

人工智能的四大哲学问题

郑祥福

(浙江师范大学法政学院, 浙江 金华 321004)

摘要:当代西方哲学中的认知转向是和人工智能的研究协调发展的, 人工智能的哲学问题已不再是人工智能的本质问题, 而是关于人的意向性问题, 概念框架问题, 语境问题以及日常化认识问题。这些问题的解决将有助于人工智能科学的基础性问题的解决。

关键词:人工智能; 认知科学; 认识论

中图分类号: N03 **文献标识码:** A **文章编号:** 1003—5680(2005) 05—0034—04

20 世纪西方科学哲学的发展, 经历了向“语言研究”和“认知研究”的两大转向, 认识论的研究在不断去形而上学化的同时, 正在走向与科学研究协同发展的道路。作为当代人工智能科学的基础性研究, 认知研究的目的是为了清楚地了解人脑意识活动的结构与过程, 对人类意识的智、情、意三者的结合作出符合逻辑的说明, 以使人工智能专家们便于对这些意识的过程进行形式的表达。人工智能要模拟人的意识, 首先就必须研究意识的结构与活动。意识究竟是如何可能的呢? 塞尔说道:“说明某物是如何可能的最好方式, 就是去揭示它如何实际地存在。”^[1]这就使认知科学获得了推进人工智能发展的关键性意义, 这就是认知转向为什么会发生的最重要原因。

由于哲学与认知心理学、认知的神经科学、脑科学、人工智能等学科之间的协同关系, 无论计算机科学与技术如何发展, 从物理符号系统、专家系统、知识工程, 到生物计算机与量子计算机的发展, 都离不开哲学对人类意识活动的整个过程及其各种因素的认识与理解。人工智能的发展一刻也离不开哲学对人类心灵的探讨。无论是强 AI 派抑或弱 AI 派, 从认识论上看, 人工智能依赖于物理符号系统模拟了人类思维的部分功能, 但是, 其真正对人类意识的模拟, 不仅有赖于机器人本身技术上的革新, 而且也有赖于哲学对意识活动的过程及其影响因素的了解。

从今天来看, 人工智能的哲学问题已不是人工智能的本质是什么, 而是要解决一些较为具体的智能模拟方面的问题。这些问题包括:

1. 关于意向性问题

人脑的最大特点是具有意向性与主观性, 并且人的心理活动能够引起物理活动, 心身是相互作用的。大脑的活动通过生理过程引起身体的运动, 心理状态是脑的特征。“确实存在着心理状态; 其中一部分是有意识的; 大部分是具有意向性的; 全部心理状态都是具有主观性的; 大部分心理状态在决定世界中的物理事件时起着因果作用。”^[2]在这样的前提下, 塞尔认为, 计算机或人工智能是无法像人的大脑一样, 既具有意向性又具有主观性的。他对一些强 AI 观点提出了批评, 认为坚持这种观点的人, 把人的思维与智能纯形式化了。而计算机程序的那种形式化、语法化的特征, 对于那种把心理过程与程序过程视为同样过程的观点是致命的。因为人心不仅仅是形式的或语法的过程, 人的思想所包含的决不只是一些形式化的符号。实际上, 形式化的符号是不具有任何语义的。“计算机程序永不可能代替人心, 其理由很简单: 计算机程序只是语法的, 而心不仅仅是语法的。心是语义的, 就是说, 人心不仅仅是一个形式结构, 它是有内容的。”^[3]塞尔认为, 机器究竟能否进行思维的关键在于: 它是否能够给对象赋予意义。“意识、思想、感情、情绪以及心理的所有其他特征远非语法所能包容。不管计算机的模拟能力有多强, 按照定义, 它也不能复制那些特征。”^[4]在塞尔看来, 计算机模拟毕竟不是现实, 例如, 我们可以用计算机模拟风暴来临时城市周边防洪的形势, 用计算机模拟仓库的火灾, 用计算机模拟车祸发生的场景, 等等。但是, 谁会认为这种模拟就是事实呢? 心毕竟是种生物学现象, 其复杂性决非是计算机所能模拟与复制的。

计算机有没有意向性, 这个问题的争论可以归结如下:

【收稿日期】 2005—05—25

【作者简介】 郑祥福(1955—), 男, 浙江常山人, 哲学博士, 浙江师范大学法政学院教授, 研究方向为当代西方科学哲学。

1)究竟什么叫作意向性?机器人按照指令从事特定的行为是不是意向性?2)人类在行动之先就已经知道自己究竟是在做什么,具有自我意识,知道其行动将会产生什么样的结果,这是人类意识的重要特征。那么我们应该如何理解机器人按照指令从事某种行为呢?3)意向性能否被程序化?塞尔认为,“脑功能产生心的方式不能是一种单纯操作计算机程序的方式。”^[9]相反,人们要问的是:意向性是不是一种可以理解的精神,如果可以理解,那么为什么就不能程序化?塞尔认为,计算机具有语法,但不具有语义。但实质上,语法与语义本身就是二而一的问题,两者是从来也不会相分离的。如果两个机器人之间可以互相交流,那么我们难道能说它们之间只有语法而没有语义吗?如果程序能把语法与语义包含在一起,那么我们还有必要分清楚语法与语义吗?塞尔的观点是,哪怕计算机复制了意向性,但复制不是原本。其实,当对人类的认知及其与其行为的关系弄得一清二楚时,我们肯定能够把我们对人类大脑的心理过程与行为的关系编成程序,输入各种我们所了解的有关人类的信息,使计算机“无所不知”。然而,到了那个时候,我们是否还能像塞尔所说的,人工智能不是智能,人工智能中没有意向性和心理过程,因为它缺乏人类的蛋白质与神经细胞吗?意向性的复制是不是“意向性”?对理解的复制是不是“理解”?对于思想的复制是不是“思想”?对于思维的复制是不是“思维”?我们的回答则是:基础是有别的,功能是相同的。依赖于不同的基础形成同样的功能,人工智能只不过是人类智能的特殊的实现方式。塞尔用意向性来否定人工智能的深度,虽然有一定的根据,然而,当人工智能能够模拟出类似于人一样的思想时,即使人们都认为人工智能和人的智能是有着本质区别的,那么我们也会感到这种区别已经不具有什么重要意义了。塞尔的观点只能将人的心灵再度神秘化!

2. 人工智能中的概念框架问题

任何科学都是建立在它所已知的知识之上的,甚至科学观察的能力也无不与已知的东西相关,我们只能依赖于已知的知识,才能理解未知的对象。知与未知永远都是一对矛盾,两者总是相互并存又相互依赖的。离开了已知,就无法认识未知;离开了未知,我们就不能使科学认识有所发展和进化。“科学学习如何观察自然,而且它的观察能力随着知识的增长而增长。”^[9]大量的证据可以证明,当人们观察物体时,观察者得到的经验并非决定于以光线进入他眼球的信号,也不仅仅决定于观察者视网膜上的映像。两个正常的观察者从同一地方观看同一个物体,并不一定得到同样的视觉经验,即使在他们的视网膜上的映像实际上是一样的。正如汉森所说的那样,观察者在观看物体时,看见的比眼球接触到的多得多。所以,夏佩尔说,“观察者在观看物体时得到的视觉经验,部分地依赖于他过去的经验、他的知识和他的期望。”观察对于科学是十分重要的了,但是,“观察陈述必须用某种理论的语言构成”,“观察陈述是公共实体,用公共的语言加以阐述,包含着具有不同程度的普遍性和复杂性的理论。”^[7]这就表明了观察需要理论,科学需要理论为先导,科学认识不是建立在未知的基础上,而是建立在已知的知识基

础上的。

概念框架也称背景知识、背景信念。之所以将人们认知的概念框架称作信念,是因为概念框架是在不断地学习与实践中形成的,得到确证的那些可资利用的可靠的信息;这些信息在过去已被证明是非常成功的,我们对它没有理由怀疑。如果说背景信念有什么不确定性,那么我们可以说背景信念是不断在增长的、变化的,它处在不断的更新中。

人们认知结构中的概念框架究竟是由哪些元素构成的?对于这个问题,科学哲学家们仅仅将其理解为已知的知识。例如,“世界图景”(图尔敏)、“研究传统”(劳丹)、“研究范式”(库恩)、“背景信念”(夏佩尔)等等。为了解释概念框架,我们首先要将认知主体看作是一个处在复杂环境中的人,他不仅是一个科学的观察者,而且是一个社会的观察者、生活的理解者、情感的关系者,总之是一个社会的有着七情六欲的人。

我们根据一个人所处的社会环境与现实社会背景,便可以具体地归纳这个认知者所具有的概念框架是由哪些因素构成的。但是,由于概念框架是一个变量,如果我们不对智能模拟的目标加以限定,那么计算机编程就会面临指数爆炸的问题。因此,对人类智能的模拟就必须把机器人的目的加以限定,让机器人做特定的、有限的工作。人脑的活动是分区域的,那么对人脑意识的模拟首先应当分功能地进行。

概念框架问题是人工智能研究过程中最为棘手的核心问题,它所带来或引发的相关问题的研究是十分困难的。在这个问题上,基础性的研究是哲学的任务,即概念框架应当包含哪些因素,日常知识如何表达为确定的语句,人类智能中动机、情感的影响其状况是如何的,如何解决某些心理因素对智能的不确定性影响。而人工智能的设计者们则要研究这些已知知识应当如何表达,机器人如何根据概念框架完成模式识别,概念框架与智能机行为之间如何联系,概念框架如何生成、补充、完善,以及在运用这个概念框架某部分知识的语境问题,等等。而至于智、情、意的形式表达方面,则是人工智能研究者的任务。

3. 机器人行为中的语境问题

人工智能要能学习和运用知识,必须具备识别语言句子的语义的能力,在固定的系统中,语义是确定的。正因为这样,物理符号系统可以形式化。但是,在语言的运用中则不然,语言的意义是随语境的不同而有差别的。

实际上, AI也就是首先要找到我们思想中的这些命题或者其他因素的本原关系、逻辑关系,以及由此而映射出构成世界的本原关系、客体与客体之间的关系。最初的物理符号系统便是以此为基础的。但是,由于人们的思想受到了来自各方面的因素的影响,甚至语言命题的意义也不是绝对确定的单个句子或原子命题的意义更是如此。因此,最初,简单的一些文字处理与符号演算完全可以采取这种方式,但进一步的发展,例如机器人之间的对话、感知外界事物、学习机等,就必须在设计时考虑语句所使用的场合及各种可能的意义。

我们再回到维特根斯坦思想的发展。维特根斯坦的早

期思想在哲学研究中遭到了来自各个方面的批评,主要的问题是语言的日常用法,是不可能按照维特根斯坦规定的那样来使用的。在日常的使用中,语言的实用用法即语境决定了语言命题的意义。“哲学不可用任何方式干涉语言的实用用法;因而它最终只能描述语言的用法。”^[9]任何语言总是有确定的意义的,关键在于它是处在什么样的场合中,如何使用,即用于一个特定场合的句子其意义是确定的,否则这个句子就无法为人们所理解,就无所谓意义,所以“我们无疑懂得这个词,而另一方面,它的含义就在于它的用法”^[9]。

约翰·奥斯汀则把语言划分为两类:一类是记述式的,或者说是陈述,具有真或假的特性;另一类是完成行为式的话语,“它要完成它所特有的任务,这就是被用于完成某种行为。说出那样一种话语,就是完成某种行为。”^[10]“完成行为式话语必须是在特定的环境中说出,这种环境与有关的行为在各个方面都是合适的。”^[11]完成行为式的话语在我们日常生活中往往体现为一定的效果,即如果这件事是如此这般地发生了,那它便是正常的,如果相反,那它便是不正常的。在奥斯汀看,我们没有纯粹的语言标准,把陈述式或完成行为式分离开来,例如火车站广播员说,“请各位旅客在越过铁轨时通过天桥行走”,它既是直陈式的,又是完成行为式的句子。不管奥斯汀对句子的划分存在什么样的问题,但是,我们从维特根斯坦后期思想、奥斯汀的语言哲学来看,语境问题是确定语言意义的极为重要的方面。

由此看来,由于语境问题的存在,人工智能的语言编码就不是一种纯粹逻辑的技术,也不是一个纯粹认知心理学的问题。功能模拟在产生特定思维过程中已经显现出超人的有效性,记忆、知觉、意象、概念形成、问题解决、言语理解等等,都被看作可用实验来检验理论的研究领域,有关这些领域,提出了叠代、递归、组块、后进先出存储、水平搜索、垂直搜索、几何图形编码或其他信息加工的组合。但是,人的精神状态以及语言在实际中的运用,则是一个更为复杂的领域。如果要模拟人的行为,那么就必然会涉及到语言及其运用,涉及到语境的问题。

福多和拉普提出了“内容整体论”概念,认为一个表达式只有作为整个语言的部分时才能够有内容。如果我相信命题 P,那么,我就得相信与 P 处于语境关系中的各种命题。内容整体论是与语言原子论相对而言的,这种原子论相信,表达式可以通过自身与语言之外的实体的关系而具有意义。“我们所说的内容整体论是认为,具有内容这种属性是整体性的,在这种意义上,只有当语言中的许多其他(不同义的)表达式具有内容时,某个表达式才能够具有内容。”^[12]当然,持内容整体论观点的最早也许是“奎因-迪昂”原理,主张句子命题的意义必须是句子处在一个整体的科学理论语境中才能确定。

而自弗雷格对词义的分析时提出语境原则以来,语境论现已普遍成为语言哲学所探讨的重要方面。弗雷格认为,一个词的意义只有在句子的语境中才有意义,例如“暮星”与“晨星”虽然是相同的对象,但其语境不同,意义也是不同的。达米特在语境论基础上,提出了“从属原则”,即如果一个词

独立于它出现时的句子,那么我们就无法确定它的意义。同样,没有领会整个思想,我们就不能掌握构成这个思想的组成部分。

在哲学上,语境论是在“概念的相对性”提出之后形成的,它作为反对形式化的一种观点,即反对人们认为可以建立一套能被普遍应用而无须考虑特殊情况的抽象形式,或者我们可以通过研究一个陈述的逻辑结构来确定它的含义的观点。语境论者包括道德语境论(伦理情景主义)与哲学的语境论。“在语言哲学中,语境主义提出,一个词的意思是通过它在一个句子中的用法或出现来确定的,也就是说,通过它对句子内容的作用来确定的。因此,句子或命题在对意思的解释上比词或概念更重要。理解一个词就是理解它如何被用在一个句子中。语境定义是以语境主义为基础的,它指的是:解释一个词,要求助于它出现于其中的句子。根据对语境主义的类比,奎因和戴维森创立了人们所说的语义和整体论,这一观点认为,一个句子的意思是由它在整个语言中的用法确定的。”^[13]奎因说道:“即使假定终究可以给同义性的概念提供一个令人满意的标准,但是,这种做法仍然只是考虑了‘意义’这个词的一种语境,即‘意义相似’这个语境。语词是否也有语言学家们应予以注意的其他一些语境呢?是的,语词肯定还有另一种语境,即‘具有意义’这个语境。”^[14]因此,我们说同义的问题时,就得确立情境的相似性,但是,“没有两种情景是完全相似的,即使在其中说出同一形式的情境也有无数的差别。”^[15]

鉴此,人工智能在设计语言编码时,就不得不考虑整个思想以及言语的各种情境条件对于句子意义的制约作用。然而,困难在于找出那些与语言情境有关联的主观成份,而对于后者,则几乎是不可能的。因为外部情境是一个极不确定的因素,每一个场景都是不相同的,这只能根据社会文化的类型大致确定几种不同情境类型,社会化的认识论则将在这方面提出它们自己的见解。

显然,从事对语境问题研究的哲学家们推动了这个转向。实际上,把世界形式化,或者形式地理解智能行为,直至目前人们认为是极为困难的,这些困难用形式化的方法是无法逾越的。由此可见,西蒙和纽厄尔的符号程序已经走到了终点。致命的局限性是由于符号程序没有看到信息加工系统是动态的、相互作用的、自组织的系统。而罗森布拉特则清楚地看到了这一点。

4. 日常化认识问题

人工智能模拟不仅要解决心身关系,即人脑的生理与心理的关系问题,而且还必须解决人脑的心理意识与思维的各个层次间的关系,以及人的认识随环境的变化而变化、随语境的变化而变化的问题。根据智能系统的层次性分析,我们可以逐步做到对各个层次的模拟,但是,智能层次性分析也只是一抽象化的分析或理想化的分析而已。实际的智能是多个层次之间不可分割的相互关联着的整体,各层次间究竟是如何发生关联的?在什么情况下发生什么样的关联?这便涉及到日常化的认识问题。

因此,建立在符号系统基础上的人工智能无法解决人类

认识的日常化问题,特别是无法解决人脑的情感、动机、意向性等心理活动功能,无法解决我们的日常认识因语境不同而意义不同的问题。

现象学大师胡塞尔则认为,世界、思想的背景、日常语境等是一个非常复杂的系统,这个系统是与那些同个人复杂的信念体系等相联系的事实组成的。但是,原则上说,我们可以将自己在世界中的存在悬置起来,而完成对人类信念系统的独立描述,这样,我们智能行为的基础就可以得以确定。但是,海德格尔却反对其导师胡塞尔的观点。我们所生存的世界,我们说话的语境等等我们日常应付自如的方式,是我们社会活动的一部分,是我们的存在方式。这种存在方式是我们无法像胡塞尔那样将它抽取出一定的成分确定化的。明斯基指出:“仅是建构一个知识基础,就成为智能研究的重大问题……关于常识性知识的内容和结构,我们还是知道得太少了。‘极小’常识系统必须‘知道’有关因果、时间、目的、地点、过程和知识类型……某些情况。在这一领域中,我们需要花力气做严格的认识论研究。”^[19] 1970年后,在明斯基的倡导下,人们开始研究“微世界”领域,打算形成系统地处理知识的方法,并且人们希望这些限定的、孤立的微世界能够逐步变得更接近现实,并且能早日成为通往现实世界的理解手段。但是,最终人们发现这种研究在目前的情况下过于困难。因为,关键的问题是我们必须在面对常识的研究中形成一组抽象原理,以与常识理解相对称。“但是,人类很可能根本不是按照通常的方式使用常识性知识的。正如海德格尔和维特根斯坦所指出的,与常识性理解相当的,很可能是日常技能。所谓‘技能’,并不是指过程的规则,而是指在众多的特定场合知道该做什么。”^[17] 如果是这样的话,理解技巧就不是以某种确定的规则为基础的了,例如道德语境、审美情境等等。那么,基于符号的AI就无法对这样一种人们在特定的语境中所做出的特定行为做出模拟了。这就迫使AI的研究从符号操作理论转向神经网络模型的建构。

由于物理符号系统是一个物理的过程,它与人的心理活动有着根本的区别。因此,要模拟人的心理与意识,在原则上是不可能的事。但是,模拟人的意识所要解决的问题主要是三个方面:

首先,限定人工智能的目标。必须使单一的机器人对人的意识的模拟特定化,即不要让机器人做太多的过于复杂的

事情,不要使编程陷入指数爆炸。根据西蒙的观点,人的认识就是解决问题,而人在解决问题时是先易后难、逐个逐个地解决,而不是一次解决无数个复杂问题。

其次,必须把人的心理与意识分层次进行模拟。只要能够展示出人脑的功能,那么尽管物理的过程与人的生理、心理的过程是截然不同的,但从功能上来说则是等效的。但是,要分层次地进行模拟,心理学的实验则是必不可少的。即为了给观察到的人类行为建立模型而编制符号系统程序,心理学对参试者观察与实验的结果便可作为构造物理符号系统的假设。

再次,必须建立理解意义的各种条件性假设。由于语句的意义在使用过程中是变化的,其意义随语境的不同而不同,这就需要在编程时设计出各种条件性假设,不同的语境有着不同的条件,只要我们设计出这些条件,那么其意义就得以确定。

【参 考 文 献】

- [1][2][3][4][5] 塞尔·心、脑和科学[M]. 杨音莱译. 上海:上海译文出版社, 1991. 15、19、23、28、30—31.
- [6] 夏佩尔·理由和求知[M]. 褚平,周文彰译. 上海:上海译文出版社, 1990. 380.
- [7] 查尔默斯·科学究竟是什么?[M]. 查汝强等译. 北京:商务印书馆, 1982. 34.
- [8][9] 维特根斯坦·哲学研究[M]. 陈嘉映译. 上海:上海人民出版社, 2001. 75、121.
- [10][11] 奥斯汀·记述式与完成行为式[A]. 戴涂纪亮主编·语言哲学名著选辑[C]. 北京:三联书店, 1988. 202、203.
- [12][13] 转引自布宁·西方哲学辞典[S]. 余纪元编著. 北京:人民出版社, 2001. 194、196—197.
- [14][15] 奎因·从逻辑的观点看[M]. 江天骥等译. 上海:上海译文出版社, 1987. 45、56.
- [16] 转引自博登·人工智能哲学[M]. 刘西瑞,王汉琦译. 上海:上海译文出版社, 2001. 440.
- [17] 德雷福斯·造就心灵还是建立模型:人工智能的分歧点[A]. 载博登·人工智能哲学[C]. 442—443.

(责任编辑 成素梅)