技术哲学· 文章编号:1000-8934(2003)06-0043-04

关于知识-机器互动机制的可能性的探讨

王克迪[1] 傅小兰[2] 黄 斌[3]

(1. 中共中央党校 哲学部,北京 100091;2. 中国科学院 心理研究所,北京 100101;
3. 中共中央党校 研究生院,北京 100091)

摘要:作者提出运用计算机系统对修正了的三个世界相互作用理论和逻辑推理进行模拟验证,使之臻于完善。建议的实验方案是,在类似于虚拟现实系统的预设运算环境中加入非人为控制的知识或物理信息,期待某种新知识产生。意义在于,从科学技术与哲学社会科学相结合角度提出一种解释人-机关系新理论,提出未来社会中知识生产新机制,尝试运用计算机系统模拟验证哲学概念推理。

关键词:知识;人-机关系;模拟

中图分类号:N0 **文献标识码**:A

本文根据修正了的波普尔三个世界理论,提出一种知识-机器互动机制,分析运用现代计算机技术对这一机制进行模拟验证的可行性和具体方案,并对有关问题进行讨论。

1

卡尔·波普尔(Karl Popper)指出,"如果不过分认真地考虑'世界'或'宇宙'一词,我们就可以区分下列三个世界或宇宙:第一,物理客体或物理状态的世界;第二,意识状态或精神状态的世界,或关于活动的行为意向的世界;第三,思想的客观内容的世界,尤其是科学思想、诗的思想以及艺术作品的世界。"[1]波普尔把这三个世界分别称为世界1、世界2和世界3。从1960年代初直到1994年他去世,波普尔比较系统地发展了这一理论,尤其着重讨论了世界3的存在方式、特点和它与其他两个世界的关系。本文不可能全面充分地阐述波普尔三个世界理论的全部内容和理论发展变化的过程,过去20年来,国内外学术界已经就波普尔三个世界理论进行过大量的介绍、研究和讨论。下面只就波普尔理论中与本文主题密切相关的三个部分作简要介绍。

(1)世界3的存在方式 在与脑科学家约翰·艾克尔斯(John Eccles)合写的《自我及其大脑》(The Self and its Brain)中,波普尔谈到,"有些世界3客体只存在于编码的形式中,如乐谱(可能从来不会演奏)、或者唱片录音。其他的——诗,可能还有理论——也可以存在于世界2客体中,如记忆,估计也以记忆痕迹的方式编码而存在于某些人的大脑(世界1)中,并随着大脑死亡而消失。"[2]波普尔还提出,"不仅地图和计划(plans)是世界3客体。行动计划(plans of action)也是,这可能包括计算机程序(computer programmes)。"[3]这意味着,波普尔认为世界3主要以语言或某种编码形式存在,承认这一点是重要的。

(2)世界3的编码特性与内容不变性 波普尔认为,内容是人类语言的产物,而人类语言反过来又是最重要、最基本的世界3客体。语言有其物理的方面,而所想或所说的内容则是某种更加抽象的东西。"内容正是我们在从一种语言翻译成另一种语言中想加以保存、保持不变的东西"[4]。在波普尔看来,语言是一种载体工具,客观知识、内容在它之中的体现,犹如理论或剧作在书本中的体现。同样的思想,可以翻译为不同的语言,也可以采取不同的编码(从本质上说,编码就是一种翻译)。"客观意义的知识不包括思想过程而包括思想内容。它包括我们用语言所表述的理论的内容,这一内容可以、至少可以近似地从一种语言翻译成另一种语言,客观思想内容是在合理的优良翻译中保持不变的内容。或者按照更加实在主义的说法,客观思想内容就是翻译者力求保持不变的东西,即使他会不断地发现这

收稿日期:2002-11-01

作者简介:王克迪(1959-),安徽合肥人,哲学博士,中央党校哲学教研部科技哲学教研室教授、副主任,中国自然辩证法研究会理事,研究方向:信息哲学,科学技术史。傅小兰(1963-),湖南长沙人,心理学博士,美国密西根大学博士后,中国科学院心理研究所认知与复杂信息环境实验室研究员、主任,博士生导师,中国心理学会理事、国际学术交流工作委员会主任,美国心理学会外籍会员,北京市青联委员,研究方向:认知心理学。黄斌(1973-),湖北荆州人,中央党校研究生院硕士研究生,研究方向:信息哲学,科技政策。

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

一任务难得简直不能完成。”[6] 简而言之,人们总是可以从某种表现中抽象出一个思想或一段信息,这个思想或这段信息不会因其语言或编码形式而有本质变化。

(3) 三个世界之间的相互关系　世界 1 与世界 3 之间以世界 2 为中介,这一点很清楚地包含在三个世界的理论中。波普尔指出,科学猜想或理论能对物理世界产生效果。波普尔进一步指出,“精神在第一世界与第三世界之间建立了间接联系。这一点极为重要。无法否认,这种由数学理论和科学理论组成的第三世界对第一世界产生巨大的影响。比如,由于技术专家的介入确实能产生这种影响,技术专家通过应用上述那些理论的某些成果而引起第一世界的变化。”[6]“世界 2 作为世界 3 和世界 1 之间的中介而发挥作用。但也正是对世界 3 客体的把握(grasp)给予世界 2 以改变世界 1 的力量。”[7]

显然,波普尔所说的三个世界相互关系,是一种简单的直线关系,即世界 1 与世界 2 可以直接相互作用,世界 2 与世界 3 可以直接相互作用,但是世界 3 与世界 1 不能直接相互作用,世界 2 是两者之间的关键性中介。基于此观点,我们可以推测,科学知识的增长主要靠人的精神活动(世界 2)与知识(世界 3)相互作用(例如,科学史研究),或者精神活动与客观世界(世界 1)的相互作用(例如,一般意义上的科学研究)实现,但是,知识与客观物质世界不能直接相互作用。

无须否认,波普尔关于三个世界的划分对于我们理解知识的增长和生产仍然具有原则性的指导意义,波普尔规定的三个世界之间的关系并没有发生根本的变化,世界 3 与世界 2、世界 3 与世界 1 分别相互作用仍然是今天的知识生产的主要形式。然而,随着信息技术成为重要的知识传播和处理手段,由于计算机技术的介入,三个世界及其相互关系出现了新的情况。在以计算机为核心的现代信息处理系统中,作为人工自然的计算机裸机系统可视为一种世界 1,而采用数字编码形态的、含有时间序列的计算程序软件可视为一种世界 3,那么世界 3 与世界 1 就可能发生直接相互作用。[8] 在这种情况下,波普尔的三个世界关系由直线形转变为环形,其中任何两个世界都可以发生直接的相互作用。在计算机技术日益广泛应用的今天,对波普尔三个世界理论尤其是他的三个世界相互关系理论进行修正,是十分必要的。

2

本文的基本思想是对上述概念推理进行机器验证,由于关键之点在于世界 3 与世界 1 的直接互动,因而称为知识 - 机器互动系统。提出的基本设想包括三个方面:①把上述机器(计算机)与程序的直接互动推广为一般的知识 - 机器互动关系;②体现这样的关系的系统可以作为新的知识产生机制;③对上述关系与机制进行验证演示。

上述三个方面中的前两个已经作为哲学推理得出,并获得一些科学研究实例的支持(如虚拟现实系统、发现洛仑兹吸引子的计算机运算系统等,详见注释 8 文献)。因此,本文重点在于阐述如何实现上述设想中的第三点,即如何对知识 - 机器互动关系与机制进行比较严格意义上的科学模拟验证。

(1) 模拟验证的可行性分析　鉴于以下三个方面的进展,我们认为模拟验证设想已具备较充分的可行性。

第一,Michael Heim 借用中国古老的“风水”概念,专指虚拟现实技术中用于描述物理科学刻画的本体论世界的信息之流,其中居于核心地位的是世界与事件发生与进行的过程[9]。Heim 指出,虚拟现实及其客体建构描绘了对于现实世界存在物的理解,这种理解已经在 20 世纪后半叶哲学中准备好,现代先进的计算机和网络技术是检验这些存在物特性及其运动的实验台。

2000 年以来,斯坦福大学知识系统试验室(KSL) Thomas R. Gruber 等人将人工智能研究与下一代万维网研究结合,提出所谓“本体论”(ontology)[10] 系统,利用高性能计算机对存储信息的内容进行语义分析,实现机器智能,预期在电子商务等领域将得到广泛运用[11]。有关思想和实验模拟代表了有关研究的最新进展。

第二,人工智能研究证明,原则上可以用图灵机模拟人类思维[12],信息处理技术最终可以实现机器对人类智能的模拟。我们对此一“强人工智能”立场有一定保留,即不认为现阶段的机器模拟可以真的模拟人脑思维机制,但认可机器模拟在科学研究上的合理性。

第三,20 世纪 80 年代以来的复杂性研究大量使用的计算机模拟试验证明[13][14][15],通过适当编制程序,在适当的外部条件下,大量粒子或其他成员组成的随机系统有可能实现自组织功能或表现出某种智能。

上述所有例证中,计算机程序作为一种时间序列相关的文本被引入世界 3 是关键性的,知识与物质(准确地说是人工自然或机器)的互动是核心。其本质在于,机器系统(包括人工智能系统)与其说是模拟人脑的思维机制,还不如说是运行人脑的思维

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

产品。在这里,重要的也许不是这样的机器系统能否模拟人脑机制,而是这样的机器系统及其运行环境能否真的实现知识与机器的互动,并创造出新的知识产品。

(2)模拟验证的具体设想 本文提出的模拟验证的具体设想是,尝试运用计算机系统对上述经过修正的三个世界相互作用理论和逻辑推理进行模拟和验证,并使之臻于完善。

进行模拟验证所运用的计算机系统表面上类似于改进的虚拟现实系统或一般意义上的人工智能系统,本质上却有较大区别。传统的人工智能系统,本质上也是机器与知识的互动系统,其主要目的在于揭示知识(程序)所蕴含着的未知知识,或者说揭示已有的知识(表现在编制好的机器的运行程序中)的尚未展示给人的自主性。可以推断,新知识的产生不会超过预设程序的蕴涵,机器的作用仅在于加快或者尽可能充分地展示这一过程,当然,这样的过程一般是人脑(世界2)不可能或者没有耐心加以揭示的。

然而,预设程序(世界3)中包含的内容的自主性或超越性(波普尔术语)还不意味着未知知识的全部,即使是最完美的程序,也不可能包含全部的未知知识,甚至也不能包含全部的已有知识。因此,我们所能够做到的仅仅是考虑来自两个方面的变化因素,即在预设的运算环境中加入非人为控制的知识(世界3)和物理世界(世界1,非线性)的信息,在这种情况下,世界3与世界1之间发生相互作用时,我们有理由期待更多的新知识产生,其形态为编码信息,可以是概念、图形、数据,也可以是某种规律或事实,甚至概念本身。

(3)模拟验证的目标 实现模拟验证的目标在于,通过建立一个能够进行计算机验证和演示、适合解释现代信息技术条件下知识、客观世界与思维主体关系的相互作用以及新知识产生的理论和概念体系,在高性能计算机系统上设计具备知识表达、概念推理和运行结果演示功能的开放性系统,对修正了的三个世界理论进行模拟和检验,并在实证和文献研究的基础上对上述理论进行充分论述与发展。

为达到上述目标,需要重点解决两个关键问题:①设计一个精巧试验,在一个(对物理世界和知识世界)信息开放的系统中模拟知识与物理世界的互动关系,该实验的关键部分是确定知识的产生机制及其算法表达。系统的开放性通过输入某种信息实现,途径之一是输入外部物理世界的非线性信息,拟由外部传感设备通过模数转换接入计算机实现;途径之二是直接输入某种知识。②设计、组建高性能知识 - 机器互动系统,这包括较强大的运算环境,受到编程控制驱动的外部设备,可以将外部世界信息转化为知识或编码信息的传感设备等。图1是模拟实验方案示意图。

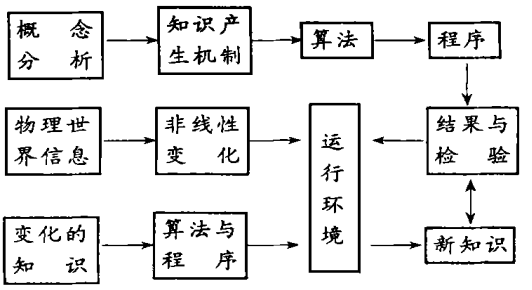

图1 知识—机器互动关系模拟实验示意图

3

本文提出的设想具有科学和哲学理论意义,表现在三个方面:首先,从科学技术与哲学社会科学相结合的角度提出一种解释人 - 机关系的新理论,实际上是把传统的人 - 机关系研究转换为知识 - 机器关系研究,可能将有助于人工智能与机器认知研究,至少为虚拟现实研究提供一种新思路,作者将就此另文专论。其次,为人 - 机系统嵌入一种新的人为控制关系,即知识控制因素,提出未来社会中知识生产新机制。第三,尝试利用计算机模拟哲学概念和进行推理,为哲学社会科学研究信息社会问题提供一个可供选择的理论平台[16]。

从方法论角度来看,本文设想缘起于哲学理论推理和概念分析,采用了一些来自科学研究,主要是虚拟现实研究、复杂性研究、混沌学研究和计算机数值模拟等方面的实例和证据,再通过引入一些新概念廓清原先的一些概念,从而达到修正原有理论(波普尔的三个世界理论)的目的。本文提出的模拟验证的基本目的是从科学意义上检验这种修正,将这种主要由哲学推理和概念分析得来的理论认识回归到科学层面的实验模拟研究,其方法论上的主要依据将是哲学推理与科学研究共同遵循的逻辑推理规则,这种在方法论上合法性是本文设想的合理性的基础。

可以期待本文提出的模拟验证将出现积极的结果,它将提示,在现代人 - 机关系研究中,哲学研究成果,特别是充分适应了现代科学发展新形势的哲学理论,有可能给予科学研究以思路上的帮助,有时甚至是直接的指导。这种帮助或影响是客观存在的,并不依某个人的意志和喜好而改变。众所周知,在以电子计算机为核心的人工智能研究领域,近几十年来,虽然来自硬件和软件两方面的进展极为巨大,但是仍然没有取得根本性的突破,迄今有关人工智能研究还存在着乐观派(例如"强人工智能派")和

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

悲观派(例如"反人工智能派")两种极端的重大分歧。对立双方都有相当多的实例和逻辑依据作为支持,难以取得共识。我们相信,最终解决这样的问题可能在一定程度上依赖于哲学理论的重大发展。

人工智能的本质在于模仿人脑的知识生产机制。人工智能派的信念是可以制造这样一种机器系统,利用它完全可以模拟人脑的思维;反人工智能派则主张在人脑机制未充分明了之前,所谓机器模拟无异于缘木求鱼,而人脑之复杂绝非今日之科学技术可以穷究探底。显然争论的关键点在于对人脑思维机制的认识和模拟上。本文另辟蹊径,建议采取不同的思路,不直接触及人脑机制问题,而采取间接方式,把人脑的思维能力通过编程即人的思维产品来表达,也就是把直接模拟世界 2 转换为用世界 3 来代替这种模拟。由于世界 3 的客观性容易获得,具有与世界 1 一样的客观性确认,这样的机器系统的合理性至少相对易于获得论辩双方的接受,因而它的实验结果也将相对引发较少争议。因此,在某种意义上,本文提出的模拟验证应该能够作为一种对于人工智能研究问题的解决方案来对待。

也许更重要的还在于,本文提出的实验检验结果还将再次回归到哲学层面:证实了,将表明哲学理论的修正是正确的,因而这种理论获得重要的科学基础;否证了,则将出现较为复杂的情况,它可能提示哲学推理或概念分析有误,也可能表明引入概念失当,还可能证明(在这个理论修正中)把科学研究事实提升到哲学概念层面的认识中具体做法有重大偏差,当然也不完全排除对于这种提升方法的全盘否定。此外,模拟研究中的试验设计、编程说明甚至编程本身乃至系统运行过程中都有可能出现差误,导致验证失败。在这种情况下,将会提出许多新课题。但是无论如何,哲学研究都会从中得到一定的启发。

最后,还有一个机器系统产生的"知识"的认证与鉴别问题,这将既包含知识形态的判断,也包含逻辑判断,还包含价值判断等多种因素。我们以为,这个问题不是目前可见的以及将来可预见的机器系统能够自行解决的,它需要人脑的参与来解决[17]。正是在此意义上,作者认为,人终归有人的用处。此外,有关问题的解决,还依赖于信息处理技术的长足进步。

参考文献

[1][6] Karl Popper. **Objective Knowledge: An Evolutionary Approach** [M]. Oxford University Press, London, 1983. 106, 155.

[2] Karl Popper, The Worlds 1, 2 and 3, in Karl Popper and John C. Eccles, **The Self and its Brain** [M]. Springer International, New York, 1977. 41.

[3][4][5][7] Karl Popper, Three Worlds, **The Tanner Lectures on Human Values** [C] ed. by Sterling M. McMurrin, University of Utah Press, Salt Lake City, 1980. 163, 159, 156.

[8][16] 王克迪. 赛伯空间之哲学研究 [M]. 当代世界出版社, 北京, 2001. 83 - 108, 序言.

[9] M Heim. The Feng Shui of Virtual Reality [J/OL]. **Crossings: eJournal of Art and Technology** 1.1. 见网站: http://crossings.tcd.ie/issues/1.1/

[10] T R Gruber. A Translation Approach to Portable Ontology Specifications [DB/OL]. **KSL** 92 - 71. 见 KSL 网站.

[11] N F Noy, et al. Ontology Development 101: A Guide to Creating Your First Ontology [DB/OL]. **KSL** 01 - 05. 见 KSL 网站.

[12] A 图灵. 计算机器与智能 [A]. M 博登. 人工智能哲学 [C]. 上海: 上海译文出版社, 2001.

[13] M 沃尔德罗普. 复杂 [M]. 北京: 三联书店, 1998.

[14] M 海姆. 从界面到网络空间 [M]. 上海科技教育出版社, 上海, 2000.

[15] Y. 卡斯蒂. 虚实世界 ——计算机仿真如何改变科学的疆域 [M]. 上海: 上海科技教育出版社, 1998.

[17] 王克迪, 什么是虚拟现实? 应如何看待虚拟现实? [A]. 中共中央党校哲学教研部编. 哲学热点问题释疑 [C]. 北京: 中国城市出版社, 2002. 75 - 82.

Research on the Possibility of Knowledge - Machine Mutual Interactive Mechanism

WANG Ke-di[1], FU Xiao-lan[2], HUANG Bin[3]

(1. Philosophy Department, The Central Party School, Bejing 100091; 2. Institute of Psychology, Chinese Academy of Sciences, Beijing 100101; 3. Philosophy Department, The Central Party School, Beijing 100091, China)

Abstract: The authors suggest to modulate and demonstrate the modified mutual interaction theory between the three worlds of Karl Popper, using a computer system, and make it a more accepted one. The demonstration scheme in planning would be to add some non-human controlled knowledge or physical information to a prepared computing environment parallel to a virtual reality, and to expect some kind of new knowledge coming out. A new theory to explain the human-machine relationship from a cross view between scientific and humanism, a new mechanism to produce knowledge in the future society, and an attempt using computer to modulate and demonstrate the inference of philosophical conceptions, would be the meaning of the suggestion.

Key words: knowledge; human-machine relationship; modulation

(本文责任编辑　刘孝廷)

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net