

# 论人工智能的整体观

On Artificial Intelligence as a Whole

聂勇军 刘凤岐 陈火旺 高宏奎

(国防科技大学计算机学院,长沙 410073)

70-73

TP18

**摘要** 人工智能研究中分析方法的应用造成了许多困难;整体观原理是构成事物复杂性的一个重要原理,人类智能和人工智能就是整体特征的表现;通过分析整体观的人工智能,指出克服困难的途径和进一步的研究问题。

**关键词** 智能,人工智能,分析,方法,整体观

推理系统 专家系统

## 一、引言——人工智能的困难

从人工智能(AI)的研究方法论上来看,人们很自然地拿自然智能即人类智能作为类比和研究的模型。人们通过分析自然智能的各个方面,从表现形式上把智能分为如下几种能力:语音识别、语言理解、图形图象识别与理解、博弈、规划、推理、综合、类比、记忆和学习等等。于是人们对AI的基本要求是能有相应的系统来模拟人的相应能力。然而除了很少几方面取得成功外,AI离这个目标还很遥远,因此AI长期以来受到各方面的责难,是什么造成AI的困难呢?

长期以来,“分析”一词在最广泛的范围内被使用,人们习惯上,毫不怀疑地把一个问题拿来进行分解,然后将分解的子问题再分解,如此下去,直到子问题还原为可以被把握的对象为止。西方人精于分析的思维方式,分析方法的思想在西方占有统治地位,故极大地推动了近代科学技术在西方的发展,例

如对基本粒子研究中分析方法发挥了重要的作用。

在认识论上,分析方法的思想主要表现为,往往将系统的性质或能力简单地归结为部分性质或能力之和,即还原论。不可否认,这种方法是一个重要的思维方法,对一个未知对象的研究,分析的重要性是有目共睹的,然而它不是万能的。正是它的过度使用,使人们在研究一类具有整体特性的问题时,造成了认识上的混乱和困难,特别是分析方法在AI研究中的应用,使AI遇到了很大的困难。

分析方法应用到AI系统中产生了如下的后果:通过对自然智能的分解,把人类智能分化为若干方面,并相应地在一些比较容易的方面作出诸如:语音识别、国际象棋、图像识别、推理系统和专家系统(ES)等AI系统。这些AI系统在许多狭窄领域的应用方面取得了一定的效果并获得很好的经济效益,但往往与人类的自然智能相比,却显得太肤浅太幼稚。例如目前最好的人类面部识别AI软件的能力还远远不及一个两岁的儿童。我们认为这是分析方法

MA冲突的预测——决策方法,并将其应用于CSCW的资源冲突的解决。

**小结** 众多学科间有着共同之处,而学科的交融又是现代科学发展的一个特征和趋势,对共性的把握无疑有益于学科本身的发展和人们对它的认识。人工智能和面向对象是计算机领域的两大热点,它们都有着思维科学基础,面向对象又和分形有关,分形和混沌是密切联系的。从而,这些学科直接或间接相关,这对于它们互补或借鉴是十分有益的,同时,我们也可以从中概括共性,总结规律,指导进一步的工作。

### 参考文献

[1] 傅京孙等,人工智能及其应用,清华大学出版社,1987  
 [2] 董军、马小虎,面向对象方法:产生、特征和应用,计算机与信息处理标准化,1996,6.

[3] 钱学森,关于思维科学,上海人民出版社,1986  
 [4] 刘泉宝、刘永清,关于人工智能的哲学思考,计算机科学,22(2)1995  
 [5] Harmon Paul, Object-Oriented AI, A Commercial Perspective. CACM, 38(11)1995  
 [6] 董军、肖少拥,面向对象的知识表达及其应用,微型电脑应用,1996,3  
 [7] 肖前等,辩证唯物主义原理,人民出版社,1981  
 [8] 爱因斯坦,爱因斯坦文集,商务印书馆,1976  
 [9] Ott Edward, Chaos in Dynamical Systems, Cambridge University Press, 1993  
 [10] 王东生、曹磊,混沌、分形及其应用,中国科学技术大学出版社,1995  
 [11] J. S. Nicolis, Chaos and Information Processing—a Heuristic Outline, Singapore, Worldscientific, 1991  
 [12] 胡上序等,人工神经元计算导论,科学出版社,1994

应用于 AI 研究中的偏差所造成的,从而导致 AI 系统缺乏智能。

由于分析方法在西方占据着统治地位,源于西方的现代 AI 研究受到了这种思维方法的极大影响,但是这种方法不适合于 AI,这是目前 AI 遇到的一个主要问题。早在五十年代,在编制了第一批跳棋、象棋、智能游戏和自动推理程序之后,一些控制论的热衷者就把这程序看作是未来的机器人和“会思维机器”的功能模拟原型。1958年 H. A. Simon 和 A. Newell 对 AI 的一些乐观十年估计,除了计算机下棋的目标比较接近外,其它的一些预测,直到今天还是预言。

## 二、整体大于部分之和——构成复杂世界的原则

整体观是东方人对认识论的一个重要贡献,同西方人的分析观完全相反,在认识事物时是从高层次、综合和整体性方面来考虑问题。中国人是这种观点的一个主要代表。

学习过中学化学的人都知道,水由两种元素组成,其分子式为  $H_2O$ 。显然水分子组成不能说不简单。我们又知道,水在常温下是液体、无色、无味、透明,常压下加热到  $100^{\circ}C$  会变成气体,温度降低到  $0^{\circ}C$  以下时会变成固体冰。相信水的这些性质对任何一个人都是不陌生。问题是水为什么会有这样的性质呢?它们是从哪里表现出来的呢?我们不能从 H 和 O 原子的性质中得出来,作为原子,H 和 O 无所谓水的性质的。能否从  $H_2O$  分子中得出呢?也不能, $H_2O$  当然也无所谓液体、透明等性质的,那么我们应从哪里去寻找水的性质呢?那就是大量  $H_2O$  的整体性,它是由  $H_2O$  分子的大量迭加而产生集团的性质,只分析一个或若干个  $H_2O$  是毫无意义的,是不可能得出水的性质。

多少年来,生命系统的起源及演化一直是科学研究的一个重大课题。一般说来,生命系统有两个与众不同的特征:复杂和组织。即使是只观察和分析一个低级细菌,它身上有复杂的功能和形态的网络。这细菌可以用多种多样的方式与外部环境相互作用,以一种有控制的方式交换物质,其体内的情形是一个井井有条的大城市。负责大部分控制功能的是细胞核,其中包含着遗传“密码”,就是使细菌得以进行繁殖的全套化学指令。控制并左右这细菌的一切行为的化学结构是由多达上百万个原子组成的分子,这上百万的原子是以复杂、高度特殊的方式组织起来的,就生命的化学基础而言,最重要的是核酸分子,即大名鼎鼎的“双螺旋”形状的 RNA 和 DNA(核糖核酸和脱氧核糖核酸)。

一个生物学上的生物体是由完全平常的原子组成的。看到这一点是重要的,不错,生物体的部分代谢功能是从环境中获得新的物质,并非泄变了质的或不用的物质,活细胞里的一个碳、氧、氢和硫原子与细胞外的原子没什么两样。一个生命系统之所以是活力的生命现象,是因为其结构的整体性使然。决不会有什么“活力”的原子使其成为生命。

看过漫画的人往往能对一些优秀作品的深刻思想产生共鸣,一幅由点和线组成的小小漫画为什么会传达如此深刻的思想呢?这是由这些点和线及布局结构的整体性而表达的,也就是说它们表达的是一个整体性层次的信息。如果单纯从低层次的点和线上看,不过是由一些点和线组成的集合而已;仅从局部上看也是一些无法理解的片断;只有从整体上才能体会漫画的内涵思想。

对于什么是智能(或自然智能)问题,人们历来都有不同的看法。有的人认为智能是人类特有的一种能力;有的人认为智能并非只是人类特有的能力,而有一种度量高低的谱系,以人类智能为最高,其次是人类的近亲狒狒和猿类,再其次是鲸类,如此等等。我们不妨把智能看作是在通常情况下,能对一定的环境状态作出有限合理性(Bounded Rationality, H. A. Simon 语)的反应能力。人类尽管有自身肢体能力和知识的限制,但却能自如地解决那些看来难以处理的棘手问题,并作出合理决策和反应,重要的一个原因是他能通过各种信道获得各种信息(如眼、鼻、耳、口和肤信道),综合利用各种信息,采用逐步尝试(试错)的方法,到达合理的目标,得到有意义的解。由于人类有这些广泛的信息,大量地使用启发式(Heuristic)的方法,故在计算机中(特别是 AI 问题)的 NP 难解的问题在人类看来,根本不是问题。从另一个角度来看,人的知识虽然有限,但人类对问题求解方法却是多种多样的,在进行问题求解时,最重要的是将这些方法综合在一起,组成一个复杂的推理过程。这就是说虽然智能的表现形式是多种多样(语音识别、语言理解、图形图象识别与理解、博弈、规划、推理、综合、类比、记忆和学习等等),但这些能力的体现不是孤立的,而是相互依存,相互制约。最后再让我们看一看人类广泛的常识应用。我们知道,人类拥有大量的常识知识,人们会在求解问题中不知不觉地应用常识知识。这些广泛的常识知识的应用,往往能极大地约束人类的思考范围,使推理成为中肯推理(Relevance Inference)。因此常识的综合应用,是人类获得一般智能的重要保证。这些各方面的能力综合在人类大脑中体现了智能的整体性。

由分析方法导致还原论的结果,在许多方面是

成功的,特别是应用于一些简单问题和系统时,其作用不容否定,然而在涉及“整体性”方面的一大类问题时它就显得力不从心了。可以说“智能”中应用分析的方法是它的一个失败之处,从而导致 AI 研究困难重重。从生物学的角度来看,我们知道,构成人类大脑智慧物质基础的核心是神经细胞,而神经细胞不过是由我们所熟知的原子所组成,如碳、氢和氧等等。然而我们却始终不明白智能的奥妙,原因何在?这就是因为“智能”是一种整体特性的表现形式,是大量组成成份神经细胞综合能力的表现。

### 三、整体性的 AI——实现 AI 智能步步提高

如果一切事物都由分析并导致还原论来决定,说任何一个复杂的事物不过是由某某部分组成而已,那么这个世界就不会是我们今天所面临的如此纷繁的世界,人类也就不可能发生进化。前面说过,现代 AI 产生于西方,在这个产生过程中,极大地受到西方分析思想的影响,然而随着现代科学技术的发展突飞猛进,这种分析思想的局限性日益显现出来。现代科学技术中整体论、系统论、超协调、耗散结构等思想方法越来越起到重要作用。下面我们就具体分析一下 AI 研究遇到的困难在整体性方面的具体表现。

AI 有两个相关的目标:(1)探索智能的本质,(2)赋予计算机(或任何机器)以智能。自从现代 AI 提出以来,由于认识上出发点的差异,人们对什么是 AI 的认识一直存在分歧,对它的定义也就各执一词。这说明探索 AI 道路的复杂性和艰巨性。

ES 是最早走出实验室取得实际应用的 AI 系统。如果我们向一个诊断皮肤病的 ES 提出一个问题:张先生的汽车表面上有一些浅红褐色的斑点,它有什么问题?系统会告诉你,你的汽车可能生了红癣,并建议它每天口服红霉素,口服四次,连服两周。显然这是很可笑的事。原因是系统只通过一个信道获得症状,没有对被诊断对象的识别能力,不能应用有关的常识知识进行综合的推理,从而导致不能判定对象的属性是否满足有关约束。这是 ES 缺乏获取信息的通道,常识知识的匮乏和运用知识能力有限所造成整体性能力不够。

基于上述的一些分析,我们认为这方面进一步的工作应集中在如下三个方面:

(1)多信道获得信息。通过多信道获得信息或知识,将这些信息融在一个大型 AI 系统中,使它们的信息得以综合发挥作用,从而提高系统的求解能力。

(2)系统集成。将已有的 AI 系统集成在一个系统中,使它们相互联系,相互制约,从而发挥整体方面的能力。前面说过,许多 AI 系统是人类能力表现

的某一方面的模拟,将许多系统集成在一起,可获得整体的效应。

(3)常识的应用是 AI 整体性的体现。为什么这样说呢?我们知道,常识的数量是巨大的,综合应用常识知识可以极大地限制问题的范围,从而可以使 AI 系统的应用范围得以拓展,可以使系统的能力具有更广的适应性。CYC 工程是整体思想的一种体现,由于常识数量巨大,仅此项工程目前已进行了十二年的努力,已取得一定的成功,当然它还处在一种非常初级的阶段。

(4)对求解方法的获取。仅有知识的 AI 系统不是有能力的系统,重要的是还要有运用知识的求解方法。特别是一些元方法(如试错方法)的开发,能导致 AI 系统对新问题探索能力的提高。

只要把这些方面在一个系统中全面体现,AI 系统的能力会不断提高的。

### 四、结论——曲折的道路光明的前途

我们认为 AI 缺乏一般能力不是不可克服的,而且这种能力并不是由任何神秘的力量来决定,而是由系统的整体性决定。整体性在现代科学技术中得到了广泛的应用,特别是在一些整体特性和大型系统的分析中普遍采用。智能就本质而言是一种整体性的体现,因此,AI 系统智能水平进一步提升,最终来源于 AI 系统研究整体性方面的工作进展。要体现 AI 是一种整体特征的能力,就必须扩展系统获取信息的通道,把现有系统的各个已有子系统能力集成到一个整体系统,并使系统获得足够多的常识知识和求解策略,这样通过我们不懈的努力,相信 AI 系统的智能水平会不断地提高。

我们对 AI 整体论思想方法的认识还很肤浅,在研究中应用这种认识论思想还是一个非常初步阶段。因此在如今这个 AI 相对低潮期,将各种主义和流派观点集中综合考虑,相互融合,取长补短,并不断吸收新的研究成果,才能一步一步地提高整体性的认识,向 AI 的目标前进。

我们坚信,只要在 AI 的研究中坚韧不拔,一方面努力挖掘智能的意义,一方面继续努力作整体性方面的研究,她的前途一定是光明的。

#### 参考文献

- [1](美)N. 维纳,陈步译,人有人的用处,控制论与社会,商务印书馆,北京,1989
- [2](英)戴维森,徐培译,上帝与新物理学,湖南科学与技术出版社,湖南,1992
- [3]D. B. Lenat & R. V. Guha, Building Large Knowledge-Based Systems, Representation and Inference in the Cyc Project, Addison-Wesley Publishing Co. Reading, MA 1990

73-77

知识发现

数据库

人工智能 (19)

计算机科学1997 Vol. 24 No. 5

# 知识发现的若干问题及应用研究<sup>\*</sup>

王清毅 陈恩红 蔡庆生

(中国科技大学计算机系 合肥230027)

TP18

**摘要** This paper first discusses several issues on knowledge discovery, such as the knowledge discovery process, the methods and algorithms of data mining and discovering association rules. Then some practical development tools and systems of knowledge discovery are given. The paper concludes with the future research work of knowledge discovery.

**关键词** Knowledge discovery, KDD, Data mining, Association rule

知识发现是一个众多学科诸如人工智能、机器学习、模式识别、统计学、数据库和知识库、数据可视化等相互交叉、融合所形成的一个新兴的且具有广阔应用前景的领域。目前国际上对知识发现的研究与开发进展很快。第一和第二届数据库中知识发现的国际会议(KDD95)和(KDD96)已经召开,第三届也于97年8月在美国的加州举行。第一本关于知识发现的国际学术杂志《Data Mining and Knowledge Discovery》于97年3月创刊。众多领域的知识发现系统和工具也不断投入市场。

知识发现的研究始于从数据库中发现有用的模式这一概念并先后有着不同的术语,如数据挖掘,知识提取,信息发现,数据模式处理,数据考古学以及数据库中的知识发现。其中数据挖掘术语多为统计学家、数据分析学家及管理信息系统团体采用;而数据库中的知识发现术语则是在于89年的第一届KDD专题讨论会上被首次采用。这一术语强调了知识是数据发现的最终产品并很快在人工智能和机器学习领域得到广泛应用。

长期以来,在知识发现领域这两个术语的范畴和使用界限一直不很清晰。直到KDD96国际会议上,知识发现研究领域的知名学者Fayyad, Piatetsky-Shapiro 和 Smyth 就这两个术语的关系作了如下阐述:KDD是指从数据库中发现知识的全部过程,而Data Mining则是此全部过程中的一个特定步骤。

## 一、知识发现的一般过程

在KDD96国际会议上, Fayyad, Piatetsky-Shapiro 和 Smyth 对KDD下了最新定义:

定义1 KDD是识别出存在于数据库中有效的、新颖的、具有潜在效用的乃至最终可理解的模式的过程。

一般说来,知识发现的过程可描述如下:

(1)熟悉应用领域的的数据、背景知识,明确所要完成的KDD发现任务性质,例如是要验证系统的假设还是要发现新的数据模式。

(2)确定与发现任务相关的数据集合。

(3)从与发现任务相关的数据集合中除去明显错误的数据和冗余的数据。

(4)数据预处理,即创建知识发现算法所需要的数据组织形式,如概念层次。如果发现任务属于验证,则还需提出相应的假设。

(5)针对所要发现任务的所属类别,例如归类、回归分析、简约、聚类、发现关联规则等等,设计或选择有效的数据挖掘算法并加以实现。

(6)测试和评价所发现的知识。例如对知识一致性、新颖性和效用性的处理。

(7)解释和运用指的是对所发现的知识进行解释以及在实际系统中的应用。

以上各个过程都涉及数据可视化技术。整个发现过程也不是简单的线性流程,步骤之间包含了循环和反复。这样可以对所发现的知识不断求精、深化

<sup>\*</sup>得到国家自然科学基金和国家教委博士点基金的支持。 王清毅 博士研究生,研究方向为机器学习、知识发现。陈恩红 博士,研究方向为机器学习,知识发现,约束满足问题。蔡庆生 博士生导师,研究方向为人工智能、机器学习、知识发现。