

关于 21 世纪信息科学发展的一些见解

Some Viewpoints on the Development of Information Science in the 21st Century

闫学杉

(北京大学信息管理系,副教授 北京 100871)

今日的信息理论像一个多学科讲坛,不同的信息问题研究者,把一些分布在不同领域里的信息概念加以阐述、系统化并提出了许多新颖的命题。传播学家呼吁:公平的信息获取才是信息对称的基本保障;通讯工程师说:当受信者对一个事件出现的验前概率估计越小时,他获得的信息量就越大;计算机科学家则说:微处理器时钟频率的大小,决定着它处理信息的速度高低;生物学家则高呼:脱氧核糖核酸上的信息是控制子代和亲代相像的唯一源泉。

凡此种种表达了不同的信息问题研究者对不同领域信息问题的见解。他们用各自学科的语言阐述自己的看法,而且都深刻地洞察了他们自己所关心的信息对象的本质。这些研究者的工作都和信息问题相关,尽管这个共性是显而易见的,但从理论研究的角度看,代表这些共性探索的真正的信息科学却不存在。虽然“信息科学”这个名称早在 40 年前就出现了,但对信息科学的研究内容却一直界定不清,它时而指计算机科学家所说的信息,时而指文献学家所说的信息,有时又代表通讯科学家所说的信息。而真正的信息科学所关注的信息,应该不是指任何个别信息领域里的特定信息,它应该反映着一切和信息有关领域里的信息现象,关怀着一切和信息有关领域里的信息问题。从 1948 年《信息论》和《控制论》的问世到现在,已经整整 50 年过去了,尝试建立这样的信息科学理论一直是许多人的梦想。昨天它没有出现,在即将到来的 21 世纪它会出现吗?21 世纪的信息科学又会是个什么样子呢?这是研究信息问题的人都关心的问题。本文试图就这些问题提出自己的一些见解。

一、信息的概念和信息科学研究的历史

本世纪 20 年代之前,信息概念在东西方的使用,尤其在英语和汉语中的使用,其范围一直局限在人类日常生活和文学作品领域里,含义多指关于人的消息或事物的报道。从 20 年代起,一些数学家开始尝试着把这个概念引进统计学。此后不久,通讯科学家把它引进到电子通讯工程研究中,但它的主要

应用范围仍旧局限于人类生活和文学作品里。这种局面一直延续到 20 世纪 40 年代后期。

从 1948 年开始,由于申农就通讯工程信道容量的研究创建了一套出色的统计理论,一时间,信息概念在通讯工程中开始流行起来,并第一次超越了它在人类生活和文学作品中的出现频率。同时,由于维纳关于在生物和机器中通讯和控制问题的研究成果的问世和普及,信息概念又迅速地扩展到了生物学、哲学、心理学、神经病理学和语言学等研究领域。据统计,至 80 年代初,信息被引进的学术领域已不下 30 门。而至 90 年代中期,这个可观的数字已扩展至 50 余门。今天,在任何一门社会科学分支中,如果不想看到“信息”的出现已经是一件非常困难的事情,而在过去看来和信息概念完全不相干的物理、化学、生物、天文、地理和部分技术科学中,也随时能够见到“信息”概念的频繁出现。

作为对信息问题理论化的关注,一开始被学术界认为只是《信息论》和《控制论》的任务,到了 1959 年,美国宾夕法尼亚大学莫尔电子工程学院首次提出了“信息科学”的概念,它被用来代表一组计算机方面的课程^[1]。但此后的计算机学术界,并没有在信息本体或更大的范围上就信息的理论问题进行任何探讨,使其仅限于计算机科学可以用信息科学来指称的形式上。60 年代中期,计算机在文献检索试验方面闪露出的曙光,极大地刺激了图书馆学家们的兴趣,他们迅速地用“信息科学”这个概念来代表他们对文献信息的研究^[2]。于是,人们只好用“图书馆信息科学”(Library and Information Sciences)和“计算机信息科学”(Computer and Information Sciences)来区分上述两种所谓的“信息科学”。

由于信息概念在使用领域上的不断膨胀,使得人们一直希望出现一门真正的信息科学,它所关注的信息,既不仅是文献学家所说的那种,也不仅仅是计算机科学家所说的那种,它的范围应该更加广泛。于是,在 80 年代末至 90 年代初,一种朝着这个方向努力的第三种信息科学出现了^{[3][4]}。这种信息科学有时被称之为广义信息论,它不仅包含了申农信息论的内容,有时还包括信号处理方面的内容。更重要

的是,它试图把工程的、自然的、生物的和人类等类型的信息也作为它的研究范围。人们期望它会成为新的信息科学的代表,但由于这种信息科学仍旧把基础建立在申农的概率信息理论之上,而申农理论至少在解释细胞信息和人类信息方面是失败的,因而这种信息科学的基础从一开始就表现出一种脆弱性。如果我们把这种信息科学叫作概率论信息科学的话,那么,在今天看来,信息研究的历史上至少有三种信息科学出现,它们是:“计算机信息科学”、“图书馆信息科学”以及“概率论信息科学”。现在,每门“信息科学”都存在着,但从公理学的角度上说,都没有建立起令人信服的信息理论,这就是今日信息科学理论在全世界的现状。

信息在不同领域的应用,导致了它在不同领域里存在着特殊含义,结果就出现了信息的基本含义和特殊含义之分。总体上说,包括上述三种信息科学在内,目前对信息的研究可以分为两大类,一类是基于信息的基本含义之上的信息科学,一类是基于信息的特殊含义之上的信息科学。如果我们把第一种信息科学称之为统一信息科学,把第二种信息科学称之为部门信息科学的话,那么在信息研究领域里,部门信息科学和统一信息科学同时发展的状况,在本世纪末已是既成事实,并将会继续成为 21 世纪初叶,也许在更长一段时间里信息科学发展的风景线。

在各种各样的信息类型里,有一个共性是存在的,那就是每种信息都有它们的发出者和接受者。让我们把目光放在这些发出者和接受者之上。根据它们的不同,我们把发出者和接受者分为机械的(包括各种电子、光学、声学机构)、细胞的(目前主要集中在动物细胞上)、人类的(以人的整个大脑为基础)、分子的(目前主要局限于化学分子上)、动物的(以低等动物的大脑为基础)等类型,依此就有机械信息学、细胞信息学、人类信息学、化学信息学、动物信息学。下面让我们以发展得比较迅速的机械信息学、细胞信息学和人类信息学为例,对部门信息学进行一个示例性分析。

二、三个主要的部门信息科学分支

机械信息学 或称电子信息学,或称计算机与通信科学。它是信息科学群体中最为重要的分支之一,是专门研究信息机械中机械信息的存在、传输、变换、行为及处理规律的科学。什么是信息机械?按照机械学理论作出的解释,机械的类型共由三大类组成。它们是动力机械、物料机械和信息机械,分别做能量、物质和信息的处理和传输工作。动力机械把一种能量变换成另一种能量,物料机械将物料进行加工或位移,而信息机械的任务,则是将信息进行加工处理或进行传输。机械信息学是站在理论的角度研究信息机械中理论问题的,它既是应用科学的分

支,也是信息科学的分支。信息机械包括传统的音视频信息机械、通讯信息机械和计算信息机械(计算机)等。传统信息机械的主要目的在于信息的传输和再现,而计算机信息技术的主要目的在于信息的处理和再生。

信息机械工作原理服从于下述一些理论,它们构成了机械信息学的核心。第一,信息的统计理论。它的基础建立在信号的概率之上。信号是信息的物理承载者,就像电子、光学或声学元器件是信号的物理承载者一样。由于申农理论被无限制地渲染,人们往往把信号理论当作信息理论来看待。即便如此,随着计算机互联网的出现,存在了差不多半个世纪,以传统的电话为核心建立起来的申农信息量理论,也开始受到了以计算机互联网为基础的信息计量理论的挑战。事实证明,申农理论已不能胜任或不能完全胜任以计算机为基础的信息通讯分析。因为复杂的网络分层协议替代了简单的电话通讯规程,非即时通讯或准即时通讯替代了即时通讯^[5]。

信息机械工作原理依据的第二个理论是电子学或微电子学。电子学的历史开始于阴极射线管的发现并仍在继续发展着。它是一部记载着数学家、物理学家、工程师和发明家卓越贡献的历史,也是一部辛勤建立起来的完整的应用科学史。电子学致力于电子在真空、气体、液体、固体和等离子体中运动时产生的许多物理现象的规律研究,致力于电磁波在上述介质中传播时发生的物理效应,以及电子和电磁波相互作用的规律的发现。电子学作为科学门类之一具有鲜明的应用性。除了为信息机械的研制服务之外,电子学还为材料、能源的研究和开发服务。

也许我们应该把程序设计与人工智能作为机械信息学中的第三个分支,虽然传统的信息机械不受制于它,但随着智能化信息机械的普及,包含程序化操作的信息机械已随处可见。新的程序设计理论的出现也只是最近 10 余年的事情,它致力于使问题空间和求解空间在结构上尽可能地一致,它不再仅仅是数学、逻辑学和语言学的结合,它还包含了一种被称之为对象理论的程序哲学,因此人们把新的程序设计学称之为面向对象的程序设计理论。人工智能研究在经历了最初的高昂热情和激动之后,冰冷的现实逼迫人工智能研究者坐下来重新审视他们最初的理想,现在人工智能学家正在寻求新的突破点来从事这项研究。

细胞信息学 细胞信息学还有另外一些名称,分别是生物信息论、生物信息学、细胞通讯研究。生物信息论和生物信息学略有不同,生物信息论是信息概念被引进生物学后,人们创立的生物学和信息论的交叉学科,其研究范围主要包括神经科学和遗传学中的信息流问题。而近年来流行的生物信息学^[6]有时指传统的生物信息论内容,有时又指计算机在生物学中的应用。生物学是维纳推广信息运动时最早涉及的领域之一,但那时维纳关注的生物学信息

问题还主要局限于神经科学。而今天的生物学中,除了神经科学和遗传学外,内分泌学和免疫学也早已成为生物学中推广使用信息概念的两个新领域。

在同一生物体不同物理位置的多种相互影响方式之间,神经的作用是最重要的一种,而在神经细胞的内部,起着这种“二仙传道”的作俑者被认为是信息,这是神经科学家们早已公认的;遗传学是第二个引进信息概念的生物学领域。随着40年代基因研究的发展,从发现脱氧核糖核酸是遗传的物质载体,到20世纪末叶人工培育的“多莉”绵羊在英国的成功,信息是亲代和子代之间遗传基因的体现者,这也已家喻户晓;激素研究一直是内分泌学中的一个重要概念,今天的生物学家和医学家已开始把激素作为内分泌系统中传递信息的使者,从而起着接受信息的靶细胞和其他细胞之间的协调作用,结果能够使得内分泌系统正常工作;在免疫学研究中,传统的免疫概念主要指机体免除疾病的能力,但在今天的免疫学家看来,机体对病原体的免疫能力是由于机体具有识别某类特别细胞的功能。换句话说,免疫机能实质上是免疫细胞能否通过抗原正确地和其他细胞进行信息通讯的问题^[7]。这样,免疫学中的一些主要问题就变成了细胞之间的信息传递问题了。

从根本上说,生命体的各种信息过程,是生命体为适应其环境而存在的。从以上信息在生物学研究的应用中我们可以看到,细胞是信息存在和流动于其间的基本个体。也就是说,信息概念在生物学中的应用完全是以细胞为基础的,它们包括信息在信息源细胞、携带信息的配体、接受信息的受体细胞中的存在与表达方式,以及信息在这些细胞之间的传递规律等。所有这些问题构成了信息在生物学中应用的主要研究内容。尽管信息概念在上述四个生物学研究领域中还未来形成统一的细胞信息理论,但用细胞信息学来表达这个信息科学和生物学的交叉分支,仍不失为一个比较准确地反映生物学中信息问题实质所在的选择。

人类信息学 人类信息学的另一个流行名称叫信息传播学。它是研究信息在人类社会成员之间传递、传播和交流规律的信息科学分支。和机械信息学、细胞信息学相比,人类信息学所关注的信息,是迄今为止所有信息类型中人们使用得最多的一种,事实它上涵盖了信息的一切原始含义,它是消息、音信、新闻、资料、指标、数据、情报、命令、指示、知识、文学和艺术话语内容等的总称。一个人没有信息可以生活和工作是不可思议的;同样一个团体或一个国家没有有效的信息保障而能够有效地运转也是不可思议的。

人类信息学由三个部分组成,它们是:信息的本体研究、信息的内部过程研究和信息的外部过程研究。信息本体研究的主要内容是:人类信息的本质,人类信息的微观结构和宏观结构,信息是怎样在人类的认知、决策和审美中发生效用的。信息的内部过

程研究是人类信息学和认知心理学的交叉领地,内容包括:信息在大脑内部的表征、注意的信息机制、信息记忆和检索、信息思维、信息和语言的关系等。信息的外部过程研究包括信息在不同的人类群体,例如小组、组织、系统、一国之内或国与国之间的传播问题。由新闻学发展而来的传播学已经对这些问题进行了长时期的探索。因此,信息的外部过程研究也叫(人类)传播学。

在一个处于稳定态的社会系统内,社会的前进一般被认为是建立在以市场为基础、以信息为导向的机制之上。这就涉及到了人类信息学和其他学科交叉而形成的最为引人入胜的地带。1994年和1996年,诺贝尔经济学奖连续两年授给了信息经济学者,而信息经济学则是由信息科学和经济科学形成的交叉学科,这已或多或少地向人们昭示了现代社会发展的基本原理。1991年阿姆斯特丹第一届信息法学国际大会所表现出来的精神,正在预示着一门新的信息法学的问世。信息法不是仅由新闻法、出版法、广告法、著作权法、专利法、艺术法、电影法、电视法、电信法、网络法、个人隐私法、国家保密法、政府信息公开法等法规的汇集,它是一部建立在一切自由的信息行为必须在法律的保障和约束下进行的法典。经济法的目的在于调整经济活动中人和机构的经济行为,而信息法的目的则在于调整社会活动中人和机构的信息传播行为,这在信息社会中尤为重要。人类信息学中第三个重要的交叉学科是信息社会学。现在有人说,人们已经进入了信息社会,有人说这种情况只适于发达国家,这种争论需要由信息社会学理论提供标准来确定。信息社会学是信息传播学、历史学和社会学的交叉学科。此外,随着人们对信息化社会认识的不断加深,发现它并不是仅由计算机加电子通讯构成的技术拼盘,一个有效的信息化社会运转所需要的基础是一个有关信息与传播理论的复合体,它包括支持信息社会有效运行所需的理论建设、文化教育、政策法规、规章制度、道德观念、人才资源等,而对这些问题的分析和研究,不仅构成了人类信息学丰富的内涵和艰巨的使命,而且越来越多的迹象表明,人类信息学——信息与传播科学的发展,将是21世纪信息科学和人文与社会科学中一个双重亮丽的分支。

三、对21世纪信息科学发展的一些预测

对信息科学研究内容的确定,目前存在着两种完全不同的方法。一种是过程性方法,一种是规定性方法。过程性方法认为,信息科学是研究信息的获取、表述、存输、识别、编码、处理等传输过程中各种信息问题的科学。规定性方法则人为地规定,信息科学的研究内容包括计算机科学、电子与信息系统科学、自动化科学,有时还包括半导体科学和光学。过程性方法是理想的,但不实用,因为不同领域里的信

息过程可能是完全不相像的。规定性方法是实用的,但不尽科学。到目前为止,尚未见到力倡这种信息科学的研究者,为这种“信息科学”定义和为这种“信息科学”的学科体系提出理论依据。过程性方法多为学术界行为,规定性方法多为政府性行为。

进入 80 年代以来,信息研究的多元化已经被学术界所公认,这种见解已经从继《信息论》和《控制论》之后的第三部信息科学经典名著《信息研究:学科之间的讯息》^[8]中明白无误地传递出来,书中 9 个不同信息研究领域里的 39 位科学家已经坦承,“信息科学”的名称不应该为一个或某些学科所垄断,这就像量子物理学不能单一垄断物理学的名称、工业经济学不能单一垄断经济学的名称一样,尽管它们各自在自己的大学科中的地位是举足轻重的。

从科学界默认的传统来看,只有统一的信息科学才是真正的信息科学,而所有的统一科学都是关于全类对象的科学,所有的分支科学都是关于分支对象的科学。虽然计算机信息科学家、概率论信息科学家和图书馆信息科学家都对某类信息问题进行了研究,但还都没有真正立足于所有的信息对象之上,因此从统一信息科学的观点上看,上述研究将在 21 世纪逐渐淡出,并重新归类到不同的部门信息科学里去。

由于统一信息科学关注的信息对象分布在不同领域,所以要发现不同领域里信息现象的共同规律是很困难的。目前仅可以看到:在一切有机的或相关的多个体的群体中,个体之间均存在着相互影响或相互作用,这些影响或作用在本质上又不能仅归因于来自个体之间的物质或能量,那么这些个体之间就必然发生通讯或传播;有通讯或传播就必然有信息存在,有信息就必然有信息研究,从而有可能产生新的信息科学分支。这就是我们从近代信息研究中得到的启示,特别是从化学信息学^[9]的出现中总结出的信息科学发展规律。由于信息是和通讯或传播联系在一起的,而分析通讯与传播问题的首要条件是必须考虑两个以上的个体,所以凡是使用信息的地方,就必然使人们产生关于对象的联想或系统化思维,这就是信息概念的生命力和活力之所在。

但上述材料还不足以构成支持真正信息科学出现的充分依据。建立统一信息科学理论是项艰巨复杂的工作,它将象征人类对纷繁的信息世界的了解有了进一步的升华。同时它也是高智力的角斗场:对许多要予以抽象的分支需要有深刻的了解,而信息涉及的分支又那样的繁多,它涵盖着从电子学到哲学的自然科学、技术科学、人文科学、社会科学中的许多基本问题。喜欢挑战性工作的人也许会对它产生兴趣,但必然会冒一定风险。因为,在部门信息科学中,现在几乎还没有一门理论已成型或被公认,其中大部分还只是一个雏形或正处于探索之中。在诸多使用信息的不同领域,信息之间的含义相去甚远,甚至永远都是互不相干的。如果统一的信息科学找

不到足够的共同信息规律来承载起信息科学基础的话,信息科学很可能蜕化为一个群体研究的指称而不再被看作一门实体的科学。

就部门信息科学的发展来看,立足于不同领域里信息问题之上的部门信息科学在 21 世纪将会有较大的发展,因为大部分部门信息科学所涉及的信息问题都和人类的生产实践直接相关。而生产实践,也只有生产实践才是推动科学发展的真正动力。除了本文主要论及的机械信息学、细胞信息学和人类信息学之外,化学信息学也将会为人们展开一部用信息观点看待古老科学问题的新篇章。了解低等动物之间通讯的动物信息学也会有所进展,但肯定是小步伐的,因为这门学科还缺乏强烈的社会需求动力。上述五种部门信息科学可以视为标准部门信息科学,因为我们可以明显地看到信息传递于其中的信息发出者和信息接受者。而另外一些信息学,主要立足于认识论或工程基础之上,但却表现出一派生机。例如,全(信)息学、信息物理学以及正处于形成或完善之中的土地信息学、地理信息学、心电信息学、医疗信息学等。起源于生物学和医学研究的全息理论,今天已经扩大到了文化和经济等现象的研究之中。这门理论在新世纪也会有较大的发展,因为自然界和人类社会中的全息或半全息现象到处存在。信息物理学,如果它想从结构是一切物质的属性角度切入信息科学^[10],它还需要接受更加严峻的考验。

统一信息科学上的悲观主义,部门信息科学上的乐观主义,这就是我们对信息科学在进入 21 世纪时的总看法。

主要参考文献

- [1] Wellisch, Hans: From Information Science to Informatics, *Journal of Librarianship*, vol. 4, (July 1972), pp. 164-187
- [2] Saracevic, Tefko, ed.: *Introduction to Information Science*, New York: R. R. Bowker, 1970
- [3] 鲁晨光. 广义信息论. 中国科学技术大学出版社, 1983 年版
- [4] 钟义信. 信息科学原理. 北京邮电大学出版社, 1990 年版
- [5] 逯绍义, 王思明著. 计算机通信网络信息量理论. 电子工业出版社, 1997 年版
- [6] Schombarg, Dietmar: *Bioinformatics: From Nucleic Acids and Proteins to Cell Metabolism*, Berlin: Walter de Gruyter, 1994, p. 1
- [7] 尤复翰, 陆佩洪等. 细胞的通讯与细胞的防御. 江苏科学技术出版社, 1982 年版, 172
- [8] Machlup, Fritz; Mansfield, Una: *The Study of Information: Interdisciplinary Message*, New York: John Wiley & Sons, 1983
- [9] 宋心琦, 周金渭, 周福添著. 分子识别、化学信息学和化学反应智能化问题. 载 21 世纪 100 个科学难题. 吉林人民出版社, 1998 年版, 403
- [10] Stonier, Tom: *Information and the Internal Structure of the Universe — An Exploration into Information Physics*, London: Springer-Verlag, 1990, p. 17 (责任编辑 孙立明)