

人工智能

知识推理

49-52

关于知识的推理新进展

TP18

黄智生

(荷兰阿姆斯特丹大学计算机科学系留学生)

摘 要 本文由上篇和下篇两个部分组成。上篇主要阐述人工智能理论研究的意义和作用,以消除人们通常对理论性研究所产生的一些误解;同时也进一步对关于知识的推理研究的意义进行分析。下篇主要介绍最近召开的关于知识推理理论问题的国际会议(TARK)情况,着重其在计算机科学和人工智能领域方面的国际新动态和新进展。

上篇

我曾经不止一次地听到这样的质问:“人工智能理论研究已经进行了几十年,召开了许多学术会议,发表了不少学术论文,然而,真正能够直接创造巨大经济效益的商品化系统并不多见。请问这种纯粹的理论研究究竟给人类社会提供了哪些有用的东西?”

诚然,作为一个人工智能的理论研究者,我也一直在思考:这种理论研究真的是理性行为吗?我想我已得出了一些比较确定的想法和答案。下面将就这个问题谈一谈。当然,要对这个问题给出一个令人信服的回答,需要许多认真细致的论证,否则不能避免论证过程中一些过大的跳跃,使人觉得理由不充足而不能接受。同时还得考虑到读者知识背景和职业特点的差异性,就一些容易产生误解的术语作不同层次和角度的描述。这将远远超过本文的篇幅所能允许的限度。然而,为了使本文的下篇和其它类似的研究产生本应该产生的效果,提供一个大概的论证仍然是必要的。

毫无疑问,一切科学研究都必须造福于人类社会,但一切能够造福于人类社会的科学研究是否都必须直接创造经济效益,那就不一定了。

首先,从科学发展史上看,人类的科学研究兴趣,常常不在于直接创造经济效益,而在于探索未知世界。科学家们为大自然所呈现

出的和谐美所震撼,把揭示这种和谐背后所蕴含的奥秘看作是对人类本身的一种挑战。众所周知,牛顿的力学理论体系和元素周期律的发现,是科学革命的二个重要里程碑,对人类文明的发展产生了巨大的推动作用。然而,这些知识体系的建立,却需要大量的理论思维工作。没有高度抽象的理论思维来概括各种离散的实验数据,这种知识体系的建立是不可思议的。研究中国文化史的人,大多得出相近的结论^[1]:使中国这个具有五千年文明历史的古国失去科学革命的机会,导致近代科技落后的原因之一是传统思维中重经验轻理论,不能将经验上升到一般理论之整个社会文化素质的先天缺陷。虽然,在中国历史上空谈理论家不乏其人,然而,他们所相信和所建立的理论并不是能够指导人们认识客观世界的知识本身,而是理论本身所能带来的功利目的。依我所见,一种好的理论至少应该满足下列三个标准:①可理解性:即用清晰简明的语言来描述其概念体系,人们经过适当的努力就能大体准确地掌握其知识体系。②客观性:即给接受者提供一个对理论本身进行自我评价的可能性。③效用性:即应该给接受者提供一个有效地分析和解决有关问题的工具。换言之,一种好的理论不应该模棱两可,令人不知所云;不应该是理论家自己想怎么说就可以怎么说,而接受者自己不能对之进行自我客观的审定;不应该是仅仅满足审美或道德上的完善需要,而忽视了认识世界

这个最重要的功能。从这个意义上讲,不难发现,历史上的许多理论都缺乏这些最基本的标准。

人类一直对人的智能本身所着迷,揭示人类的智能奥秘一直是对科学家的一种挑战。电子计算机的诞生,使人类发现了一个可以较好地类比大脑的模型,人工智能学科就应运而生了。尽管,对于人的大脑是否本质上只是一台计算机这个问题仍有许多争论,但认知心理学的近几十年的结果表明,大脑的信息处理模型是人类认识中到目前为止最好的大脑类比模型。尽管目前人工智能所能取得的成就与早期人工智能科学家所持的过于乐观的估计相差甚远,但人工智能理论对相关学科如心理学、哲学、逻辑学、语言学、乃至社会学所产生的影响仍然是激动人心的。这种影响也以下面将要提到的关于知识的推理的发展而表现出来。可以这样断言:除非人类对人的思维本质问题有其它较大的突破,否则人工智能将伴随着计算机科学的存在一直成为科学家感兴趣的课题之一。

实际上,许多人工智能的理论问题都来自如何运用计算机来实现人的部分智能这个非常实际的问题,但许多人工智能理论研究者们在理论层次上付出了大量的工作却不能使有关工作成为具体的应用系统。非单调推理的研究就是一个明显的例子。然而,这个趋势的存在并不表明理论研究与实际工作的脱节,相反正说明问题本身的高难度。从方法论上讲,这还是有道理的。如果我们对每个问题首先想到的是如何尽快地把它变成具体的应用系统,而忽视了问题本身所涉及的许多基础问题,表面上看来这似乎是跨出了一大步。然而,当问题有所改变时,我们又得重新设计新的系统。由于各自所面临的具体应用问题各不相同,其它研究者也得根据各自的情况设计系统,相互之间缺乏一般性的知识可以共享,各自都是在原地跨出一大步,又回到原地这样的循环。如果事先就尽力把基础问题及其所有的可能性搞清楚,我们就能实现研究者之间的最大知识共享。尽管在这个问题上一个研究者只能跨出一小步,但其

它研究者可以在这个基础上继续往前跨出一小步。从长远来看,这个面向基础的小步稳定地跨越,更能推进人类的文明发展。

这几十年的人工智能理论研究过程中,出现了许多与思维和推理有关的抽象概念和术语,它们大多数可以用优美的数学语言来描述。这些抽象的理论术语的出现大大地丰富了人们对思维问题的描述,也促使人类相互之间更有效地更能避免误解地进行知识传播。人类的思维能力的发展,也是一个从较为直观形象的概念体系到较为抽象的概念体系的发展过程。正像本文的论证一样,如果没有使用“智能、理性、推理”等抽象概念,就不能很好地表达作者的意图。特别是以数学抽象符号为媒介的抽象概念的表述,更能有效准确地实现知识传播,因为依赖自然语言的抽象概念常常使人产生误解更甚于理解。当然,我们也看到有些理论工作者故意搞一些玄而又玄的抽象概念,以势吓人,这是不可取的。

虽然人工智能的理论研究主要出于实现机器智能的考虑,但由于它关注于思维和推理问题,从而使其许多理论成果可相应于人类本身的社会实践。关于知识的推理这个学科分支所涉及的一系列问题,如相互知识和共同知识问题、多来源知识的协调问题和信念修改的问题也给我们在日常行动中提供了一个相应可参照的思考框架。如果我们不了解有关理论,我们只能依靠我们的经验和直觉。掌握了有关理论,使我们对有关问题就会产生更多有效的思考,使人思维更敏捷,思路更清晰,行动更理性。

当然,值得注意的是,本文的目的不是在宣称理论工作比实际工作重要。就中国的实际情况看,把主要的研究力量投入到与经济效益直接相关的领域是完全合理和必要的。本文的目的只是想说明:理论工作仍是十分需要的。综上所述,其重要性主要表现在:①理论思维是在更深的层次上影响人类的文明进程,使人类行动更理性;②理论思维把可能遇到的问题尽可能地事先考虑好,可避免具体工作的无效反复;③理论思维所产生的一系列抽象概念给人类提供了知识有效传播和

共享的可能性;④理论思维可以用于指导人们的具体实践,使人们思维更敏捷,思路更清晰,行为更理性。

所以,并不奇怪,许多发达国家在理论研究上投入了大量的人力和财力;也不奇怪,人工智能的理论研究虽然与早期预言家的预言相差甚远,但仍然是一个充满生机和活力的研究领域;更不奇怪,关于知识的推理这个仅仅诞生十年的新分支已成为计算机科学家、人工智能研究者、哲学家、心理学家、经济学家、社会学家和语言学家共同关注的领域之一。

下篇

一九九四年三月在美国加州面向太平洋的 Asilomar 会议中心召开了第五次关于知识的推理理论问题的国际会议。会上共宣读了两个特邀报告和二十一篇论文,内容涉及到计算机科学、人工智能、哲学、心理学、经济学、社会学、语言学和逻辑学。就所涉及的主题来看,主要有:知识和信念逻辑基础、行动逻辑、信念修改理论、对策论与推理、多主体系统、逻辑万能问题、知识与机器人规划、非单调推理、知识与进程理论、共同知识问题等。关于知识的推理这个学科的发展历史及其所涉及的一系列专业术语,读者可参见文[3,5,6]。

下面将主要介绍与计算机科学和人工智能直接相关的一些新动态和研究成果。

Hector Levesque 在会上作了《情形计算中的知识、行动和能力》的特邀报告。情形计算(Situation Calculus)最早是由人工智能创始人之一 John McCarthy 提出的一种用于描述行动及其变化的形式语言。后来,Reiter 在情形计算的基础上对框架问题提出了一种解决方案,但 Reiter 的方法只适用于确定性的基本行动。Levesque 在会上提出了一种涉及到知识和复合行动的框架问题的解决方案。

Darwiche 和 Pearl 在题为《迭代信念修改的逻辑》一文中,对原有的信念修改的经典假设(即 AGM 假设)提出了四条新的假设。他们认为原有的 AGM 假设太弱而不能合理地处理迭代性的信念修改过程。他们提出的

新假设是:

(C1)如果 $\alpha \models \mu$, 那么 $(\psi \cdot \mu) \cdot \alpha \models \psi \cdot \alpha$.

(如果有两个新证据到来且第二个比第一个更具体,那么第一个证据是多余的,即仅第二个证据就能产生相同的结果。)

(C2)如果 $\alpha \models \neg \mu$, 那么 $(\psi \cdot \mu) \cdot \alpha = \psi \cdot \alpha$.

(如果有两个互相矛盾的证据到来且后一个压倒前一个,那么仅后一个证据就能产生相同的结果。)

(C3)如果 $\psi \cdot \alpha \models \mu$, 那么 $(\psi \cdot \mu) \cdot \alpha \models \mu$.

(如果一个新证据与原来的信息包含了另一个证据 μ ,那么在迭代修改中,证据 μ 应该被保留下来。)

(C4)如果 $\psi \cdot \alpha \not\models \neg \mu$, 那么 $(\psi \cdot \mu) \cdot \alpha \not\models \neg \mu$.

(如果一个新证据修改原来的信息后与另一个证据 μ 不相矛盾,那么在迭代修改后,也应不相矛盾。)

本文作者与 Peter van Emde Boas 在会上宣读了题为《多主体来源的信息获取》(Information Acquisition from Multi-agents Resources)的论文。我们关注的是从多个信息源中获取信息后可能会导致所获得的信息被所有原来信息提供者所反对这样一个不合理的问题。这个问题通常被称为 Schoenmakers 问题,是由 Schoenmakers 于一九八六年提出的。该问题曾被生动地用一个称为“判官疑难问题”(Judge Puzzle Problem)的故事来描述。Schoenmakers 通过这个故事提出了一个涉及多信息源知识获取和多个知识库合并共享等方面所存在的一个严重问题,即可能存在着一些不合理的结论,而如何去发现和这些结论却是十分困难的。我们针对这个问题提出一系列新的方案。首先我们证明了涉及有关问题的两个普通性定理。通过这些定理,我们知道了寻找 Schoenmakers 问题的绝对安全解是不存在的,而且不合理的结论只会出现在信息提供者的信念可能互相矛盾的情形下。这就给解决 Schoenmakers 问题确立了一个大体的方向,即寻找安全解就必须涉及到信息提供者的部分原有信念。然后,我们采用信念依赖的逻辑对寻找安全

解提供了一个可计算的算法,从而为解决 Schoenmakers 问题迈出了一步。信念依赖的逻辑是本文作者提出的一个用于描述和处理多主体间信念交流和依赖的逻辑^[4]。

在加拿大的一位中国留学生林进新的论文是《不协调情形下的协调信念推理》。由于经典逻辑不能很好地用来处理不协调的信念推理,而实际上一个主体的信念常常是不协调的,所以该作者引入了两种信念模态操作符 B^c 和 B , 分别来表示协调的信念和一般的信念。 B^c 具有分离规则(modus ponens)的能力,所以主体能够进行正常的协调性信念推理来得出协调性的结论。 B 允许不协调性,而且通过 B , 主体还能够对其不协调信念进行推理。该论文对这种逻辑提供了一个完备的公理系统,而且进一步讨论了如何运用这个新的逻辑系统来对一组主体中的间接知识进行推理以及如何从一个知识库中消除不协调的问题。

近年来,运用知识和时间逻辑进行分布式协议推理取得了许多成功的研究成果。Brafman 等人受此启发在文《知识作为带不确定性运动规划的一种工具》中探索了运用知识逻辑进行不确定性机器人运动规划的问题。虽然,在机器人的运动规划研究中,无论这种不确定性表现在控制还是在外部感知上,都已经有许多成果,但他们的论文的新颖之处在于:他们定义了这个领域中一个新的且自然的问题,即给定初始和目标位置,如何来获得一个正确而完备的终止条件。而且,他们提出了一个时间和知识逻辑,可以用来对不确定性机器人的运动规划进行推理。使用这个逻辑,对一大类运动规划都能得出正确和完备的终止条件的存在性结论,而且还可以较好地刻画和描述最佳的正确终止条件。

van Hulst 和 J.-J. Ch. Meyer 的论文《并行进程的知识证明系统》提出了一个使用分布式知识进行并行政程序的正确性证明系统。首先他们定义了一种类 Kripke 的程序语义,然后针对这种程序语义提出了公理化证明系统。最后证明了这个系统的正确性和完备性。

这种引入知识概念的并行政程序正确性证明系统可以较好地描述主体(即进程)间的相互知识特征。

值得一提的是,在这次 TARK 会议上有许多篇关于知识和对策论相结合的论文。现代对策论近几十年来突飞猛进的进展,使得它成为经济学和社会学上重要的理论工具。在这次会议上,我就有很深的感受,觉得现代对策论的许多方法可以用来很好地描述和分析实际生活中人际之间经济或政治冲突的动态特征,可以很好地帮助我们在实际生活中更为深刻地思考问题和分析问题。引入知识和逻辑的现代决策论如虎添翼,正在成为具有广泛运用前景的生机勃勃的学科之一。

颇引人注目的是这次国际会议的主席 Fagin 等人合作的新著《关于知识的推理》的出版^[2]。这本新著洋洋五百多页,汇集了这个学科的许多研究成果,并配备了大量的练习题,为这个新学科成为计算机科学及其相关领域的硕士和博士研究生的一门新课程起了有力的促进作用,也对这个新学科感兴趣的所有研究者提供了一本不可多得的参考书。

参 考 文 献

- [1] Fagin, Ronald, (ed.), Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1994 Conference, Morgan Kaufmann Publishers, 1994.
- [2] Fagin, Ronald, et al., Reasoning about Knowledge, The MIT PRESS, 1994.
- [3] Halpern, Joseph, (ed.) Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference, Morgan Kaufmann Publishers, 1986.
- [4] Huang, Zhisheng, Logics for Belief Dependence, in: Borger, E. et al. (eds.), Computer Science Logic, Lecture Notes in Computer Science 533, Springer, 274-288, 1991.
- [5] 黄智生, 关于知识的推理, 计算机科学, 1991第1期.
- [6] 黄智生, 信念修改的理论和方法, 计算机科学, 1991第6期.
- [7] 吴荣政, 王锦贵(主编), 简明中国文化史, 湖南师范大学出版社, 1991.