

从基于逻辑的人工智能到社会智能的发展^{*}

戴汝为

(中国科学院自动化研究所复杂系统与智能科学重点实验室, 北京 100080)



摘要:从思维科学和系统科学交叉发展的观点回顾人工智能半个世纪发展的历程: 它经历了以“物理符号系统”为假设的传统人工智能(逻辑思维的模拟); 人工神经网络为代表的分布式人工智能(形象思维的模拟); 现场人工智能(逻辑思维、形象思维与运行环境相结合的模拟); 智能体技术、人工社会研究所涌现的社会智能(社会思维和群体智慧与运行环境相结合的模拟)。钱学森于1990年发表了开放的复杂巨系统及处理这类系统有关复杂问题的方法论, 接着提出: 我们要研究的不是智能计算机, 是人机结合的智能系统, 以此为基础, 建立了信息空间综合集成研讨体系(Cyberspace for Workshop of Metasynthetic Engineering), 为社会智能涌现提供了可操作的技术平台^[1]。

关键词: 思维; 系统; 人工智能; 人一机结合; 社会智能

中图分类号: N94

文献标识码: A

Intelligence Development from the Logic Artificial
Intelligence to Social Intelligence

DAIRUwei

(Lab of Complex Systems and Intelligent Science Institute of Chinese Automation Beijing 100080, China)

Abstract: This paper summarizes the development of Artificial Intelligence(AI) from the viewpoint of Noetic sciences as well as System sciences as thus: Traditional AI with physical symbol system hypothesis (the simulation of logical thinking), Artificial Neural Networks representative of distributed AI (the simulation of thinking in imagery), Situation AI (the simulation of logical thinking and thinking in imagery in their environments), the Emergence of Social Intelligence from the research of agent technology and artificial society (the simulation of social thinking and collective wisdom in their environments). Qian Xuesen proposed a new discipline of science—the study of Open Complex Giant System and its methodology in 1990, and then pointed out what we study was not intelligent computer, we must pay attention to the human-centered man-computer cooperated intelligent system. On the basis of that, Cyberspace for workshop of Meta-synthetic engineering has been established to provide an operational intelligent platform from which social intelligence can be emerged.

Key words: thinking; system; artificial intelligence; human-computer cooperation; social intelligence

^{*}: 原载自然杂志, 2005, 28(6)。

收稿日期: 2006-12-22

基金项目: 国家自然科学基金重大项目: 支持宏观经济决策的人机结合综合集成体系研究(79990580)

作者简介: 戴汝为(1932-), 男, 云南昆明人, 研究员, 中国科学院院士, 研究方向为自动控制、模式识别、人工智能、复杂性科学及思维科学。

20世纪 40年代是个令人鼓舞的科技大发展时期,有两件对人类起着极其重大影响的事件。一是数字计算机的问世,二是控制论的发表。控制论这门关于动物和机器中控制与通讯的科学的发展,起初曾受到无知者的批判,但 CYBERNETICS(控制论)一词的前缀 CYBER 一至今被反复地加以应用。至于数字计算机,对人类所起的重大作用,则人所皆知。1956年,在美国达特莫茨,一些科学家经过讨论,把用计算机来模拟人的思维的研究,命名为人工智能(Artificial Intelligence简称 AI)。

智能与思维密切相关,关于思维的探讨与研究早就开始。钱学森高屋建瓴地提出了现代科学技术体系,把思维科学纳入其中。1980年,他发表文章说:“我们要把逻辑学扩大为思维学,包括一部分我们已经研究得很多而且很有成绩的逻辑思维,还要包括其它的人的思维过程。这在国外已逐步地引起重视,他们是从搞机器人、人工智能这方面考虑的;搞人工智能、机器人,就要搞一个人工智能、机器人的理论,这个理论他们叫认知科学(Cognitive Sciences),我们用‘思维科学(英文定名为 Noetic Sciences)’。更确切一点,就是包括逻辑思维,也包括各种思维过程,形象思维等^[2]”。

根据哲学家熊十力的观点,人的智慧包括性智与量智。钱学森把智慧与思维相结合展开为量智与逻辑思维相联系,而性智则包括了与形象思维联系的象智及多种思维综合集成的高层次性智,即大成智慧^[8]。下面从思维科学与系统科学的交叉研究来回顾人工智能的发展。

1 基于逻辑的人工智能一度辉煌

早期人工智能的起源是基于心理学对经验知识研究的结果,即发现启发式知识在人类思维过程的作用。这类知识表达成逻辑形式加以利用,这是人工智能最初的模型、称为“基于逻辑的心理模型”。传统的人工智能可以概括为符号表达、逻辑推理、启发式编程,或者称之为对“深思熟虑”的思维的模拟。在达特莫茨会议的一段时间内取得了一些引起人们关注的结果^[3]:

山依的下棋程序;纽厄尔、肖和司马贺的通用一般问题求解 GPS;鲁滨逊(Robinson)的归结原理;王浩的机器定理证明;尼尔森(Nilsson)的 A* 搜索算法;塞木尔的下棋程序;塞夫伯利支的小鬼模型;明斯基的语义信息处理等,这里不一一列出。

在上述成绩面前,一些醉心于人工智能发展远景的专家们,在远隔重洋的美国,在头脑发热的 1958年作出了种种乐观的预言,有的西方学者曾充满信心地说:不出 10年计算机要成为国际象棋冠军;不出 10年计算机将能谱写具有优秀作曲家水平的乐曲;不出 10年,大多数心理学理论将在计算机上形成。有些人工智能的专家甚至断言:照此趋势下去,20世纪 80年代将是全面实现人工智能的年代。到了公元 2000年计算机的智能就可以超过人了。

从人类思维的模拟来看,所完成的一些工作,可以说是对人的“量智”相关的逻辑思维的模拟与复制,其中最具有代表性的“一般问题求解”,推理在问题求解中起到核心作用,而推理中采用的搜索过程是启发式的。这项工作,及所提出的以数字计算机为原型的“物理符号系统假设”深刻地影响了整个人工智能的发展。当人们沿着这一途径继续进行工作时,所遇到的困难远远超过原先的想象^[4]。

2 模拟形象思维的人工神经网络及分布式并行处理

人们在人工智能的研究上又另辟新的途径,就是采用并行处理,及分布式表达的方法。在某种程度上,有意或无意地以模拟与复制形象思维相关的“象智”的人工神经网络方法一时间成为热门。

人工神经网络的特点是:1)以并行的方式处理;2)以分布方式表示信息。也就是用有若干个结点,每两个结点间可以连接起来的网络表示信息,以往用以表示知识的语义网络是一个结点与一个概念对应,而人工神经网络是以结点的一种分布模式以及加权量的大小与一个概念对应。这样即使每个结点上的信息属性发生了畸变与失真,也不致于使网络所表达的概念的属性产生重大的变化。另外有些共同的单元上的信息也可以用来表达相类似的概念,这就是在谈到人工神经网络时,人们所津津乐道的“并行信息分布处理”的精髓所在。比较著名的人工神经元网络模型有霍普费尔德(Hopfield)网络、反向传播(Back-Propagation)网

络、自适应共振理论(Adaptive Resonance Theory)网络等。在20世纪80年代,霍普费尔德的工作及反向传播理论受到关注。

人工智能的研究途径概括为模拟逻辑思维,以符号处理为核心的传统方法;及模拟形象思维,以人工神经网络联接为主的联接机制(Connectionism)方法。从模拟人的思维的角度来考虑,这是很自然的。人的两种主要思维方式是逻辑思维和形象思维(直感思维)。1985年钱学森即指出:形象思维非常重要,它是研究思维科学的突破口,因为一旦搞清楚,前科学的那一部分,别人很难学到的那些形成科学以前的知识、经验、感受等,都可以挖掘出来了,这样就把我们的智力开发大大地向前推进一步。关于形象思维虽然人们认识到它的重要性。但用现在的计算机来模拟形象思维是相当困难的,需要在计算机的体系结构上有新的突破。人们对网络模型结构抱有很大希望^[5]。

3 现场人工智能和现场认知

在有关系统的研究中,人们越来越清楚地认识到,系统的复杂程度一方面取决于系统本身,另一方面取决于系统的运行环境。1990年钱学森指出:“这是说人工智能是一项系统工程,要用多种学科,只有所谓‘人工智能’是远远不够的,要用控制论(讲反馈外部情况),用系统科学,还有思维科学^[3]”。

3.1 系统与运行环境一体的现场人工智能

美国MIT的年轻教授布鲁克斯在1991年的国际人工智能联合会议报告了他的研究成果“没有推理的智能”,研究了能在办公室里活动的人造动物。成为大会设立的“计算机与思维”奖的得主之一。他在研究中突出的4个概念:

1) 所处的情境(Situatedness)

机器人处于一定的世界之中。

2) 具体化(Embodiment)

机器人有躯干。

3) 智能(Intelligence)

智能是由于与周围进行交互的动态所决定。

4) 涌现(Emergence)

在考察系统时也要考虑系统运行的环境,所谓系统运行的“现场”,包括自然的与社会的两种情景。从系统与周围世界的交互以及有时候系统的部件间的交互涌现出智能。智能是某种复杂系统所涌现的性质,从而引导了现场人工智能的发展^[6]。

3.2 认知、思维与智能的交叉研究

认知科学是1983年由诺贝尔经济奖及计算机科学的图灵奖获得者司马贺在中国讲学期间引入中国^[7]。他在与中国学者的交流中了解到中国在开展思维科学的研究,他在致钱学森的信中说:如果两者的目标相似,在中国及国际方面飘扬起独一面旗帜,将大有所得,以便有完满研究讯息。司马贺对钱学森倡导的思维科学研究的突破点在形象思维以及创造性思维才是智慧的源泉,颇有同感,以致他在1995年获国际人工智能学会终生荣誉奖时在发表的演说:“不可言传”的解释——AI中的直感(Intuition)、顿悟(Insight)和灵感(Inspiration)话题中,专门谈了这些与智能密切有关的‘只可意会、不可言传’的现象及其模拟问题^[3],进一步表明了思维的多种类型对人工智能的引导作用。在钱学森的建议下,为了把认知科学和思维科学与智能研究结合起来,1992年我国在“攀登计划”中开展了认知科学若干前沿重大问题——思维与智能模拟的研究,取得了一定成果,从而为学科交叉发展做了相当重要的工作。

3.3 现场认知和智能涌现

认知的研究与人工智能的发展息息相关,现场人工智能提出后,形成了“现场认知理论(Situated Cognition)”,以斯坦福大学克兰西(Clancey)为代表的学者在现场认知领域进行了一些工作:例如新的机器人的设计;有关大脑的新的模型;用新的方法在各领域中使用计算机等。

如果将人的大脑看成一个系统,外部环境又是一个系统。系统与其运行的环境两者间的关系是不可分开的,这个人与环境组成的大系统通过一种动态自适应机制,而呈现出了一种包含了人与环境的相互融合的整体性质,这种性质是进化的,从适应性的观点来看这些是有助于揭示智能本质的一些工作。

人工智能的研究使人们掌握了智能模拟的技术,从思维科学得到启示可以研究人类思维能力的提高及智慧的产生,而现场认知更加突出了系统与其所处的环境之间交互的重要作用。对于人类而言,就是人和由人组成的社会之间交互、学习而涌现出了智慧。因此这种人类智能的社会性成为当前时代的特征^[8]。

4 人工智能与思维和系统研究相结合拓展

4.1 人机结合的智能系统

现场人工智能强调的是智能系统,例如研制一个智能机器人,着眼点在于机器人的各个部件之间的相互作用,由于部件间的交互作用,最终体现出机器人适应环境的能力,也就是“适应性”的体现。

20世纪90年代我国在研究智能计算机的方面付出了很大努力。但经过一阶段就吸取了教训。钱学森曾经在早期指出我们要研究的问题不是智能机,而是人与机器相结合的智能系统。不能把人排除在外,是一个人一机智能系统。在国外,人工智能专家莱纳特(Lenat)和有知识工程奠基人之称的费根包姆(Feigenbaum)在1991年也明确提出“人机结合作预测”是知识系统的“第二个纪元”。从这样的人机系统中将出现超人的智能和能力。在这阈值之外,有着我们如今无法想象的奇迹。

人一机系统认知过程的最大特点就是交互作用,人一机系统的概念随着系统的复杂性而体现其拓宽。人的社会性和集团化日益突显,机器的集群性和自组织功能得到发展,所以人一机系统这种交互作用的数量、范围是在迅速增长。社会化的人们之间扩大着相互影响;机器群组之间流动着日益增长的信息;知识组合选择、人群和机群之间感受和知识交相流动;定性和定量认识循环转化;这样巨大庞杂的数据、信息、知识、概念,彼此间在网络中流动、碰撞、选择组合直至综合集成。当这种交互作用由量变达到一定阈值后,就会有质的转化;当综合集成到一定新的层次后会有智能产生。诞生在开放复杂巨系统环境中的这种人一机系统,正是人和环境的交互作用,焕发了巨大认知能力,成为智慧涌现的不竭源泉^[3]。

4.2 智能体研究和智能的社会性

近年来,人工智能领域研究新概念、新手段不断涌现,其中“智能体”等研究与应用方兴未艾。当前以Internet为中心,智能分布式计算与知识发现的网络技术,高速宽带的数字移动智能通讯手段,营造了真实的、动态变化、相互联系却难以精确预测的复杂信息环境,因而智能体(Agent)技术以及所导致的理念应运而生,基于智能体的计算和面向智能体的程序设计成为软件开发的一项重大突破。智能体除其自主性、反应性、主动性之外拥有移动性、学习与适应性、交互及社会性,与多个智能体能构成一定的社会环境;从而表现出类似人的特性及其人类的社会属性。特别是多智能体系统环境除其“知识层”、“组织层”之外其“社会层”日益突出,其社会性和复杂性更加增强,有社会智能体之说。多智能体(Multi-Agent)系统中的社会性和自主性及协作性体现的更为鲜明,便于深入到问题的基础,重视集体行为,突出社会因素,体现为智能的社会性^[9]。

4.3 “人工社会”研究的启示

人工社会源于20世纪90年代初的“仿真社会”的研究,研究者们采用基于代理(即智能体)的建模和模拟方法,打破学科界限,从各种活动和现象的动态交互入手,综合地由个体的行为模型开始分析社会结构和群体规律。人工社会的研究必须有实际的预测,社会问题的研究是复杂和困难的,但是计算机和人工社会的方法为社会问题的研究提供了一线新的希望,同时也为开放的复杂巨系统理论及综合集成方法论的应用开辟了新的领域。

回顾从传统人工智能、现场人工智能,至所谓计算智能,从模拟人的逻辑思维、形象思维到依托强大的计算机和信息技术,从虚拟情境中演示人工社会,预示着人的智慧是发展的,智能研究必然融入社会,走向未来。

5 社会智能的涌现

5.1 社会思维和群体智慧

社会思维是指人作为社会整体对客观现实的认识,它是在整个社会时间、社会关系的基础上,无数个人思维和各种群体思维交互作用、多元复合的观念体系。社会思维是人作为社会主体的整体思维,从思维主体范围的角度来看,社会思维包括个人思维和群体思维。群体思维是以若干思维个体组成的群体作为思维主体形成系统的特有功能,从而产生单个个体所不能达到的整体思维能力。它给我们带来了新的扩大了的认识能力。

在群体进入这样的思维状态下,就会使思维能力大大提高,从而发挥其前所未有的水平,使思维的结果实现跨越,涌现出群体智慧。钱学森把社会思维学融入到从定性到定量的综合集成法,提炼出从定性到定量的综合集成研讨厅体系,他认为这是思维科学的一项应用技术,从而为实现群体智慧提供了方法论依据和基本的手段。

5.2 人文与科技相融合——社会智能的涌现

人类在21世纪所面临的许多困难,诸如能源的大量需求、环境污染、资源耗竭、人口膨胀等,单靠现在的科学成就是很不够的,必须走科技交叉发展的道路。多种学科在形成之始,是在还原论思想影响下,属于不同领域,辩证的系统论观点,通过综合集成的多种学科、交叉发展,最终表现为人文和科技交融,服务于自然和社会和谐。正如马克思所预言“自然科学往后会把关于人类的科学总括在自己下面,正如同人类的科学把自然科学总括在自己下面一样”。

从思维科学和系统科学来分析,人类智慧的发展可以概括为3个方面,1)对于人类的个体而言,擅长于形象思维易形成象智,象智与量智相结合形成高层次的性智;2)对于人的群体而言,综合个体的智慧着重与情境相联系所产生并通过多层次的学习所丰富的群体智慧等;3)对于人类的社会而言,社会思维激发群体智慧的涌现,在一定社会环境下形成与此相联系的社会智能。个人与其所处的环境所呈现出来的智慧,群体与环境交互所涌现出来的智能,进一步如果上升到人类社会,那么就必然要研究社会思维以及与其紧密相联系的社会智能了^[8]。

综合集成研讨厅是一项思维系统工程,也是处理开放的复杂巨系统相关复杂问题的可操作平台。基于智能技术的“信息空间综合集成研讨体系(Cyberspace for Workshop of Metasynthetic Engineering)”,它是一个高技术的“智能化空间”,可以面向网络实时跨平台操作,多种资源共享,形成一个知识生产和服务体系,从而用途广泛,功能强大,在支持科学决策和咨询上实现了思维系统工程^[10]。“研讨”是在激活了的社会思维状态下,产生的是集古今中外之大成的群体智慧,采用的是信息网络、多媒体现代技术。这就体现了以人为本,发挥了专家系统、机器系统、知识信息系统的综合优势,成为涌现社会智能的可操作的技术系统。正如钱学森所预言,这表明我们在系统科学、复杂性研究以涌现社会智能的实践上高一个层次,体现了我国科学家在为国家建设和社会服务方面所做出的努力。

参考文献:

- [1] 戴汝为.从工程控制论到综合集成研讨厅体系——纪念钱学森归国50周年[J].自然杂志,2005,27(6):366—370.
- [2] 钱学森.关于形象思维的一封信[Z].中国社会科学,1980.
- [3] 戴汝为.从现代科学技术体系看今后智能系统的工作[A].许国志.系统研究:纪念钱学森85寿辰论文集[C].杭州:浙江教育出版社,1996.143—156.
- [4] 戴汝为,王珏,田捷.智能系统的综合集成[M].杭州:浙江科学技术出版社,1995.
- [5] 戴汝为.第十一届国际人工智能联合会议简介[A].神经网络及其应用学术讨论会论文集[C].北京:第十一届国际人工智能联合会议,1989.
- [6] 戴汝为.人工智能发展的几个问题——1991年国际人工智能大会简介[J].模式识别与人工智能,1992,5(2):69—78.
- [7] 司马贺.人类的认知——思维的信息加工理论[M].北京:北京科学出版社,1986.
- [8] 戴汝为.社会智能科学[M].上海:上海交通大学出版社,2006.
- [9] 史忠植.智能主体及其应用[M].北京:北京科学出版社,2000.
- [10] 戴汝为.支持科学决策和咨询的技术——思维系统工程[J].中国工程科学,2005,7(1):17—20.