

类脑智能研究现状与发展思考^{*}



徐 波^{1,2**} 刘成林^{1,2} 曾 毅^{1,2}

1 中国科学院自动化研究所 北京 100190

2 中国科学院脑科学与智能技术卓越创新中心 上海 200031

摘要 近年来人工智能研究的许多重要进展反映了一个趋势：来自脑科学的启发，即使是局部的借鉴都能够有效地提升现有人工智能模型与系统的智能水平。然而，想要真正逼近乃至超越人类水平的人工智能，还需要对脑信息处理机制更为深入的研究和借鉴。类脑智能研究的目标就是通过借鉴脑神经结构及信息处理机制，实现机制类脑、行为类人的下一代人工智能系统。文章从受脑启发的新一代人工神经网络、基于记忆、注意和推理的认知功能模型、基于生物脉冲神经网络的多脑区协同认知计算模型等角度，并结合研究团队在类脑智能领域的研究进展，论述类脑智能的研究进展、发展方向和对未来发展的思考。

关键词 类脑智能，人工神经网络，记忆，注意和推理，脉冲神经网络，多脑区协同，自主学习

DOI 10.16418/j.issn.1000-3045.2016.07.008

人工智能学科自诞生之初便奠定了其模拟、延伸、扩展人类智能的宏伟目标。60年来人工智能发展取得了巨大的成就，但是离学科诞生之初提出的实现人类水平智能的目标仍然有很大的距离。以往人工智能的研究成果大多数属于行为尺度模拟部分智能的计算模型。这是因为，一方面传统人工智能的主要研究者来自信息科学领域，另一方面脑科学研究以前还很难支持从更深入的机制上探索智能的本质及其计算实现的机理。近年来，随着脑与神经科学领域新技术的不断涌现，以及人工智能所依赖的深度学习、计算能力和数据发展，可以在更深刻的层面支持人工智能研究者对智能本质的探索。类脑智能研究逐步引起学术界的广泛注意，其核心是受脑启发构建机制类脑、行为类人的类脑智能计算模型。

本文将结合中科院自动化所“十一五”末以来在类脑智能研究方向的布局和研究实践，从脑与神经科学对人工智能的潜在启发、新一代人工神经网络、记忆认知功能模型以及脑区协同认知模型等角度概述近年来类脑研究关联现状与发展，并对类脑智能研究发展提出思考。

^{*}资助项目：中科院战略性
先导科技专项项目（B类）
（XDB0200000）

^{**} 通讯作者

修改稿收到日期：2016年6
月9日

1 脑与神经科学对人工智能的潜在启发

脑与神经科学的进展，特别是借助新技术与新设备的研究支持研究者通过不同的实验方法（如生物解剖、电生理信号采集与分析、光遗传技术、分子病毒学、功能影像分析等）得到对脑的多尺度、多类型的生物证据，正在尝试从不同侧面来揭示生物智能的结构和功能基础。从微观神经元、突触工作机制及其特性，到介观网络连接模式，再到宏观脑区间的链路及其协同特性，这些实验及机理研究得到的有关脑的多尺度结构与功能机制将对构建未来类脑智能计算模型提供重要启发。

在微观层面，生物神经元和突触的类型、数目等在不同脑区中具有较大差异，且能够根据任务的复杂性实现结构和功能的动态适应。现有实验结果表明，兴奋性神经元在前馈神经网络的应用中表现出较好的分类效果^[1]，而更加结构多样和功能复杂的抑制性神经元由于计算资源和学习方法的限制，尚未能像兴奋性神经元一样在人工神经网络的训练和学习过程中展示出应有的潜力，这将是未来值得探索的重要研究方向。突触方面，如时序依赖的突触可塑性（Spike-Timing Dependent Plasticity, STDP）是一类时序依赖的连接权重学习规则，突触权值的变化主要依赖于细胞放电发生于突触前神经元和突触后神经元的先后时刻，通过对放电时间差与权重更新建立数学映射关系，来描述网络中的神经连接强度的变化情况。该原则的生物基础已经在众多的生物实验中被证实，可以分为二相STDP、三相STDP，以及部分类STDP机制，如电压依赖的STDP等^[2]。

在介观层面，特异性的脑区内部的连接模式和随机性的网络背景噪声的有效融合，使得生物神经网络在保持了特定的网络功能的同时，兼顾了动态的网络可塑性。例如生物神经网络中的泊松背景噪声对生物神经网络的学习和训练过程起到极大的促进作用^[3]。此外，脑与神经科学研究者普遍认为神经元连接构成的网络结构对认知功能的实现具有决定性的支撑作用^[4]。神经元之间构

成的网络基序及基序结构的组合对神经信息处理过程也发挥着决定性作用。由于神经元类型的不同，使得神经元之间的网络连接更为复杂。例如实验表明，有些神经元倾向于与同类型的神经元相连接，有些神经元倾向于与其他类型的神经元连接，而有些神经元则只与其他类型的神经元连接^[5]。此外还发现，认知功能相近的神经元更容易形成突触连接^[6]。值得思考的问题是：不同的连接模式对应的功能差异是什么？对于认知功能的实现具有何种意义？实践表明这些结论都对未来类脑神经网络的设计有重要的潜在启发。

在宏观层面，不同脑区之间的协同使得高度智能的类人认知功能得以实现。如哺乳动物脑的强化学习认知功能，长时、短时记忆功能等都是通过不同脑区功能的协同实现更为复杂的认知功能。脑区之间的连接不仅决定信号的传递，而且反映了信息处理的机制。如脑区之间的前馈连接可能反映了信息的逐层抽象机制，而反馈连接则反映了相对抽象的高层信号对低层信号的指导或影响。此外，有些脑区负责融合来自不同脑区的信号，从而使对客观对象的认识更为全面（如颞极对多模态感知信号的融合），而有些脑区在接收到若干脑区的输入后则负责在问题求解的过程中屏蔽来自问题无关脑区的信号。

要实现人类水平的智能，需要计算模型能够融合来自微观、介观、宏观多尺度脑结构和信息处理机制的启发。实现跨尺度机制的融合，才能够实质性颠覆现有计算模型，实现有深远影响的原始性创新。

2 新一代人工神经网络模型研究

虽然传统的人工神经网络在神经元、突触连接等方面初步借鉴了脑神经系统在微观尺度的概念和结构，但是在信息处理机制上真正从脑科学借鉴的机制并不深刻。近年来发展起来的深度神经网络（Deep Neural Network, DNN）模型抓住了人脑在脑区尺度进行层次化信息处理的机制，在计算和智能模拟能力上取得重要突破，并在模式识别和人工智能应用领域取得了巨大成功。

卷积神经网络 (Convolutional Neural Network, CNN) 作为 DNN 的一种, 是受生物视觉系统的启示, 将生物神经元之间的局部连接关系 (局部感受野) 以及信息处理的层级结构应用到计算模型中^[7]: 当具有相同参数的神经元应用到前一层的不同位置时, 可以获取具有某种不变性的特征。CNN 在从低层到高层的过程中感受野越来越大, 逐渐模拟了低级的 V1 区提取边缘特征, 再到 V2 区的形状或者目标的部分等, 再到更高层的 V4、IT 区等, 高层的特征是低层特征的组合, 从低层到高层的特征表示越来越抽象。研究者详细对比了 DNN 的高层与灵长类动物 IT 区在物体识别任务中的关系^[8,9], 发现 DNN 的高层能够很好地反映出 IT 区的物体识别特性, 证实了 DNN 与生物视觉系统在某种程度上的相似性。

基于 CNN 的深度学习方法在视觉、语音领域的诸多任务中均取得突破性的进展, 其端到端的建模和学习能力颠覆了传统的“特征+分类器学习”的固有模式, 使得特征和分类器不再有明确的界限, 它们均能在 CNN 中一体化地学习。

神经科学研究已证明脑皮层中反馈神经元连接比前馈多得多, 但反馈的神经机制和作用尚待深入研究。传统的深度神经网络模型里一般只有前馈连接, 尚缺乏对反馈的建模。为了更好地模拟人脑, 最近有很多研究探索如何将反馈引入神经网络模型。例如, Liang 等人^[10,11]提出在 CNN 的卷积层加上层内连接的方法, 使每个单元可同时接收前馈和反馈的输入; Wang 等人^[12]通过 top-down 的反馈连接和乘法机制引入注意力模型; Cao 等人^[13]在 CNN 的卷积层加上层间的反馈连接, 将高级的语义和全局信息传到下层, 通过语义标签的反馈, 可以激活特定的与目标语义相关的神经元, 从而实现自顶向下的视觉注意, 定位复杂背景中的潜在目标。

强化学习是人类通过交互进行学习的重要方式之一。Google DeepMind 团队将深度神经网络与强化学习融合^[14], 构建的深度强化学习模型可用于自动学习打 49 种电脑游戏。该模型借鉴了人类通过与环境进行交互, 自

动学习选择最佳策略、采取最佳行动的学习过程。在此基础上开发的 AlphaGo 在与韩国围棋棋手李世石的交战中以 4:1 的成绩获得胜利^[15]。该团队巧妙地让两个围棋机器人互相切磋棋艺, 在交互中采用强化学习, 对策略网络和价值网络进行更新。强化学习在一定程度上解决了深度学习的大数据依赖问题, 如在上述例子中, 人类棋谱的数量是有限的, 而强化学习产生了更多可供学习的棋局。此项工作还说明了不同功能网络的协同, 能够提升智能系统的水平。

虽然现有 DNN 模型从某种尺度初步借鉴了人脑信息处理的部分原理, 但总体而言依然是初步的尝试。对脑信息处理机制的深度借鉴来提升现有模型仍然具有很大空间。目前在神经元的类型、突触的类型及其工作机理、网络权重更新、网络背景噪声等方面, 神经生物学的研究都取得了可以被计算模型应用的进展。许多研究团队已经构建了一系列满足不同尺度生物实验证据和约束的计算模型, 如生物神经元模型^[16-18]、生物突触模型^[19,20]、生物脉冲神经网络计算模型^[21]。这些都为未来脉冲神经网络 (Spiking Neural Network, SNN) 的进一步研究奠定了坚实的基础, 提供了创新源泉。人脑的神经系统存在很多反馈连接, 例如自顶向下的视觉注意就是来自于从高级认知脑区 (如 PFC、LIP 等) 的脑活动到初级视觉脑区的反馈信号, 现有模型中虽然有些神经网络引入了反馈的概念^[10,11], 但是反馈如何影响低层的输入信号以及跨层的反馈等, 模型中并没有深入考虑。此外, DNN 的领域特异性强, 扩展和泛化能力相对较差, 不同领域之间很难实现知识共享, 而人类不同的感知模态之间存在着很强的相互作用, 不同模态的知识能够很好地共享。

通过与脑科学的紧密联系和深度交叉, 构建更加类脑的神经网络计算模型和学习方法是类脑智能的研究方向之一。

3 基于记忆、推理和注意的认知功能研究

认知体系结构研究是类脑认知计算模型研究发展历

程上代表性的方向之一，其中最具代表性的成果是思维的自适应控制（Adaptive Control of Thoughts）认知体系结构。该模型受到了从人工智能视角研究认知科学的早期实践者之一 Allen Newell 思想的影响^[22]。该模型的特点是覆盖了感知、决策、语言、运动等广泛的认知功能，模块在行为层面模拟了对应的脑区功能，并具备坚实的认知心理学依据。然而该模型全部采用产生式实现认知功能，本质上是一个规则系统，且认知功能间的协同并非自组织。由于其计算模型没有采用神经网络，其介观和微观实现机理与人脑还存在较大区别，且很难处理大规模现实世界问题。

近年来深度学习除了在视觉、语音信息处理领域的长足进展与成功应用外，在自然语言处理领域也取得了重要突破。不仅在文本语义特征表示方面具有明显的优势^[23,24]，在关系分类^[25]、情感识别^[26]和信息检索^[27]等应用任务中也优于基于人工构造特征的传统方法。记忆（Memory）、推理（Reasoning）和注意（Attention）等机制逐渐成为神经网络领域的新研究热点。

2015年12月，由Facebook人工智能研究院的Jason Weston 牵头在人工智能领域顶级会议 NIPS2015上组织了一项关于记忆、推理和注意机制的研讨会，简称RAM（Memory, Reasoning, Attention）研讨会^[28]。参会者讨论认为，合理采用记忆、推理和注意机制，可以有效地解决人工智能的很多核心问题，并提出目前主要注意的问题。（1）记忆单元中存储哪些内容？（2）神经记忆单元中记忆的表达形式？（3）记忆单元规模较大时如何进行快速语义激活？（4）如何构建层次化记忆结构？（5）如何进行层次化信息推理？（6）如何对冗余信息进行遗忘或压缩处理？（7）如何评价系统的推理和理解能力？（8）如何从人类或动物记忆机制中获得启发？^[29]

Chaudhuri 和 Fiete^[29]指出：“记忆”是指神经系统中任何活动或连接变化的总称，这种变化由信号刺激或大脑状态触发并且持续时间要长于触发时间。“记忆”具有自适应性，具有记忆单元的智能体可以从经验中学

习，概括能力更佳，可以利用先验信息在不完整数据中进行更好地推理和预测。传统方法将记忆分为“短时记忆”和“长时记忆”。从认知神经科学角度看，短时记忆是由刺激产生神经元状态的持续性变化，而长时记忆是指神经元之间突触的连接和强度变化^[29]；从智能应用角度看，短时记忆由当前环境数据产生的状态编码更新和存储，而长时记忆是对历史信息进行高度经验性概括的编码，如概念、实体和结构化知识的表示。

早期利用神经网络对信息进行编码记忆的模型注重于神经记忆单元的结构化设计，如在端到端编解码模型中采用长短期记忆（LSTM）单元^[30,31]。该记忆单元采用累加器和门控神经元进行记忆细胞信息编码的更新，其中遗忘门控神经元决定细胞丢弃哪些编码信息，输入门控神经元确定当前时刻要更新的信息，状态更新累加器神经元由遗忘门和输入门共同更新细胞状态，而输出门控神经元则决定了当前时刻细胞的输出状态。目前基于LSTM单元或其他相关改进记忆单元已较好地应用于各种编解码任务，如对话系统^[32]。

然而，神经记忆模型将外部输入信息压缩到固定长度的向量化编码中存在较大的局限性。当外部信息量较大时，记忆编码容易丢失细节信息，使得智能系统的语义解析能力变弱^[33,34]。针对此问题，一种解决方案是增大记忆单元的编码维度，如在机器翻译任务中采用4 000 维以上神经元^[33]。该方法大大提高了神经网络模型的计算复杂度。

近两年，几位研究者从不同角度出发分别提出了采用非定长记忆单元和注意机制进行信息动态提取和融合的解决方案^[33,35,36]。其中，非定长记忆编码是一种由输入信息长度动态调整记忆编码长度的模式，可改善传统编解码记忆模型中将输入信息全部压缩到一个固定长度编码的局限性。采用非定长记忆编码机制后，编码端的记忆编码量有所增加，若全部作为解码端的输入，反而会因为语义信息冗余而获得较差性能。为此，研究者们相应地提出了注意机制进行动态地编码信息融合。注意机制是由外部刺激

引发注意转变,从环境和非定长记忆编码单元中选择重要信息进行融合,得到当前时刻刺激下有限长度的语义向量。另外,在一些自动问答等任务中往往需要通过推理才能从记忆单元中挖掘到有用的语义信息,一些研究者又采用多轮注意机制迭代的方式从记忆单元中逐次激活语义信息完成推理,找到目标信息^[35,37]。目前相关工作已分别在机器翻译^[33]、自动问答^[35]和图灵机模拟^[36]等任务中取得了良好表现。

上述机制不但在自然语言处理领域存在深入研究和广泛应用的潜力,其中某些机制在其他认知功能的建模方面也已取得了初步成果。例如注意机制在视觉信息处理的研究中已成为一个重要分支。人类视觉系统能够在复杂场景中迅速地将注意力集中在显著的视觉对象上,这个过程称之为视觉选择性注意。人类的视觉注意过程包括两个方面:由刺激驱动的自下而上的视觉注意过程和由任务驱动的自上而下的视觉注意过程。目前的视觉注意计算模型集中在对自下而上的注意机制的建模,其中 Itti 的显著图模型^[38,39]最具有代表性,该模型将多种尺度下的多种特征通过中心-周边算子得到显著性度量结果。相对于自下而上的视觉注意计算模型,自上而下的视觉注意过程研究相对较少,其中的工作可以追溯到 Yarbus 等人^[40],他们发现实验者观察照片时的眼动模式与被问的问题有关。有研究还发现大部分的注视点(fixation)集中在与任务相关的位置上。在自上而下的视觉注意计算模型方面主要的工作来自 Itti 研究组^[41-43]。

神经网络计算模型对推理、注意、记忆的借鉴虽然已在应用层面取得了很好的效果,然而进一步的发展还需要更为深刻地在机制与原理层面受脑研究的启发。例如在推理方面,认知神经科学的研究就分别探讨了不同形式的归纳推理、演绎推理的神经机制^[44,45]。而对于记忆的探索,脑与神经科学研究中也从不同视角对记忆进行分门别类的研究,如发现支持长时记忆、短时记忆、工作记忆、陈述性记忆、程序性记忆等不同类型的记忆的神经环路和神经机制上都具有较大差别^[46]。对这些脑机制

的深入探索与应用必将为神经网络计算模型的进一步发展带来新的契机。

4 基于脉冲神经网络的脑区协同认知模型研究

虽然来自神经科学的启发已在人工神经网络学习方面显现出巨大潜力,但是仅借鉴神经网络的学习机制还不能够完全支持实现类似于人脑的通用智能能力,其原因在于人类的认知功能多种多样,真正通用的类脑智能研究需要融合不同认知功能对应神经机制和认知行为机制的启发,构建面向通用智能的类脑认知计算模型。

从神经网络出发的脑区协同模型代表性工作是2012年加拿大滑铁卢大学研制的 SPAUN 脑模拟器^[47]。SPAUN 将 250 万个神经元模块化地分割组织为 10 余个脑区,通过构建不同的工作流实现了模拟笔迹、逻辑填空、工作记忆、视觉信息处理等能力。SPAUN 采用了简化的 SNN,通过脑区之间模块化的组织实现了特定认知功能的初步建模。虽然相对其他模型而言,SPAUN 已经部分接近真实大脑工作原理,但是该模型仍然具有极大的提升空间,主要表现在 3 个方面:(1)目前 SPAUN 感知区域全部用深度网络代替,仍然是对感知功能粗略的建模;(2)SPAUN 针对不同的认知任务绘制不同的工作流,不能自主决策任务的类型,不能对任务进行自主建模;(3)SPAUN 脑区之间的连接是逻辑连接,没有真正采用生物脑的约束,没有实质性地通过借助脑区之间的各种连接、脑的工作机制来提升智能水平。

人类多数高级认知功能的实现都与脑皮层密切相关,许多认知计算模型的研究希望构建通用的皮层计算模型,并以此突破通用智能。其中分层时序记忆(Hierarchical Temporal Memory, HTM)模型便是其中的代表性工作^[48]。该模型受皮层组织及皮层信息处理机制启发,实现了对皮层的不同层次功能的初步建模。该模型整合了空间和时间编码,与人类认知功能更接近,可实现时序数据信息抽取和模式预测,目前已广泛应用于多种智能数据处理领域,如交通流量预测、股票交易数据

预测、服务器流量预测等。然而，该模型对皮层层次所对应功能只是进行了粗略对应，与真实皮层结构仍有较大差距，且模型是对皮层微柱的建模，没有脑区层次的启发和协同机制，也没有自动问题建模的能力，认知功能的实现还无法达到自组织。

在构建类脑认知模型中，脉冲神经网络（Spiking Neural Network, SNN）被认为是能接近仿生机制的神经网络模型，其模型基础和运算方式与哺乳动物脑更为接近^[49]。与传统神经网络模型最大的区别是：SNN的神经元以电脉冲的形式对信息进行编码，这更接近真实神经元对信息的编码方式。由于采用了此种方式进行编码，使得SNN能够很好地编码时间信息。考虑到DNN所忽略的很多生物规则可能恰恰是实现类脑智能的关键（如对于时间的编码、抑制性神经元在网络中的特殊作用等），将这些生物规则加入到SNN中将有可能使现有网络获得更强大的计算和任务适应能力，这将是提升类脑神经网络模型最有效的手段与研究方向之一。

为了实现面向通用智能的类脑智能计算模型，中科院脑科学与智能技术卓越创新中心多个团队协同，围绕SNN自身模型的优化和多脑区系统模型^[47,50]（图1）等，正在进行类脑认知计算模型的研制。在SNN优化方面，将受生物现象启发的部分学习规则加入到SNN的学习框架中来，以尝试提升现有的SNN模型，并初见成效^[51]。面向学习与记忆，构建了前额叶与海马区记忆模拟系统，面向感知-决策构建了以基底神经节若干脑区为核心的类脑强化学习系统。这些系统在微观层面采用全脉冲神经网络实现，并引入了兴奋性、抑制性神经元模型、不同类型的突触计算模型以及可塑性模型。在保持了结构与机制类脑的基础上，初步实现了认知功能上表现出类人行为，如记忆模拟系统在实现记忆的同时表现出很好的抗噪性能，类脑强化学习系统在特定任务中学习效率接近人类被试平均水平。此外，还构建了面向亿级类脑神经网络建模的计算平台，支持跨尺度的生物神经网络模拟与类脑认知计算建模^[52]（图2）。

多尺度、多脑区、多认知功能融合的认知计算模型与类人学习结合是实现类脑智能的一个重要方向，使智能系统不仅在模型结构上类脑，而且具有类人终生学习的能力，从而自主产生智能认知功能并不断自适应进化。

5 关于类脑智能模型的进一步思考

脑是自然界中最复杂的系统之一，由上千亿（ 10^{11} ）神经细胞（神经元）通过百万亿（ 10^{14} ）突触组成巨大网络，实现感知、运动、思维、智力等各种功能。大量简单个体行为产生出复杂、不断变化且难以预测的行为模式（这种宏观行为有时叫做涌现），并通过学习和进化过程产生适应，即改变自身行为以增加生存或成功的机会。类脑智能研究需要加强人工神经网络和生物神经网络在机构、功能和学习机制上的融合，尤其迫切需要围绕两个方向进行科研攻关。

（1）迫切需要发展更加高效能的新一代人工神经网络模型。目前DNN一定程度上已经借鉴了神经系统的工作原理，并具备相对完整的编解码、学习与训练方法，但从发展和应用的眼光看，该类模型还存在巨大的提升空间。而大部分SNN在学习与训练算法方面更多地借鉴了神经元、突触等微观尺度的机制，其在学习方式上更加接近于无监督学习，计算效能也比深度网络高出一个量级，但由于网络训练只考虑了两个神经元之间的

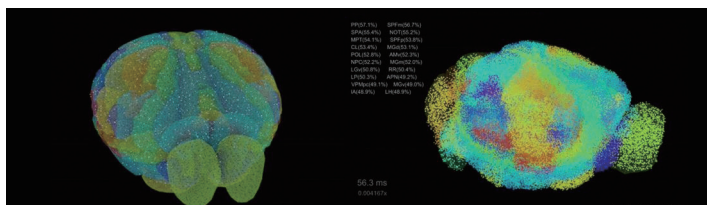


图1 鼠脑多尺度计算建模（包括213个脑区、7100万神经元、脑区之间的连接权重依据生物实验结果）

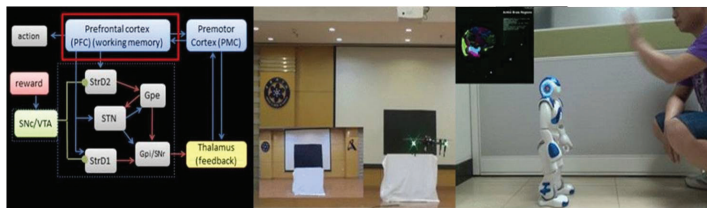


图2 自主强化学习的多脑区协同计算模型应用于无人机故障和机器人自主学习

局部可塑性机制,对介观(如神经网络连接、皮层结构)、宏观尺度(如脑区之间的网络连接)的借鉴非常缺乏,因此在性能上与DNN等模型还存在一定差距。两个模型都需要不断从脑科学中吸取营养并不断融合,发展出性能更好、效能更高的新一代人工神经网络模型。

(2) 迫切需要发展可自适应的类脑学习方法与认知结构。在类人认知行为的机器学习方面,越来越多的研究着眼于提高神经网络、认知计算模型和智能系统的自适应能力。让机器像人一样不断地从周围环境对知识、模型结构和参数进行学习和自适应进化,是机器学习的最高目标,这种学习方式被称为终生学习(Life-Long Learning)或永不停止的学习(Never-Ending Learning)^[53,54],里面混合监督学习、无监督学习、半监督学习、增量学习、迁移学习、多任务学习、交互学习等多种灵活方式。最新的基于生成模型的贝叶斯程序学习^[55]体现了人脑普遍的个例学习能力。认知科学认为,一个概念的形成具有组合性和因果性,因此认知一个新概念时用到了已有的经验积累,从而具有个例的举一反三能力。贝叶斯程序学习借鉴了这些认知科学里的概念,对字符图像进行笔画的分解和组合性的学习和推理,让程序学会如何学习,从而能举一反三地辨认新样例和产生新字符,在一个特定的视角通过了视觉图灵测试。发展可持续的类人学习机制,需要通过脑科学建立适合这类学习机制的认知结构;同时直接从大网络中通过学习演化出类脑的认知结构则更是期望的基础性突破。

6 结语

本文从类脑智能研究的视角论述了脑与神经科学对未来人工智能的启发,随后介绍了近年来在神经网络计算模型方面取得的进展与发展趋势,并简要介绍了中科院先导专项支持下的类脑智能计算模型与方法研究的设想与进展。类脑智能计算模型研究需要脑与神经科学、人工智能与计算科学、器件与系统的科研人员深度实质性协同与融合,才能够真正设计并实现出机制类脑、行

为类人的通用类脑智能计算模型。

参考文献

- 1 Ghosh—Dastidar S, Adeli H. Spiking neural networks. *International Journal of Neural Systems*, 2009, 19(04): 295-308.
- 2 Chrol—Cannon J, Jin Y C. Computational modeling of neural plasticity for self-organization of neural networks. *BioSystems*, 2014, 125: 43-54.
- 3 Dayan P, Abbott L F. Theoretical neuroscience: computational and mathematical modeling of neural systems. cambridge: The MIT Press, 2001.
- 4 Seung S. Connectome: How the brain's wiring makes us who we are. New York: Houghton Mifflin Harcourt, 2012.
- 5 Jiang X L, Shen S, Cadwell C R, et al. Principles of connectivity among morphologically defined cell types in adult neocortex. *Science*, 2015, 350(6264): 9462.
- 6 Lee W C, Bonin V, Reed M, et al. Anatomy and function of an excitatory network in the visual cortex. *Nature*, 2016, 532(7599): 370-374.
- 7 Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 1980, 36(4): 193-202.
- 8 Cadieu C F, Hong H, Yamins D L, et al. Deep neural networks rival the representation of primate it cortex for core visual object recognition. *PLoS Computational Biology*, 2014, 10(12): e1003963.
- 9 Yamins D L K, DiCarlo J J. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, 2016, 19(3): 356-365.
- 10 Liang M, Hu X L. Recurrent convolutional neural network for object recognition// *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. United States: IEEE Press, 2015: 3367-3375.
- 11 Liang M, Hu X L, Zhang B. Convolutional neural networks with

- intra-layer recurrent connections for scene labeling// Advances in Neural Information Processing Systems 28 (NIPS 2015). United States: Curran Associates, Inc., 2015: 937-945.
- 12 Wang Q, Zhang J X, Song S, et al. Attentional neural network: feature selection using cognitive feedback//Advances in Neural Information Processing Systems 27 (NIPS 2014). United States: Curran Associates, Inc., 2014: 2033-2041.
- 13 Cao C S, Liu X M, Yang Y, et al. Look and think twice: capturing top-down visual attention with feedback convolutional neural networks// Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV 2015).United States: IEEE Press, 2015: 2956-2964.
- 14 MnihV, Kavukcuoglu K, Silver D. Human-level control through deep reinforcement learning. Nature, 2015, 518(7540): 529-533.
- 15 Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search. Nature, 2016, 529(7587): 484-489.
- 16 Hodgkin A L, Huxley A F. Propagation of electrical signals along giant nerve fibers. Proceedings of the Royal Society of London. Series B, Biological Sciences, 1952, 140(899): 177-183.
- 17 Anthony N B. A review of integrate and fire neuron model. Biological Cybernetics, 2006, 95(1): 1-19.
- 18 Izhikevich E M. Simple model of spiking neurons. IEEE Transactions on Neural Networks, 2003, 14(6):1569-1572.
- 19 Tsodyks M, Pawelzik K, Markram H. Neural networks with dynamic synapses. Neural Computation, 1998, 10(4): 821-835.
- 20 Tsodyks M, Wu S. Short-term synaptic plasticity. Scholarpedia, 2013, 8(10): 3153.
- 21 Maass W. Networks of spiking neurons: the third generation of neural network models. Neural Networks, 1997, 10(9):1659-1671.
- 22 Anderson J R. How can the human mind occur in the physical universe? Oxford, UK: Oxford University Press, 2007.
- 23 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521(7553):436-444.
- 24 Collobert R, Weston J, Bottou L, et al. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 2011, 12:2493-2537.
- 25 Zeng D J, Liu K, Lai S W, et al. Relation classification via convolutional deep neural network//Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014). United States: Association for Computational Linguistics, 2014: 2335-2344.
- 26 Socher R, Pennington J, Huang E H, et al. Semi-supervised recursive autoencoders for predicting sentiment distributions// Proceedings of the Conference on Empirical Methods in Natural Language Processing. United States: Association for Computational Linguistics, 2011: 151-161.
- 27 Xu J M, Wang P, Tian G H, et al. Convolutional neural networks for text hashing// Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015). United States: AAAI Press, 2015: 1369-1375.
- 28 Reasoning, Attention, Memory (RAM) NIPS Workshop 2015. [2016-05-10]. <http://www.thespermwhale.com/jaseweston/ram/>.
- 29 Chaudhuri R, Fiete I. Computational principles of memory. Nature neuroscience, 2016, 19(3): 394-403.
- 30 Cho K, Van Merriënboer B, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint, 2014: arXiv:1406.1078.
- 31 Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation, 1997, 9(8): 1735-1780.
- 32 Vinyals O, Le Q. A neural conversational model. arXiv preprint, 2015: arXiv:1506.05869.
- 33 BahdanauD, ChoK, Bengio Y. Neural machine translation by jointly learning to align and translate. arXiv preprint, 2014: arXiv:1409.0473.
- 34 Shang L F, Lu Z D, Li H. Neural responding machine for short-text conversation. arXiv preprint, 2015: arXiv:1503.02364.

- 35 Sukhbaatar S, Weston J, Fergus R, et al. End-to-end memory networks// *Advances in Neural Information Processing Systems* 28 (NIPS 2015). United States: Curran Associates, Inc., 2015: 2431-2439.
- 36 Graves A, Wayne G, Danihelka I. Neural Turing machines. *arXiv preprint*, 2014: arXiv: 1410.5401.
- 37 Peng B L, Lu Z D, Li H, et al. Towards neural network-based reasoning. *arXiv preprint*, 2015, arXiv: 1508.05508.
- 38 Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254-1259.
- 39 Itti L, Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 2000, 40(10-12): 1489-1506.
- 40 Yarbus A L, Riggs L. Eye movements during perception of complex objects, *Eye Movements and Vision*, Springer, 1967, 171-211.
- 41 Navalpakkam V, Itti L. Modeling the influence of task on attention. *Vision Research*, 2005, 45(2): 205-231.
- 42 Baluch F, Itti L. Mechanisms of top-down attention. *Trends in Neurosciences*, 2011, 34: 210-224.
- 43 Borji A, Sihite D N, Itti L. What/where to look next? Modeling top-down visual attention in complex interactive environments. *IEEE Transactions on Systems, Man, and Cybernetics, Part A - Systems and Humans*, 2014, 44(5): 523-538.
- 44 Goel V, Dolana R J. Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition*, 2004, 93(3): B109-B121.
- 45 Jia X Q, Liang P P, Lu J M, et al. Common and dissociable neural correlates associated with component processes of inductive reasoning. *NeuroImage*, 2011, 56(4), 2292-2299.
- 46 Gazzaniga M S, Ivry R B, Mangun G R. *Cognitive neuroscience: the biology of the mind* (4th edition). New York: W. W. Norton & Company, 2013.
- 47 Eliasmith C, Stewart T C, Choo X, et al. A large-scale model of the functioning brain. *Science*, 2012, 338(6111): 1202-1205.
- 48 George D, Hawkins J. Towards a mathematical theory of cortical micro-circuits. *PLoS Computational Biology*, 2009, 5(10): e1000532.
- 49 Maass W. Networks of spiking neurons: The third generation of neural network models. *Neural Networks*, 1997, 10(9): 1659-1671.
- 50 Zenke F, Agnes E J, Gerstner W. Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nature Communications*, 2015, 6: 6922.
- 51 Zhang T L, Zeng Y, Zhao D C, et al. HMSNN: hippocampus-inspired memory spiking neural network. *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics*, IEEE Press, 2016.
- 52 Liu X, Zeng Y, Xu B. Parallel brain simulator: a multi-scale and parallel brain-inspired neural network modeling and simulation platform. *Cognitive Computation*, Springer, 2016.
- 53 Kirstein S, Wersing H, Gross H M, et al. A life-long learning vector quantization approach for interactive learning of multiple categories, *Neural Networks*, 2012, 28: 90-105.
- 54 Mitchell T, Cohen W, Hruschka E, et al. Never-ending learning// *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI 2015)*. United States: AAAI Press, 2015: 2302-2310.
- 55 Lake B M, Salakhutdinov R, Tenenbaum J B. Human-level concept learning through probabilistic program induction. *Science*, 2015, 350(6266): 1332-1338.

Research Status and Developments of Brain-inspired Intelligence

Xu Bo^{1,2} Liu Chenglin^{1,2} Zeng Yi^{1,2}

(1 Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;

2 Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China)

Abstract Recent advances in Artificial Intelligence (AI) have manifested an important trend, namely, inspirations from brain science can significantly improve the level of intelligence for AI computational models. With only local and partial inspirations from the brain, great advancements have been achieved. Nevertheless, deeper investigations and inspirations from the brain are needed to realize and exceed human-level intelligence. The ultimate goal of brain-inspired intelligence is to bring inspirations from brain structures and information processing mechanisms to brain-inspired cognitive computational models, so as to realize next-generation artificial intelligence models and systems with general intelligence. In this article, we review recent advances and discuss trends of brain-inspired computational models, including new models of artificial neural networks, and cognitive computation models. We also briefly introduce the research of brain-inspired cognitive computation models and methods supported by the strategic priority research project of Chinese Academy of Sciences.

Keywords brain-inspired intelligence, artificial neural networks, memory, attention and reasoning, spiking neural networks, multiple brain region coordination, autonomous learning

徐 波 中科院自动化所所长、研究员，中科院脑科学与智能技术卓越创新中心副主任，中国中文信息学会副理事长。曾任国家“863”计划信息技术领域专家组专家。长期从事人工智能研究，主要研究领域包括：类脑智能、类脑认知计算模型、自然语言处理与理解、类脑机器人。E-mail: xubo@ia.ac.cn

Xu Bo Professor, president of Institute of Automation, Chinese Academy of Sciences, and deputy director of the Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences. He also serves as an associate president of Chinese Information Processing Society of China. He was a steering committee member of National high-tech Programme (“863” Program). His main research interests include brain-inspired intelligence, brain-inspired cognitive models, natural language processing and understanding, brain-inspired robotics, etc. E-mail: xubo@ia.ac.cn