一类智能计算思想:按自然法则计算

董占球

刘岩 韩承德

(中国科学院研究生院 北京 100039) (中国科学院计算技术研究所)

摘要 本文从介绍计算的演化入手,引出了当前人工智能研究中存在的本质困难,指出了这些困难与近代数学的关系,也指出了 AI 的出路和所应采取的策略。最后,结合具体算法,介绍了按自然法则计算的某些主要特征。

关键词 智能计算,按自然法则计算,近代数学。

机器计算大致有几个方面,由最初的数值计算,到符号计算,到后来的神经计算。数值计算的局限性是明显的,它与人类智能显然有着较大差距,而符号计算的缺陷却逐渐暴露出来。首先,符号计算必须依靠先验知识来学习和建立规则,是构造性的;其次,符号计算不便于交互,因为自然界的知识本质上不是以符号形式存在的;符号系统与自然界的交互过程要经过“编译”处理;最后,符号计算本质上是集中式的、串行的,而人类智能表现出相当分布的、并行的特点。

在符号计算因这些缺陷走入低谷时,神经计算应运而生。神经计算声称不需要先验知识(事实上,它也需要先验知识,只是所需知识是特征级的),就能完成知识的记忆、学习和泛化(Generalization)。然而,随着研究的深入,神经计算的缺陷又渐渐表现了出来。首先,它存在着很多局部极值点,极值点未必表示所需的知识,真正的知识也未必都收敛于极值点;其次,结点数对知识的学习存在着饱和态(例如 Hopfield-tainer 模型,通常 100 个点的 ANN 大至可存 14 条知识);最后,人工神经网络(ANN)不易分析(重要的原因是其中存在着时空意义下的高阶非线性和随机性)。

后来,人们又试图在计算中采用多种层次,进而将数值计算、符号计算、神经计算结合起来,从不同等价类(即不同角度)组合处理问题。这方面较成功的工作有 INTSE 系统等^[1]。但所有这些努力本质上都是模拟的,非本征的,它们的工程价值远胜过其理论意义。为了从根本上为问题的解决探索一条道路,我们研究了按自然法则计算的思想。

它的含义是以具有(而非模拟)人类智能的某些根本属性(非线性,不确定性,不完备性,满意性)的

计算模型为载体,能够解决非递归问题的计算。

在“Physical Computation”一文中,G. C. Fox 给出了 Physical Computation 的概念,并给出了它的几大分支。我们研究了 Fox 关于 Physical Computation 的概念,并称之为“按自然法则计算”^[2]。

1 当前 AI 的困难

当前人工智能的主流仍然是符号处理,这在本质上是属于非数值计算的一类。经典的非数值计算建立在布尔运算基础上,因此符号处理也主要建立在布尔运算基础上,而且在程序设计与执行时要涉及某些递归结构(DTM 可解的问题都是局部递归问题)。以上过程有以下局限性:

1) 布尔量的取值为真、假,在布尔代数中表示为 0、1。因此,符号主要是在 $[0,1]$ 域上生成的。这就带来表示方法的某些局限性:·离散性(无法精确刻画存在的连续现象);·判定性(无法表达客观的模糊语义);·互斥性(无法采取恰当的折中行为)。

2) 在问题求解前,必须给出确定算法。在有关问题的启发信息很少时,很难有效地解决问题,进而产生了很多难解问题(如 NPC 类问题和 NPH 问题)。

3) 处理的结果是唯一的,即使问题本身有不稳定性,最后也要用确定的形式表现出来。这就无法表示人类智能所固有的测不准性和不完备性。

4) 图灵意义下的可计算问题都是递归的,递归问题都是顺序的,在此意义下不存在全局性的并行算法。

由此可见,建立在传统 DTM 基础上的 AI 研究,实质上是纯粹模拟人类智能外特性的研究工作,无法真正实现人类智能的内涵。事实上,就连 Hopfield 本人都感觉到,ANN 要有真正意义的突破,得

依赖于人类对生物神经系统研究的突破。

2 AI困境与近代数学^[4]

近代数学是完善的、严谨的、优美的形式系统，它处理问题的方法可归纳如下：1)建立数学模型；2)提出定理并给予证明；3)利用定理对模型施行演绎，并求出结果。近代数学有以下特点：

1)结构不变性。近代数学一旦为某一现实问题抽象出一个数学模型，虽然在求解这一模型时，可以在不同的变量条件下，得出不同的结果，但模型本身是不变的，这一不变性反映出该模型所刻画现实系统的不变性。然而，人的思维过程是动态重构过程，AI研究涉及大量的可变性。

2)封闭性。近代数学讨论中的任何可讨论问题必在一个论域内存在。而AI研究表现为不断更新的开放性，这在形象思维和顿悟思维中表现得尤为突出。

3)逐步求精性。近代数学的一切目的在于追求“精确解”，即使对问题本身而言是近似解。而AI研究的大量问题有时只需一个满意解或满意解集。

4)线性。近代数学长于处理变量的叠加和迭代，而对非线性系统则少有作为。不幸的是，AI研究中涉及到大量的非线性，ANN即为一例。

5)收敛性。任何发散问题在近代数学中都无法讨论。而人类思维活动并不仅仅是收敛的，AI不应仅被局限在收敛意义下。

3 AI的问题

事实上，尽管计算和计算机科学(此处不包括计算机技术)历经了数十年的沧桑变换，人工智能也曾几度沉浮，它们所要回答的基本问题仍然是初始的图灵命题，即：机器能思维吗？^[5]任何人工智能(乃至计算及计算机科学)的实质性进展都取决于对此问题的回答程度，亦即取决于任何人工智能系统向自然系统的逼近程度。因此，我们认为，AI应尽量避免确定的、精确的、事先设定的诸如此类的概念和方法。这主要是由于人类智能本身存在着不确定性、不完备性、非线性、满意性等特征。AI应尽量避免(或少用)理性主义和简化论的思想和方法，因为人类智能本身存在着感性的领悟和对复杂内涵的理解能力。同时，我们应尽可能地在AI系统中包含高级的语义，在AI中允许或实现不确定性，其中包括：①问题前提的不确定性(包括模糊条件和非线性结构)；②变量的不确定性(包括可变参数和非线性关系)；

③求解过程的不确定(包括随机优化和非线性重组)；④问题结果的不确定性。(要求问题的最后解是满意解或解集，其生成是随着以上诸不确定因素的变化而变化的。)

由于非线性的广泛存在，AI系统的智能是和它所存在的环境以及环境对系统的影响密不可分的，同一AI系统的外部环境不同，系统应表现的行为也不同。

因此，我们寻求建立具有本征特点的智能模型，以从根本上改变传统人工智能的非本征的外特性模拟方法。

4 AI方法论

上述种种，可概括为：

一个AI系统对其环境给它的影响和刺激在自适应的优化处理以后进行反应。

在此应强调AI系统不应只是刺激-反应模型，必须充分强调系统自身的自主能力(亦即自适应优化的处理能力)。为此，我们必须找到一组给出条件(刺激/响应)，系统自主进行寻优。根据某种判据给出可取结果(反应)。事实上，问题至此可归结为复杂系统(足够多的变量，足够强的变量之间的相互作用，足够多的外部环境控制及其与系统间的相互作用)的调控优化问题。

总之，AI研究不应只停留在外特性的模拟层次，应该更强调实现新的载体机制，当前迫切需要基本计算模型的本质变革，并要求新模型能够本征地表现出非线性、不确定性、不完备性和满意性等。

5 按自然法则计算

按自然法则计算之类的计算思想是比较符合上述问题求解策略的。使用此类算法需要三个条件^[6]：(1)按自然法则计算所处理的问题必须是优化问题。(2)按自然法则计算在优化过程中必须是有可供调控的参数。(3)必须找出优化目标函数，以及优化过程与受调控参数之间的内在联系，并设计出优化判据。

此类算法(例如SA,GA等)还表现出与传统方法截然不同的特点，这集中体现在解决问题时的三个不需要上^[2]：(1)不需要对问题本身抽象出一个精确数学意义上的解析式的数学模型。(2)在求解某一具体问题前，不需要给出求解该问题的确定的算法步骤。(3)结果无需精确解，而是追求满意解(或满意解集)。

按自然法则计算的分支都在不同程度上表现出与人类智能相似的本征特性。(有关BI的本征特性,参见文[2]。)以下分别讨论这种相似性:

非线性:此类算法具有非线性特征。例如,在有序组合的遗传算法中(IGS-O: Integrated Genetic Search-Ordered Combination),交叉算子(Crossover)赋予了算法重要的非线性特征。有人在改进遗传算法时,引入了一致交叉重组(Uniform Combination)算子,该算子与突变算子(mutation)相配合,从种群中随机地选择两个点 X_i, X_j ,并用如下方式产生第三点 X_k ;对于 X_k 的每一位 $X_k(1), X_k(2), \dots, X_k(n)$,以概率 P 随机地置其值为0,1;如未能赋值,随机地从 X_i, X_j 中选一个并将相应位的值赋予该位。在某些情况下,“一致交叉重组”的确有助于加快搜索速度和优化种群,但经分析,此算子只在处理线性问题时具有优势。面对自然界中广泛存在的非线性问题,此算子表现出极坏的性能。相反,有序交叉重组对于非线性问题十分有效。事实上,有序交叉算子来源于基因科学中的染色体重组,它能很有效地描述遗传学的以下现象:1)个体A的染色体上具有 x 基因,进而导致A具有一个非常好的性状;2)个体B的染色体上具有 y, z 基因,使B具有另一个非常好的性状;3)然而,有可能对于一个同时具有 x, y, z 基因的个体C,它既不具有A的相应好性状,也不具有B的相应好性状,而是具有了一个非常差的性状。这种现象就是自然界中所普遍存在的“非线性基因重组”。在解决此类问题时,一致交叉由于很容易打破由 x, y, z 组成的构成块(Building Block),又称关键模式(critical schema),而用传统的有序交叉,则可以得到好得多的效果^[5,6]。

不确定性:此类算法还具有不确定性特征。例如,模拟退火(SA, Simulated Annealing)算法中,搜索序列在产生新的当前解时,以概率 $e^{-\Delta E/kT}$ 接受较差解^[7],同样在遗传算法(GA)中,无论是一致交叉,还是有序交叉,都存在以某概率 P 给某特征位赋值的问题^[8]。

不完备性:虽然可以证明GA, SA在合理的调控参数控制下,理论上以概率1收敛于全局最优解,但此方法本身相对于回溯法而言是一个不完备算法,其搜索时间可能大于相应的回溯算法。事实上,这正与人类智能的不完备性相对应。

满意度:此类算法多以现实满意度为终止判据。

例如,对于SA的改进形式MTRSA^[9],终止判据可以如下给定:

• 记在某一 T_i 下得到的 $NS(i)$ 为 $NS(T_i)$;

• 若自某一 i 起,有:

$$NS(T_i) = NS(T_{i+1}) = \dots = NS(T_{i+p});$$

表明温度连续下降 p 次后解的最优性未得到改善,此时可认为继续降温已无意义。

• 在算法中设一阈值 P_s ,使当 $p \geq P_s$ 时退火过程停止。

又如,对于遗传算法,存在关于 P_{min} 的设定问题,如果 $P \leq P_{min}$,则算法停止。显然,上述两个终止判据都隐式遵循一个基本原则:以现实满意度为标准。

6 结论

综上所述,当前的AI研究存在着很多困难。在图灵框架内,这些困难是本征的,不可逾越的。对现存各模型在当前框架内的改进,从某种程度上来说,只具有工程意义,突破图灵模型才是解决问题的根本办法。按自然法则计算恰恰具有人类智能的某些属性,它的研究有望大大缩短传统模型框架与智能的天然载体—智能模型的距离。虽然,现有的科学,技术水平和按自然法则计算的研究现状还不足以导致真正的智能模型在短期内出现,但这必定是一条极具前景的道路。

参 考 文 献

- [1]史忠植,神经计算,电子工业出版社,1993年
- [2]刘岩、董占球、韩承德,一类非递归结式的计算思想,计算机科学,本期
- [3]Turing, A., On Computable Numbers, with an application to the Entscheidungs problem, Proc. London Math. Soc. Ser., 1936
- [4]董占球,人工智能(AI)研究中的若干问题,计算机研究与发展,1991年第6期
- [5]David H. Ackley, Stochastic iterated genetic hill-climbing, 1987 Carnegie Mellon CMUCS-87-107
- [6]Y. Liu, Research on the Theory and Application of Physical Computation, Master Degree Thesis, Harbin Institute of Technology, 1993
- [7]S. Kirkpatrick et al., Optimization by Simulated Annealing, 1983 Vol 220, Science