

中图分类号: TP18 文献标识码: A 文章编号: 1004-8634(2018)04-0026-(08)
DOI: 10.13852/J.CNKI.JSHNU.2018.04.004

人工智能与人类智能: 两套概念, 两种语言游戏

李国山

(南开大学 哲学院, 天津 300350)

摘要: 人工智能的快速发展引发了一系列哲学问题, 尤其是人工智能与人类智能的关系问题。维特根斯坦虽未能见证人工智能的大发展, 却对机器感觉与思维乃至机器情感与意志问题做过不少论述, 表达出对人工智能技术的前瞻性见解。他通过对与人工智能相关的诸多概念的语法考察, 明确指出这些概念只有在人工智能研究领域内才能获得正当使用, 而一旦超出这一领域与人类智能的相关概念相混同时, 便会引发哲学上的混乱。这样的混乱正是维特根斯坦试图加以清除的。可借助他提出的“锤子比喻”来理解他关于人工智能与人类智能之关系的看法及其启发意义。

关键词: 人工智能; 人类智能; 锤子比喻; 语法考察; 维特根斯坦

维特根斯坦于1951年去世, 未能见证人工智能的蓬勃发展。然而, 作为一名早年从事工程技术研究的哲学家, 他一生都对技术领域的新进展保持着高度的兴趣和敏感, 并在哲学探讨中时常涉及技术哲学问题。他生逢人工智能萌芽时期, 又和图灵有过密切接触和激烈争论, 这势必引发人们关注他与随后兴起的人工智能的关系。我们知道, 维特根斯坦被公认为20世纪最重要的语言哲学家和心灵哲学家。他的思想成就已被运用于几乎每一个人类理智探究领域, 而像人工智能这样一个尖端的技术领域, 更少不了他的身影。事实上, 近几十年来, 不断有人探究如何从他的哲学思想出发来看待人工智能问题。学者们集中争论的话题包括: 维特根斯坦在何种意义上否认机器可以思维? 他所倡导的语法研究对于我们探究人工智能哲学问题有何启示? 本文拟从他的一个比喻入手, 结合他的语言哲学和心灵哲学思想进一步探讨上述问题。

在《蓝皮书》中, 维特根斯坦谈到“有意识的思想”与“无意识的思想”的区别时, 用到了锤子比喻。‘可是, 说某人在谈论有意识的思想和无意识的思想时是以两种不同的方式使用‘思想’一词, 这种说法是否正确呢? ——当我们用锤子敲进钉子, 或者当我们用锤子把木塞打进孔里, 我们是否以两种不同的方式使用锤子呢? 当我们把这个塞子打进这个孔里, 另一方面又把另一个塞子打进另一个孔里, 我们是否以两种不同的方式还是以同一种方式使用锤子呢? 或者, 当我们在一个场合下把某种东西打入另一种东西, 在另一个场合下把某件东西打碎, 我们是否只是在谈论不同的使用方式呢? 或者, 是否在所有这些场合下都是谈论锤子的同一种用法, 只有当我们把锤子用作压纸的重物时, 我们才能谈论一种不同的用法? ——在哪些场合下, 我们说的是以两种不同的方式使用一个词, 而在哪些场合下是以同一方式使用这个词? 如果我们说以两种(或更多种)

基金项目: 教育部规划基金项目“维特根斯坦语言哲学研究与文献翻译”(15YJA720002)的阶段性成果

作者简介: 李国山, 南开大学哲学院教授, 博士生导师, 主要从事分析哲学、西方认识论研究。

不同的方式使用一个词，那么我们并没有用这种说法提出一个关于它的用法的观念。我们只不过提出一个具有两个（或更多个）分栏的图表，用以描述它的用法，以便使人们能够看清楚这个词的用法。下面这种说法是正确的‘我们用这个锤子做两件事：我把一颗钉子敲进这块木板，把另一颗钉子敲进那块木板。’……当我们讨论有意识的思想和无意识的思想时，我们是以两种方式还是以一种方式使用‘思想’这个词？那些说‘这的确是两种不同的用法’的人，已经决定使用一个具有两个分栏的图表，并且用他所说的话表达了他的这个决定。”^①

维特根斯坦用这个比喻本来是想说明“有意识的思想”与“无意识的思想”这两个术语之间的关系。这里，我想借用它来探讨“人工智能”与“人类智能”这两个术语之间的关系。维特根斯坦生前也非常关注类似的问题，并做过一些思考。他在《哲学语法》《蓝皮书》《哲学研究》等著作中都探讨过诸如“机器思考”“机器牙痛”“人工脚”“脚上的人工疼痛”等话题。这些探讨虽显得零碎，却清楚地表明，维特根斯坦已敏锐地注意到人工智能的发展势头，并预见到了由其发展会引出的一系列哲学问题。面对这样一些问题，维特根斯坦主张：要像对待其他一切领域内的哲学问题一样，通过语法探讨去消解它们。

一

由于“人工智能”和“人类智能”这两个术语都包含了“智能”一词，所以，我们自然就会被引导到这样的问题上来：人工智能和人类智能是两种完全不同的智能吗？它们之间有相似性吗？人工智能会不会超过、甚至取代人类智能？诸如此类的问题都已超出了人工智能专家、脑生理学家或心理学家所能回答的范围，因为它们不再是专门的科学问题。在维特根斯坦看来，它们都是人们在语言的诱导之下提出的哲学问题。我们被这样的问题吸引，却又百思不得其解。这给我们带来深深的困扰。那么，到底该如何应对这些问题呢？

维特根斯坦在《哲学研究》中明确提出通过语法考察探讨哲学问题，并将我们所使用的语词比作工具：“想一下工具箱里的工具：有锤子、钳子、锯子、螺丝刀、尺子、胶水盆、胶、钉子、螺

丝。——这些东西的功能各不相同；同样，语词的功能也各不相同（它们的功能在这一点那一点上会有相似之处）。当然，我们听到这些语词，看到写出来印出来的语词，它们的外观整齐划一，而这让我们感到迷惑。它们的用法却并非明明白白地摆在眼前——尤其在我们从事哲学的时候！”^②

可以说，“人工智能”和“人类智能”这两个术语中的“智能”一词就相当于维特根斯坦说到的那把锤子，而要应对上述那些问题，就需要仔细考察这两个术语各自的用法。事实上，在他的锤子比喻中，维特根斯坦已清楚地指出了锤子的两种不同用途：敲击某种东西和压住某种东西。而我们倾向于认为，第一种用途是锤子的常规用途。第二种用途则不那么常见，不过，既然这样用了它，说它就实际发挥了镇纸的作用，也未尝不可。

那么，我们是不是可以说，与人类智能相关的“智能”一词的使用是常规性的，而与人工智能相关的使用虽显得反常却也不失为一种正当的用法，毕竟人们已经渐渐接受了这种用法？维特根斯坦会同意这种说法，因为他对语言用法不断变化和拓展的现象做过非常形象的描述：“我们的语言可以被看作是一座老城，错综的小巷和广场，新旧房舍，以及在不同时期增建改建过的房舍。这座老城四周是一个个新城区，街道笔直，房舍整齐划一。”^③对照这种描述，我们可以将人工智能视作智能这座老城的一个新区，任其发展。如此一来，只要我们把人工智能限定在一个独立的新区内，它的发展对老城区便构不成任何威胁。不过，维特根斯坦的锤子比喻还有更深层次的用意，需要我们更仔细地去体味。这样做或许有助于我们进一步理解人工智能与人类智能之间的关系。

我们注意到，维特根斯坦在谈到锤子的常规用法时还做了更细致的区分：一是，把钉子敲进木板；二是，把木塞敲进孔里；三是，把某样东西打碎。尽管这些都可以称为锤子的敲打功能，但它们毕竟是不同的。假如把人工智能和人类智能都和锤子的敲打功能对应起来的话，我们会倾向于说，人类智能接近于前两者，而人工智能更接近于第三种功能，因为锤子的这种功能相对于前两者更不常见。我们还可以再试着把人工智能和人类智能只对应于锤子的“把某种东西敲进某种东西”的功能，那么，是不是可以很确定地说人类智能对应于把钉子敲进木板，人工智能对应于把木塞敲进孔里？事实上，随着考察的深入和细化，我

们会觉得越来越难以做出明确的区分。

维特根斯坦在《哲学研究》第62节中做了类似的评论“同样,我们也会说某些东西是作这个用的,作那个用的。本质之点在于:这是一盏灯,它是用来照明的——拿它来装饰屋子,填充一块空白之类则是非本质的。然而什么是本质的、什么是非本质的,并不总划然有别。”^④这表明,如果对话词的用法做近距离的观察,我们便会发现很多被忽略掉的区别。而在维特根斯坦看来,正是由于未能注意到这些细微的区别,我们才会陷入一个个哲学困惑,难以自拔。

事实上,维特根斯坦的锤子比喻还涉及另一个更深层次的区分,那便是锤子的某一具体功能的两次发挥之间的区分。前文引自《蓝皮书》中的那段话里有这样一句“下面这种说法是正确的‘我们用这个锤子做两件事:我把一颗钉子敲进这块木板,把另一颗钉子敲进那块木板。’”在维特根斯坦看来,只有具体到这个层次,我们才能有把握地做出确定的区分。那么,在这个层次上,我们能否把人类智能对应于第一件事,而把人工智能对应于第二件事呢?恐怕不能。况且,即便可以这样做,我们也禁不住会反问:这种区分和对应还有什么意义呢?因为两者已经没有什么区别,人工智能和人类智能做的是同样的事情!这里已达到标记同一(token identity),而前面几种情况只是类型同一(type identity),比如人类智能对应于锤子的敲进功能,而人工智能对应于锤子的敲碎功能。显然,人工智能和人类智能之间无法达到标记同一。

二

至此,对于人工智能与人类智能之间的关系,我们会倾向于类型同一,而不是标记同一,因为它们毕竟是两类不同的智能。换言之,我们考察“人工智能”和“人类智能”这两个术语的语法时应参照锤子的两种不同功能来进行。这便将我们引入了它们各自的语言游戏。

智能是数百万年生物进化的结果,是人类引以为傲、区别于其他生物的禀赋。关于“智能”和“人工智能”有各种各样的界定和描述,我这里选取《牛津哲学词典》的定义“最一般地说,智能就是灵活而有效地处理实践和理论问题的能力。”^⑤“人工智能是关于制造能做跟人一样的事情的机

器的科学。……强人工智能是这样的哲学论点:适当编程的计算机同我们一样拥有心灵。弱人工智能是这样的方法论信念:坚持认为探究心灵的最佳途径是假定上述论点是对的,而对那些会引起对强论点的不快的二元论遗产不置可否。”^⑥

这两个定义简单明了,便于使用。我们说,智能(一般只在同人工智能并提时才加上“人类”两字)主要是脑生理学家和心理学家研究领域。这些科学家研究由大脑控制的人的感知能力、反省能力、记忆能力、情绪能力、意志能力、思维能力、判断能力、推理能力、语言能力、学习能力、交际能力,等等。他们借助一套专门语言进行学术研究、展开学术讨论、发表学术成果,致力于解答人们为什么得以灵活而有效地处理实践和理论问题。援用维特根斯坦的术语,我们说,这些科学家在进行着自己的语言游戏,而这也正是他们独特的生活方式。当然,科学家的术语也会通过各种途径进入日常生活,成为我们日常话语的一部分,供大众使用。

人工智能是较晚兴起的一个科学和技术研究领域。它的标志性特点是,通过研究制造出能成功模拟人类智能的机器人。这种机器人可以在某些方面表现出类似于、甚至超过人类的智能,或许将来有一天还有可能制造出全面超越人类智能的机器人。这一领域的科学家自然也是在用他们的一套专门术语做着自己的语言游戏。不过,这些术语多是从先前的脑生理学、心理学和工程学等领域借用过来的,只是做了一些改造,或者加了一些限定。比如,他们会使用这样一些术语:智能机器、智能搜索、机器学习、深度学习、知识获取、机器感知、机器思维、人脸识别、人工生命、专家系统,等等。与此同时,这些术语也迅速进入日常话语,并被大众接受。

既然人类智能和人工智能均拥有属于自身的语言游戏,只要愿意遵守游戏规则,人们便可自如地参与到这些语言游戏中去。不过,问题在于,当这两种语言游戏彼此渗透、发生交叉的时候,势必会引起进一步思考:两者到底是什么关系?或许,两个领域内的专家不会太多地受到这个问题的困扰,因为他们只求把本领域的研究尽量做好。但是,同时熟悉这两个研究领域而又具有哲学头脑的人士则禁不住会对上述问题苦思冥想,比如那些坚持强人工智能说的学者。另外,普通民众里也会有一些喜欢思考的人纠结于这样的问题,因

为他们的日常语言游戏里会同时出现本属于这两个专门研究领域的术语，它们之间会相互纠缠，难分轩轻。

三

维特根斯坦对机器能否思考等话题的探究，所针对的正是人们在这一领域内所遇到的哲学困惑。他与“人工智能之父”图灵之间的思想交锋，是人们津津乐道的话题。1939年，图灵参加了维特根斯坦在剑桥开设的关于数学基础的讲座课程。其间，两人就机器能否思考的问题发生了激烈的争论。后来，图灵承认维特根斯坦对他的思想产生了很大影响。图灵此前发表了著名的论文《论可计算数及其在判定问题上的应用》（1936），维特根斯坦读过此文并为之吸引，甚至把它用作课程的阅读材料。两人争论的焦点就是能否制造出可以思考的机器。图灵持乐观态度，而维特根斯坦则坚定地认为，机器无法像人一样思考。

事实上，在图灵出现在他的课堂上之前，甚至在他发表那篇著名的论文之前，维特根斯坦已经对机器是否具有智能的问题做了很多探讨。例如，在写于1933—1934年的《蓝皮书》中，他就探讨了“机器能思维吗”“机器能患牙痛吗”等问题，并明确给出否定性的回答：“有人反对这样一种说法：思维是某种像手的活动那样的东西。人们可能说，思维是我们的‘私人经验’的一部分。它不是物质的，而是私人意识中的事件。这种反对意见也表现在‘机器能思维吗’这样的提问中。我们后面还会谈到这一点，现在只提出一个类似的问题‘机器能患牙痛吗？’你肯定会说‘机器不能患牙痛。’我目前仅仅想请你注意你对‘能够’这个词的用法，并且问你‘是否你对此想说的是，你过去的全部经验表明，任何机器从来都不会患牙痛？’你所谈的那种不可能性是一种逻辑的不可能性。问题是：在思维（或者牙痛）和进行思维、感到牙痛等等的主体之间是一种什么样的关系？我目前对此只谈这么多。”^⑦

他这里所指的“后面的”探讨是“有时我们觉得，仿佛个人经验这种现象是一种处于空气上层之中的现象，它与处于地面上的物质现象是对立的。有这样一种看法，即认为当物质现象达到一定复杂程度时，便形成这种处于上层之中的现象。例如，当一种动物发展到某种复杂程度时，便

出现精神现象、感觉经验、意志，等等。这种看法似乎有某些明显的真实性，因为阿米巴肯定不会说话，不会书写，也不会讨论问题，而我们却能够做到这一切。另一方面，这里出现一个可以用如下提问加以表达的问题‘机器能否思考？’（是否可以借助于物理规律或者也许借助于其他某种与机体行为相关的规律，来描述和预测这样一种机器的活动。）我们在这个提问中表达的困难，并非真正在于我们还不知道能够思考的机器。这个问题不同于一百年前人们可能提出的那个问题：‘机器能否使气体液化？’这里的麻烦毋宁在于：‘一架机器在思考（感知、愿望）’这个句子无论如何看起来都是荒唐的。它仿佛向我们提问‘数字有颜色吗？’（它可能有什么颜色呢？因为它显然不具有我们所知道的任何一种颜色。）”^⑧

在如上两段论述中，维特根斯坦抓住的是这样一个核心问题：像“机器能否思维”和“机器能否有牙痛”之类的问题不能通过观察某台（或某些）机器加以回答，因为这种经验研究所得出的结论无法作为上述哲学问题的解答。他认为，当我们说机器不能思维或不能患牙痛时，要表达的是一种逻辑不可能性，而不是经验不可能性，因此无法通过经验研究加以证实。这就可以解释，维特根斯坦为什么不赞同图灵关于人工智能的设想以及他试图通过测试来判定一台机器是否拥有“智能”的做法。维特根斯坦的正面主张是，这样一些问题的提出都是因为我们错误地使用了语言：当我们问一台机器是否可以感知、思考或表达愿望时，我们是在不适当地使用“感知”“思考”“意愿”等词语，因而，这样的问题是荒唐的，永远没有答案。此时，我们已深陷于哲学的泥潭之中。只有当我们意识到这样的问题毫无意义并停止追问之时，我们才可以摆脱这一困境。当然，如果把这样一些语词的使用严格限定在人工智能的专门领域之内，或许就不会引起哲学上的误解了。

在后来创作的《哲学研究》一书中，维特根斯坦再次明确表达了类似的看法：“机器会思想吗？——它会疼吗？——该把人体叫作这样一台机器吗？它可是极接近于这样一台机器啊！”“但机器当然不会思想！——这是一个经验命题？不是。只有说到人，以及和人相似的东西，我们才说他思想。我们还这样说布娃娃，当然还有精灵。把‘思想’一词当作工具来看看它！”^⑨“我们说到无生命的东西有疼痛：例如在和布娃娃玩的时候。

不过,疼痛概念的这种用法是次级用法。”^⑩

我们说,维特根斯坦在他的哲学著作中对人工智能哲学问题的探讨,完全是治疗性的。他是想提醒人们密切注意那样一些既被用于描述人类智能又被用于描写人工智能的词语在不同的语言游戏中被使用的具体情况,而不应该脱离使用背景,孤立、抽象地询问:机器的思想与人的思想是不是一样?哪一个更优越?机器的思想会不会超越并最终取代人的思想?这就是说,他的探讨是想向人们呈现“智能”一词丰富多彩的用法,引导人们实现对这一术语用法的综观,看清它们之间的联系和区别。回到他的比喻,可以说,他是想通过自己的细致考察,向人们展示锤子的多种多样的功能,并帮助人们认清它的这些功能之间的异同。我们需要防止的是各色各样的“语法虚构”,^⑪正是这样一些语法虚构把我们引向了哲学困惑。

后期维特根斯坦还提出了一个非常重要的概念,那就是“家族相似性”。他认为,传统哲学所探讨的所有概念都应被视为“家族相似性”概念。这其中当然也包括像“心灵”“智能”“思想”“感觉”“情感”“意志”“学习”“阅读”“理解”等这样一些概念。对它们进行哲学考察,不是探讨它们所指代的那些现象(这是各门科学的任务),而是观察其中每一个术语的实际用法“我们不分析现象(例如思想),而分析概念(例如思想的概念),因而就是分析语词的应用。”^⑫

既然“智能”“思想”等都是家族相似性概念,那么人工智能与人类智能、机器思想与人类思想之间的相似性到底表现在哪些方面呢?这绝不是一个可以给出明确答案的简单问题。维特根斯坦的考察比我们想象的要细致和慎重得多。例如,他在《哲学研究》中关于机器思想的探讨就是对比人的思想、狗的思想甚至椅子的思想而展开的。^⑬事实上,维特根斯坦对机器思想与人类思想之间的相似性是不大认同的,他更倾向于主张这两个概念属于不同的语言游戏。在晚年撰写的《心理学哲学评论》中,他这样写道“图灵的‘机器’。这种机器的确是一个进行计算的人。也可以把这个人所说的话用游戏的形式表达出来。这种饶有趣味的游戏可能是这样,即人们按照规则却得出一些没有意义的指令。”^⑭他这里想表达的意思是,假如我们把机器思想与人类思想相等同,势必会导致对相关语词的看似符合规则却是无意

义的使用。因此必须区分这两类不同的语言游戏。人工智能与人类智能之间的相似性远远低于动物智能与人类智能之间的相似性,因为机器人根本就不是一种生命体。或许,我们所说的“机器思想”就像“布娃娃思想”一样,最多不过是“思想”这个词的一种派生性的次级用法。

维特根斯坦为各种哲学病症所开列的药方,都远不是速效的,相反,这些病症都需要长期的医治过程。对于人工智能领域产生的哲学困惑,也只能慢慢疗治。而他之所以坚决反对将那些本来用于描述人类智能的术语用于描述机器智能,主要考虑的是从根本上消解深深困扰着哲学家们的心身关系问题。我们知道,自笛卡尔以来,一代代哲人为解决这一问题投入了巨大的努力,提出了各色各样的心身关系理论。然而,无论是心身二元论还是各执一端的心灵一元论和身体一元论都面临着无法克服的理论困难。其实,将人体视作一架机器的想法早已有之,如笛卡尔、拉美特利等都持此说。赖尔就曾把笛卡尔的理论称作“机器中的幽灵说”。前面引用的维特根斯坦《蓝皮书》的论述中就涉及将人的经验意识看作事物演化之结果的观点。维特根斯坦承认这种观点具有某种明显的真实性。而他紧接着就探讨机器能否思维的问题。给人的印象是,他抓住的是这样一个思路:是否可以借助日益先进的技术极大地缩短这个演化过程,直至可以在很短的时间内人工制造出具有跟人一模一样的经验意识的机器?(眼下,我们会经常听到人说,人工智能技术是以级数方式进步的。)维特根斯坦不会意识不到技术进步的巨大力量,因为他本人就是这一行出身的。但他显然放弃了接续任何一种心身关系理论的努力,因为他清楚地认识到,无论是心身平行论还是心身还原论都无法自圆其说。因此,他转而劝导人们把关注的目光转向哲学术语的日常用法,并提出:那些恼人的哲学问题全部源于语言用法的混乱。他强烈要求“我们把语词从形而上学的用法重新带回到日常用法。”^⑮事实上,维特根斯坦的哲学工作就是帮助人们看清那些被哲学家赋予了“哲学含义”的科学术语的最初用法。这些术语的发明离不开日常语言,并在科学普及的过程中重新获得在日常语言中的固定位置,但“好事者们”总是禁不住诱惑将它们的“哲学含义”船载以入。维特根斯坦就是要扮演那只“断其喉、尽其肉”的猛虎角色,去消灭这头“黔之驴”。

四

如前所述，维特根斯坦是通过将哲学术语视作工具来考察它们的用法进而消解用它们表述出来的哲学问题的。他的“锤子比喻”形象地表达了这层意思。本文一开始便将“智能”一词比作一把具有多项功能的锤子。让我们对维特根斯坦的锤子比喻做这样的引申：锤子把钉子敲进某物里，是使之更坚固；而锤子把东西敲碎，是一种改变，也是一次打开。这或许有助于我们准确理解人工智能，从而消除围绕它所产生的那样一些无益的哲学问题。

我们知道，人类开发人工智能的初衷是更好地帮助人们实现仅凭人类智能难以达到的目标。事实上，电脑在计算速度和准确度、程序化任务的执行能力等方面都已大大超过了人类智能，并切实为人类带来了极大的便利。这种积极的作用类似于锤子的敲进功能，无疑是未来人工智能研究的主要目标。人工智能必须为人类服务——这是大家的共识。

当然，人工智能作为一项新兴的尖端技术，它也肩负着技术革新的使命。它要突破和战胜的是旧有的、落后的技术。我想，目前全球范围内的人工智能技术大战，正是这样一场除旧布新的大竞争，谁也不甘落后。这里必须注意的是，人工智能绝不能，也绝不该与人类智能为敌。然而，可怕的是，人工智能可能会被用于战争和犯罪，甚至会制造出“人工智能武器”，用于对付人类智能，直至实行人类杀戮。锤子会被用来敲碎人的脑袋——这是人类必须携手防范的大灾难！

我们相信，人类有足够的智慧，像控制核技术泛滥一样，控制人工智能技术的滥用，甚至有必要及早联手，把它控制得更好。这里用到的已经是维特根斯坦的锤子的“敲开功能”了。人工智能为人类打开了一扇了解自身和宇宙奥秘的大门。我们知道，对人工智能的研究除了制造机器人之外，还有一个更进一步的目标，那就是通过机器的模拟帮助我们了解人类智能的奥秘。由于大脑构造极其复杂，直接对之做神经生理研究是一项几乎难以完成的工作。而人工智能则开辟了一条了解人脑工作原理的新通道。另外，人工智能的兴盛必将大大增强人类认知宇宙的能力，让我们在茫茫太空中看得更远、飞得更远。甚至有人设想，

未来的人工智能技术有望帮助人类实现太空移民的梦想。

从人类文明进步的视角看，作为一项尖端的科学技术的人工智能完全是无可厚非的。我们甚至可以设想，维特根斯坦也会满怀热忱地关注人工智能的发展，假如他能像罗素一样长寿的话。然而，他一定不会赞同强人工智能说的哲学主张。那么，我们到底该如何看待维特根斯坦对于人工智能的贡献与期许呢？

著名人工智能哲学家奥托·诺伊迈尔（Otto Neumaier）发表过一篇题为《一个关于人工智能的维特根斯坦式观点》的长文，详细探讨了维特根斯坦哲学与他逝世以后才兴起的人工智能之间的关联，并得出了这样一个结论：“维特根斯坦绝没有为人工智能的发展打开大门；所以，不能把他看作人工智能的一位先行者。然而，维特根斯坦却让使人工智能成为一项有意义的事的大门敞开着；确保这一点的条件也已明朗：唯当我们的生活方式发生了根本的改变，人工智能的各项目标才会实现。”^⑩

奥托·诺伊迈尔的这一结论自然有其可取的一面，但也有可以商榷的地方。首先，维特根斯坦以其对数学基础、数理逻辑和工程技术的强烈兴趣和深入研究，对当时已出现发展苗头的人工智能技术表现出极大的关注。他虽然没有像图灵那样着手进行专门的理论和应用研究，但他对这项技术的发展前景保持着清醒而谨慎的看法。他虽没有用到“人工智能”这个字眼，但他对机器感觉、机器意向、机器思维等“智能表现”以及人工肢体、人工疼痛等非自然造物都做过探讨。因此可以说，他实际已开始思考这项技术的实质内涵了。当然，作为一名哲学家，他在这一领域内的职责，正如奥托·诺伊迈尔所说的，乃是保障这一技术成为一项有意义的事。我们把他的锤子比喻用到这一领域，正是想借以呈现他在面对这项影响巨大的新技术时所表现出来的一位哲人的良知。这里，我们可以谈及维特根斯坦之锤的“敲响功能”。它将在人工智能的发展进程中，不断敲响警钟，唤醒人类的理智和良知，使其更好地认识自己、规范自己的行为，更有效地防止人工智能发展过程中的不良倾向，帮助人类奔向美好的未来——人工智能必须成为一项有意义的事！

凭我个人的观察，目前人工智能研究确实出现了一些不大好、至少不值得鼓励的倾向。一种

倾向是技术方面的。如把机器人的外形造得越来越接近于真人,甚至制造出了“性爱机器人”。这一方面是为了拉近人机距离,使机器人更容易被人们接受;可另一方面恐怕只是出于商业考虑。另一种倾向是社会方面的。如急于给某台机器人以居住权、公民身份等,甚至还会很快有人考虑和机器人结婚。在我看来,这两种倾向不只是为了哗众取宠、谋取利益,里面还潜藏着哲学误导:把机器人的外形做得更像人的身体,是想掩盖模拟人类智能进展不力。强人工智能说的倡导者们孜孜以求的、类似人的心灵的东西一直无法在机器中“现身”,所以只有在外形上大做文章了。而授予居住权和公民权的做法更是一种露骨的、对尚不成功的心智模拟的乔装打扮。面对这些不良倾向,我们最好抡起维特根斯坦之锤,为之敲响警钟。从维特根斯坦的视角看,这些做法均不属于正常的人工智能语言游戏,必须加以校正,才有助于人工智能研究沿着正常的轨道前进。

至于奥托·诺伊迈尔说到的实现人工智能健康发展的那个条件,我觉得并不一定能代表维特根斯坦本人的意思。维特根斯坦确实非常明确地说过:想象一种语言游戏,就是想象一种生活方式。但是,并不能由此断言,人工智能语汇的进入,会实质地改变人们的日常语言游戏,从而实质地影响并改变人们的生活方式。当然,不可否认的是,时至今日,人工智能技术的很多成果已广泛渗透进我们的日常生活。甚至可以说,很多人的生活已经离不开人工智能了。比如,我们有这样的感觉:离开智能手机,就像丢了魂似的。但是,人们的生活形式已经或者将要发生根本的改变么?人类历史上的历次技术进步确实改变了人们的生活形式,比如石器制作技术、制陶技术、冶炼技术、种植技术、书写与印刷技术、工业技术、电气技术,等等,但是,人工智能技术是否可以算得上一场伟大的技术革命,从而为人类生活带来巨大变革呢?现在还很难做出判断。

不过,维特根斯坦在《哲学研究》第420节已经做了一个有趣的设想“但难道我无法设想我周围的人——尽管他们的行为方式一如既往——都是机器人,都没有意识?——如果我现在——独自在我的房间里——这样设想,我会看见人们目光凝滞(有点像发呆时那样)干这干那——这想法也许有点吓人。但试试在寻常交往之际,例如走在大街上的时候,坚持这种想法!你对你自己

说‘那边的孩子都只是些机器人;他们活蹦乱跳,却都是自动装置发动的。’要么这些话对你什么都没说,要么会在你心里产生某种吓人的感觉,或诸如此类的感觉。”^⑩他接下来做出了这样的总结“把一个活人看作机器人,就像把一个形象看作边界事例或另一个东西的变体,例如把一个窗格看作乐字符。”^⑪

事实上,随着人工智能技术的发展,机器人会越来越多地进入我们的日常生活,而不需要我们去把活人设想为机器人了。维特根斯坦所设想的是人们如何与假想的机器人相处,而我们很快就会和很多真的机器人共同生活了。但是,不应忘记的是,无论人工智能技术发达到什么地步,我们都完全可以把机器人与活人区分开来,并以适当的方式与它们打交道。很难想象机器人的介入会对人们的生活方式带来根本性的改变,这就像许许多多的自动装置的进入并未对我们的生活造成实质性的影响一样。正如维特根斯坦所总结的那样,我们只消略微转换一下看问题的方式,便可有效地避免无意义的言说,同时克服心中的不适乃至恐惧。如此一来,人类智能和人工智能便可相安无事、友好相处了。

有人说,人工智能或许会成为人类的最后一项大发明,因为这可能会使人类自身成为它的牺牲品。如果真有这么一天,我们只能承认这样一个悖谬性的事实:人类智能被它的产物——人工智能——彻底击败了。这听起来非常类似于这种情况:早期分析哲学家们构造出来的人工语言最终取代了自然语言,成为人们日常交流的工具。但是,这种情况并没有发生。恰恰相反,人工语言的理想早已化作了泡影,而自然语言依然如故。另外,当有人急切地盼望着“奇点降临”的时候,我们又会不由自主地联想起曾让多少人痴迷不已的永动机。岂不知,就像任何一台机器的持续运转都离不开能量消耗一样,任何一个机器人的连续行为都有赖于人类智能的介入。“人工智能”这个术语中的“人工”一词就已经在某种程度上限定了它的使用范围,从而明确引导着我们对之做出语法综观。

因此,我们似乎可以断言,当奥托·诺伊迈尔设想人们必须从根本上转变自身的生活方式以适应人工智能的大发展时,我们所看到的是对这项技术的一种过于理想化的想象和期待。我们自然会热忱希望人工智能技术继续为人类带来福祉,

也会以开放的心态接纳人工智能术语源源不断地进入我们的日常语言,使我们的语言游戏更加丰富,使我们的生活更加多彩、更为便利。我们也完全可以设想,人工智能技术会在越来越多的可程序化的理智领域内大显身手。但是,它在人类情感和意志等方面的模拟将难以实现根本的突破。而只要在这些方面尚未实现突破,即便会有各色各样迷惑人的新鲜事儿粉墨登场,那也将无法从根本上改变人们的生活方式。

人类会时时防备着这项技术可能带来的种种负面影响。真正值得信赖的,最终还是,也只能是我们人类自身的智慧。在享受人工智能带来的福利的同时,“维特根斯坦之锤”会不时为我们敲响警钟……

注释:

①涂纪亮主编《维特根斯坦全集》第6卷,涂纪亮译,河北教育出版社2001年版,第77—78页。

②维特根斯坦《哲学研究》,陈嘉映译,上海人民出版社2001年版,第10页。

③同上书,第13页。

④同上书,第46—47页。

⑤《牛津哲学词典》(重印本),上海外语教育出版社2000年版,第195页。

⑥同上书,第26页。

⑦涂纪亮主编《维特根斯坦全集》第6卷,第23页。

⑧同上书,第63—64页。

⑨维特根斯坦《哲学研究》,第174页。

⑩同上书,第148页。

⑪同上书,第157页。

⑫同上书,第81页。

⑬同上书,第173—175页。

⑭涂纪亮主编《维特根斯坦全集》第9卷,第270页。

⑮同上书,第73页。

⑯Stuart Shanker and David Kilfoyle, eds. Ludwig Wittgenstein, *Critical Assessments of Leading Philosophers, Second Series, vol. III*, Routledge, 2002, pp.361—362.

⑰维特根斯坦《哲学研究》,第194页。

⑱同上。

Artificial Intelligence and Human Intelligence: Two Concepts and Games of Two Languages

LI Guoshan

(College of Philosophy, Nankai University, Tianjin 300350, China)

Abstract: The rapid development of AI has given rise to a series of philosophical problems, especially that of the relation of AI and human intelligence. Having not witnessed the great advances of AI, Wittgenstein, nevertheless, wrote much about the sensation and thinking of machine, or even the feeling and will of machine, and expressed his forward-looking views about AI technology. He made it clear that the relevant concepts of AI could be used properly only within its own domain by making close grammatical investigations of relevant concepts. Once they were taken out of its field and confused with the concepts of human intelligence, many confusions in philosophy arose. It is such confusions that Wittgenstein aimed to remove. His hammer metaphor can help us to understand his view of the relation between AI and human intelligence and its enlightening significance.

Key words: AI, human intelligence, hammer metaphor, grammatical investigation, Wittgenstein

(责任编辑: 知 鱼)