维普资讯 http://www.cqvip.com

计算机科学2003Vol.30№.7

人工智能的神经系统动力学融合表示模型研究*)

危 辉

(复旦大学计算机科学系,智能信息处理开放实验室,脑科学研究中心 上海200433)

Artificial Intelligence's Fusing Representation Model Based on Dynamics of Neural System

WEI Hui

(Department of Computer Science & Intelligent Information Processing Lab, Fudan University, Shanghai 200433)

Abstract The problem of representation is the most kernel and most key problem in artificial intelligence. The hypothesis of inconsequence of the lower order structure is our traditional choice, and use standing instead of standing for to implement representation. Stemming from computational cognitive neuroscience, thus based on neuron information processing, massive representation by neural network and cognitive behaviors' dynamics model of neural system, this research aims to build a common and consistent groundwork within neuroscience to explain kinds of intelligent behaviors, and fuse information, system and processing procedure together fluently. This research is great significant for the structure simulation method of AI and the probe of intelligence's neural mechanism.

Keywords Artificial intelligence, Dynamics of neural system, Representation

1. 引言:不能融合的传统人工智能模型

尽管以专家系统为代表的人工智能方法三十年来获得了很大的成功,但依然存在常识问题和知识获取的困难,造成这些问题的原因是忽视了人真实的知识结构中语义的依赖性和建构过程中的发展性和连续性。由于它们中的大多数被设计成针对特别的任务,仅对一类相对狭窄的问题是有效的,而不是基于对智能本质有一个系统认识的理论根基,因此期望使用同样的技术最终能在AI的其它领域获得同样的成功的想法被证明是将思维简单化和表面化了,这反映出当前人工智能研究方法论上的一个弊端:用已有的工具去解决好解决的问题。开发一个受限域的智能系统和通用的智能系统根本就是两回事。

让我们来看一个例子,这是一个在真实环境中的展现智能的行为:出租车司机接送乘客。在这个过程中司机可能会频繁地碰到以下问题:行车路线规划、交通信号与标志判别、路况判别、解答乘客的问询、导游/导购服务等,这些都是典型的人工智能问题。其中任何一种在不作任何简化的情况下都是一个复杂的问题求解,更何况这些任务往往是交织在一起的,如何顺利地过渡就涉及到信息流的融合问题。试想一个司机要完成它们,需要先验知识(如地图、交通规则与指挥手势、当地的风土人情、驾车技能、解题步骤等)、需要融合多通道的信息(如交通灯和道路的视觉刺激、喇叭和乘客语音的听觉刺激、驾车的本体感觉等)、需要把知识和经验与当前的情景天衣无缝地整合到一起,最终高效地演化出一个可行的解题方案。如果换成一个智能机器来充当司机,我们就会发现这个看似平常的行为所隐含的模拟智能的难度是我们当前理论与技术手段都无法克服的,因为不可想象把传统人工智能中解决它们的各种分立的方法,如自动驾驶、机器规划、模式识别、自然语言理解、推理、问题求解、搜索技术以及机器学习简单地集成在一起就能达到目的,这些现成的方法(单单孤立地来审视这些技术,就会发现它们还仅能应对简单与理想的情形,没有普适性)在理论依据、所面向领域、问题描述、求解机制、知识采纳、系统表示体制等方面存在巨大的差别。对人而言该例中的行为是一个平凡的智能任务,其中发生的认知操作是普遍存在的,所有知识被有效地提取与使用、子任务间的流畅切换、利用不同模态的线索为问题求解服务、大跨度的调度与精细衔接问题求解模式等,所有操作都依赖于知识和经验在记忆中按唯一的意义特征被无界、无差别地表示。

该项"人工智能的神经系统动力学融合表示模型"研究的哲学基础是"智能的物理主义",即认为智能(作为意识的一部分)不是什么虚无缥缈的东西,世界上没有什么纯精神的东西,一切思维活动都是遵守物质世界运动规律的脑活动对同样遵守物质世界运动规律的环境刺激的反映,任何对智能现象的描述如果不能深入到对智能进行解释,就不能作为思维的模型。不承认这一点就是唯心主义,就是形而上学,就不是一种科学的世界观。

2. 基于神经系统动力学的直接表达

通过前面的例子揭示出我们想探询的问题是"在认知科学的广泛范畴内,研究个体智能共同的基础或统一的认知结构,为人工模拟更为复杂的智能行为以适应更为复杂多变的应用环境做准备"。传统的符号化人工智能采用的是"再现智能"的方法,设计的模型往往是对一种以某种形式表现、应用于某一局部领域的智能现象的模拟,对模型的认知意义并不刻意追求。这种指导思想对智能就事论事,缺乏对其本质的一个深刻洞察,没有将此智能现象作为人所有智能的一部分放在人的心智体系中进行全面系统地考察、放在认知发展全过程的前因后果序列中考察、放在无法回避的认知约束下进行考察。我们已经不满足淤陷其中的形式推理,而转向实现更朴素和本源的基于语义的推理。要想根本性地提高人工智能系统的性能,实现向"解释智能"的转变就必须将人工智能放在

*)本研究获得复旦大学青年科学基金(项目批准号01-13)的资助。危 辉 副教授,博士后,研究方向:人工智能,认知科学。

维普资讯 http://www.cqvip.com

认知科学这个交叉学科的背景下进行修正，后者通过适当地使用还原和整合的策略说明高层次的智能现象如何突现。认知是包括知识获取、存储、回忆和使用的全部精神活动。如果对认知取信息加工的观点，那么认知的相关性就是认知结构在所加工信息的形态、加工流程、结构的组成、学习对结构的影响等方面的依赖关系。对这些相关性的研究对于知识获取、知识表示和知识应用的分阶段模型和联合模型的结构，以及构造过程研究具有重大意义。认知的相关性约束要求我们考虑智能的发生、发展和完善的过程，以不割裂的观点看待智能行为的模拟，这就是智能模型构造的系统观点。通过对认知发展过程的研究使我们能更清楚、全面地看到智能的本质，而不仅仅是它的外表。认知相关性要求和系统的智能观是建立在统一的认知结构的基础之上，这个结构的基础和统一的体现是对信息的直接表达，这是一种对知觉信息的表达，是对语义的直接表达，它是心智的计算理论中心智表达与心智计算的核心，通过基于神经元网络的直接表达的实现能够找到整体主义和还原主义、天赋论和建构论的结合点。语言是最具代表性的智能行为，对个人言语的计算结构研究能最大限度地体现认知心理学对于言语的信息加工算法、生理心理学对言语加工模型的体系结构、计算神经科学对于这个加工结构的实现、发展心理学对计算模型的进化作用。对言语计算结构的研究是认知相关性和智能模型构造的系统观点的实践。所以，表示是人工智能系统结构的核心，它依赖于一个真实的物理组织形式。然而表示问题与记忆问题密切相关，即如何从宏观的行为还原到微观的机理，正是这样一种生理心理学的观点，使我们认为神经系统动力学模型有望成为人工智能的新范型，因为它能够将不同的认知任务、不同类型与作用的知识、不同的信息加工过程、不同的信息加工结构在神经元信息处理层次上融合表示，成为天衣无缝的一体。

目前这方面的研究还较多地集中在微观层次[4,5,9,10,11,14]，比如单个神经元对信号的处理；还有一些是从神经元群的角度来讨论神经元网络对刺激的响应[6,12,15]，如对信息的表达；现在更有一些从信息加工的认知行为角度寻找神经生理基础的研究[3,7,8,13]，如形成视觉的模型等。

3. 研究目标、内容与方法

“人工智能的神经系统动力学融合表示模型”致力于从认知科学探询“How we know”的角度，而不仅仅是传统人工智能探询“What we know”的角度，以将人的智能看成是一个不可分割的系统整体为基本前提，为知识的获取、储存、回忆与使用找到一个共同的、基于神经系统动力学模型的表示基础，即使不同来源、不同领域、不同类型、不同形式、不同作用方式的知识能够融合成一个紧耦合的语义系统；使目的不同、过程不同、表现不同、部位不同的各种智能行为能够融合成一个高度共享、过渡自然流畅、紧密合作、能使全局最优的复合体；同一件认知任务的不同加工阶段能无缝地融合到一起。在这样一个建立在神经元信息处理、神经元集群信息表征、认知的动力学过程基础上的智能的解释模型上，通盘综合考虑人工智能中的一些典型问题，如语义问题、知识表示问题、串并行问题、符号与连接问题等。

“人工智能的神经系统动力学融合表示模型”研究围绕这样一个核心问题“人工智能类脑模型的核心----基于神经元信息处理机制的集群表示模型及其系统动力学行为”展开，涵盖从大粒度的心理功能向微粒度的单个神经元信号处理，包括以下一些子问题：

(1)单个神经元进行信号处理的物理与数学模型，包括信号经树突输入与经轴突输出这个过程中发生的处理过程，输入信号的采集范围与输出信号的投射范围，这样的分布方式与信息处理与表达的关系，神经元细胞体对信号的处理等；

(2)神经元网络的组成方式，如此组成在信息处理与表示上的意义，多个神经元级联连通对信号处理的意义，以及网络模型在结构上满足的数学约束，及其网络性能的数学分析；

(3)网络生成与发育过程中的学习行为，包括作用于整个智能系统的、有意义的高丰度刺激如何分解为作用于局部网络或单个神经元个体的、低维的信号，网络或单个神经元特性的学习可塑性；

(4)基于记忆的神经机制的信息表达模型，包括知觉信息表达的神经元网络模型，它如何体现为人工智能中的语义表达，这一直接表达方式的连贯性与完整性，以及它与人工智能中物理符号系统假设的实质关系；

(5)各感觉通道信息在前述信息表达模型上的融合表达；

(6)神经元网络的系统动力学过程，这个过程的不同模式是不同智能行为的基础；

(7)信息表达模型进行记忆与回忆的实验，实验结果与日常记忆现象的对照及其结果分析，分析结论用于对模型进行修正；

(8)联想记忆的形成与实现机理，以及这样一种神经元网络的反应过程对问题求解、推理、综合等人工智能问题实现方式的影响；

(9)人工智能中的知识表达结构化模型所对应的基于记忆生理过程的解释与重构；

(10)其它人工智能问题，比如模式识别、计算机视觉、语音识别、自然语言理解、问题求解、形象思维、规划、机器学习等，如何统一到这个核心表达模型上来，即一个统一的认知结构。

“人工智能的神经系统动力学融合表示模型”研究的关键在于解决这样三个问题：一是基础，即神经系统动力学过程赖以发生的单个神经元的信息处理机制，由神经元的生物模型向物理模型转变，包括神经元的结构模型与信息处理的数学模型[16,17]；二是发展，神经元集群是如何连接的，如此连接对信息表示与处理的意义，以及这样的结构如何受到学习的进一步影响[18]；三是效应，神经系统动力学行为的不同如何对应于不同的智能表现。

由于本研究是一个涉及多学科的交叉性研究，涉及认知科学的各个方面，包括认知心理学、心理生理学、发展心理学、神经生理学、神经心理学、人工智能等，但它们分别是在不同的层次研究人的智能实质。但是由于存在太多的未知，目前还缺乏将所有这些层次完全贯穿起来的科学基础(脑高级功能的细胞与分子基础)，因此它们提供的理论或是太抽象和笼统(如学习的行为学表现)，或是太集中于细微之处(如神经递质对脑功能的调控)。这也就决定了我们的研究方法不能偏向它们任何一方所采用的典型研究方法，否则就偏离了计算机这个现实条件，因此我们希望在借鉴可用的成熟理论同时，通过研究与合理假设弥补不足的部分。我们的目标是力求构造一个尽量接近客观生物视觉的机器认知模型，但它的结构与算法必须同时满足可在计算机上实现的要求。所以在研究方法上我们更多地采取信息加工的认知观点，走“解释智能”的道路，采取仿生学的方法。

维普资讯 http://www.cqvip.com

人的各种智能表现在信息加工的心理功能上都以记忆与学习为核心，在生理基础上都可还原为神经元对信号进行加工的物理与化学过程，及其连通网络从宏观上的整合。因此我们认为人工智能类脑模型的核心是基于神经元信息处理机制的集群表示模型及其系统的动力学行为。所以我们将围绕记忆与学习开展研究，并逐步拓展。

我们需要将宏观模型和模型的实现基础作一个区别，上层模型(整体)必然更靠近心理学模型，下层实现(细节)必然更靠近生理学模型，它们间的过渡与衔接是一个心理生理学的示例。记忆机制是融合表示问题的关键，我们先考察记忆与神经元间信息传递的生理机制；再设计基于神经元学说的记忆表达模型，包括网络的结构模型、模型的数学描述、动力学过程的数学模型等；再设计基于结构特征的记忆编码算法；然后进行包括稳定性、分辨率等指标在内的网络性能分析；最后以得到的模型为核心，逐步建立对人工智能各类问题的解释。

参考文献

1 Kirsh D. Foundations of AI: the big issues. Artificial Intelligence, 1991,47:4～10

2 危辉. 认知相关性与智能模型构造的系统观点:[博士后研究工作报告]. 杭州:浙江大学研究生院,2000

3 Taddei-Ferretti C, Musio C. Downward processes in the perception representation mechanisms. Singapore: World Scientific Publishing, 1998

4 Koch C. Biophysics of computation: information processing in single neurons. New York: Oxford University Press, 1999

5 Burdet G, Combe P, Parodi O. Neuronal information processing: from biological data to modeling and applications. Singapore: World Scientific Publishing, 1999

6 Riegler A, Peschl M, von Stein A. Understanding representation in the cognitive sciences: Dose representation need reality?, New York: Kluwer Academic / Plenum Publishers, 1999

7 Uttal W R, et al. Computational modeling of vision: the role of combination. New York: Marcel Dekker Inc., 1999

8 Squire L R, Kosslyn S M. Findings and current opinion in cognition neuroscience. Cambridge, Massachusetts, The MIT Press, 1998

9 Gordon L. Fain, Molecular and cellular physiology of neurons. Cambridge, Massachusetts, Harvard University Press, 1999

10 Bower J M. Computational neuroscience: trends in research. New York: Plenum Press, 1997

11 Bower J M. Computational neuroscience: trends in research. New York: Plenum Press, 1998

12 Quinlan P T. Structural change and development in real and artificial neural networks. Neural Networks, 1998, 11: 577～599

13 Haken H. Synergetic computers and cognition: a top-down approach to neural nets. Heidelberg, Springer-Verlag, 1991

14 Levitan I B, Kaczmarek L K. The neuron cell and molecular biology(2nd Edition). New York: Oxford University Press, 1997

15 Longstaff A. Instant notes in neuroscience. London: BIOS Scientific Publishers Limited, 2000

16 危辉, 何新贵. 表象式直接知识表示. 计算机学报, 2001, 24(8): 891～896

17 危辉, 潘云鹤. 从知识表示到表示: 人工智能认识论上的进步. 计算机研究与发展, 2000, 37(7): 819～825

18 危辉. 一种基于结构学习和迭代自映射的自联想记忆模型. 软件学报, 2002, 13(3): 438～446

(上接第131页)

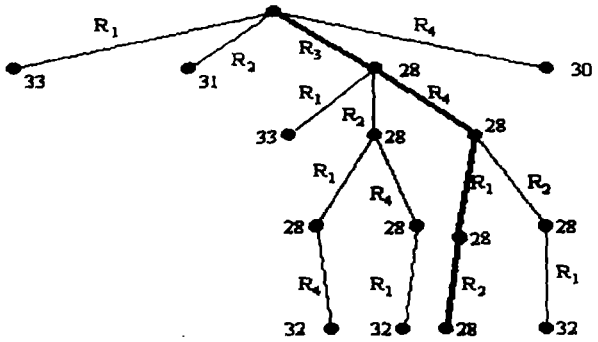


图2 分支定界分析

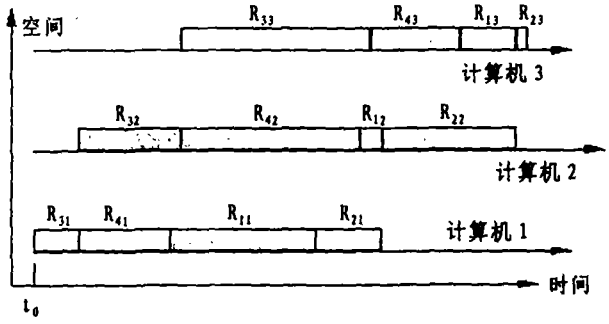


图3 另一种调度方案

顺序，可以获得最佳的效果。把处理时间从原来的40个单位时间减少到28个单位时间。在系统中实际运行的情况见图3，可见提高了系统的并行性。

研究网络并行处理的体系结构和实现技术，是今后计算机应用的一个重要方向，包括多线程的体系结构和预取、后存技术，在网络计算环境下算法的高效实现问题。

参考文献

1 Pritchard J. COM and CORBA Side by Side Architecures, Strategies, and Implementations Addison Wesley Longman. Inc. 1999

2 Wu Jie. Distributed System Design. CRC Press LLC, 1999

3 Buyya R. High Performance Cluster Computing, Volumeland 2, Prentice-hall, Inc. 1999

4 胡金初. 分布式处理在实时应用系统中的实现. 计算机工程, 2002, 12