEEE Trans Power Syst*, 2013, 28: 1609–1616
- 37 Balakrishnan N, Rad A H, Arghami N R. Testing exponentiality based on Kullback-Leibler information with progressively type-II censored data. *IEEE Trans Reliab*, 2007, 56: 301–307
- 38 Park S. Testing exponentiality based on the Kullback-Leibler information with the type II censored data. *IEEE Trans Reliab*, 2005, 54: 22–26
- 39 Xie H, Zheng R, Guo J, et al. Cross-fuzzy entropy: a new method to test pattern synchrony of bivariate time series. *Inform Sci*, 2010, 180: 1715–1724
- 40 Chen X, Dai W. Maximum entropy principle for uncertain variables. *Int J Fuzzy Syst*, 2011, 13: 232–236
- 41 Li X, Liu B. Maximum entropy principle for fuzzy variables. *Int J Uncertain Fuzz*, 2007, 15: 43–52
- 42 De Oliveira D L L, do Nascimento M Z, Neves L A, et al. Unsupervised segmentation method for cuboidal cell nuclei in histological prostate images based on minimum cross entropy. *Expert Syst Appl*, 2013, 40: 7331–7340
- 43 Zhang Y L, Zhang W, Xiao G, et al. Fast thresholding based on improved minimum cross entropy. In: *Proceedings of 12th IEEE International Conference on Control, Automation and Systems (ICCAS)*, Jeju, 2012. 1108–1112
- 44 Rao C R. Information and accuracy attainable in the estimation of statistical parameters. *Bull Calcutta Math Soc*, 1945, 37: 81–91
- 45 Li X, Cheng Y Q, Wang H Q, et al. Progress in theory and applications of information geometry. *Sci Sin Inform*, 2013, 43: 707–732 [黎湘, 程永强, 王宏强, 等. 信息几何理论与应用研究进展. *中国科学: 信息科学*, 2013, 43: 707–732]
- 46 Efron B. Defining the curvature of a statistical problem (with applications to second order efficiency). *Ann Stat*, 1975: 1189–1242
- 47 Chentsov N N. *Statistical Decision Rules and Optimal Inference*. Rhode Island: American Mathematical Soc., 1982
- 48 Amari S I. Natural gradient works efficiently in learning. *Neural Comput*, 1998, 10: 251–276
- 49 Amari S I. *Differential Geometry Methods in Statistics*. Berlin: Springer, 1985
- 50 Amari S I, Nagaoka H. *Methods of Information Geometry*. New York: Oxford University Press, 2000
- 51 Carter K M, Raich R, Finn W G, et al. Fine: Fisher information nonparametric embedding. *IEEE Trans Pattern Anal Mach Intell*, 2009, 31: 2093–2098
- 52 Cont A, Dubnov S, Assayag G. On the information geometry of audio streams with applications to similarity computing. *IEEE Trans Audio Speech Lang Process*, 2011, 19: 837–846
- 53 Garcia V, Nielsen F. Simplification and hierarchical representations of mixtures of exponential families. *Signal Process*, 2010, 90: 3197–3212
- 54 Goh A, Vidal R. Unsupervised Riemannian Clustering of Probability Density Functions. In: *Proceedings of Machine Learning and Knowledge Discovery in Databases*. Berlin: Springer, 2008. 377–392
- 55 Pereyra M, Batatia H, McLaughlin S. Exploiting information geometry to improve the convergence properties of variational active contours. *IEEE J-STSP*, 2013, 7: 700–707
- 56 Peter A M, Rangarajan A. Information geometry for landmark shape analysis: unifying shape representation and deformation. *IEEE Trans Pattern Anal Mach Intell*, 2009, 31: 337–350
- 57 Sira S P, Li Y, Papandreou-Suppappola A, et al. Waveform-agile sensing for tracking. *IEEE Signal Process Mag*, 2009, 26: 53–64

- 58 Unnikrishnan J, Huang D, Meyn S P, et al. Universal and composite hypothesis testing via mismatched divergence. *IEEE Trans Inform Theory*, 2011, 57: 1587–1603
- 59 Varshney K R. Bayes risk error is a Bregman divergence. *IEEE Trans Signal Process*, 2011, 59: 4470–4472
- 60 Westover M B. Asymptotic geometry of multiple hypothesis testing. *IEEE Trans Inform Theory*, 2008, 54: 3327–3329
- 61 Wuhler S, Shu C, Rioux M. Posture invariant gender classification for 3D human models. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Miami, 2009. 33–38
- 62 Yang Z, Laaksonen J. Principal whitened gradient for information geometry. *Neural Netw*, 2008, 21: 232–240
- 63 Gianfelici F, Battistelli V. *Methods of information geometry* (Amari S I, Nagaoka H. American Mathematical Society and Oxford University Press, 2000)(book review). *IEEE Trans Inform Theory*, 2009, 55: 2905–2906
- 64 Burgin M. Information theory: a multifaceted model of information. *Entropy*, 2003, 5: 146–160
- 65 Daft R L, Lengel R H. Information richness: a new approach to managerial information processing and organization design. *Manage Sci*, 1986, 32: 554–571
- 66 Liang J Y, Qian Y H. Information granules and entropy theory in information systems. *Sci China Ser-F: Inf Sci*, 2008, 51: 1427–1444
- 67 Teng S H, Lu M, Zhang J, et al. The entropy theory and information granularity in information system. *Comput Eng Sci*, 2012, 34: 94–101 [滕书华, 鲁敏, 张军, 等. 信息系统中的熵理论和信息粒度. *计算机工程与科学*, 2012, 34: 94–101]
- 68 Gaziano G, McGrath K. Measuring the concept of credibility. *J Quart*, 1986, 63: 451–462
- 69 Alberts D, Garstka J. *Basic Principles and Measurements of Network-centric Actions*. Beijing: National Defense Industry Press, 2007 [Alberts D, Garstka J. 网络中心行动的基本原理及其度量. 北京: 国防工业出版社, 2007]
- 70 Darilek R, Perry W, Bracken J, et al. *Measures of Effectiveness for the Information-Age Army*. Report MR-1155-A of the RAND Corporation. 2001
- 71 Niven D. Objective evidence on media bias: newspaper coverage of congressional party switchers. *J Mass Commun Quart*, 2003, 80: 311–326
- 72 Lü X K. *Informatization in China*. Beijing: Electronics Industry Press, 2002. 17 [吕新奎. 中国信息化. 北京: 电子工业出版社, 2002. 17]
- 73 Zhong Y X. Principles of information conversions—an integrated theory of information, knowledge and intelligence. *Chinese Sci Bull*, 2013, 58: 1300–1306 [钟义信. 信息转换原理: 信息, 知识, 智能的一体化理论. *科学通报*, 2013, 58: 1300–1306]

Research on metrics and models for objective information

XU JianFeng^{1*}, TANG Jun², MA XueFeng¹, XU Bin¹, SHEN YanLi¹ & QIAO YongJie¹

1 *China Academy of Electronics and Information Technology, Beijing 100041, China;*

2 *Department of Electronic Engineering, Tsinghua University, Beijing 100084, China*

*E-mail: xjfcetc@163.com

Abstract In this contemporary era, the importance of information is undisputed, but there has never been a common understanding of information, nor has a unanimous conclusion been reached by studies conducted on information metrics. On the basis of previous studies conducted, this paper analyzes the important achievements in research conducted into the properties and metrics of information, as well as their main insufficiencies, and discusses the essence, connotations, mathematical expressions, and basic problems related to information. On the basis of the understanding of the objectivity of information, it then proposes definitions and a sextuple model of objective information, discusses the basic properties of information, and explores the definitions and mathematical expressions of nine kinds of information metrics: extensity, detailedness, sustainability, richness, containability, delay, distribution, validity, and matchability. Through these metrics, this paper establishes a basic underlying framework for objective information theory that provides comprehensive support for systematic analysis and research on information and information systems.

Keywords information, information entropy, objective information theory, sextuple model, metrics

XU JianFeng was born in 1961. He received a Ph.D. degree in Software Science from Nanjing University, Nanjing, in 2001. He is currently a senior researcher at China Academy of Electronics and Information Technology. His research interests include information systems and information theory. Dr. XU is a member of the Chinese Institute of Electronics.