



孔澜颖 摄

刘晓力,中国人民大学哲学学院教授、现代逻辑与科学哲学研究所所长、博士生导师。研究方向:科学哲学、逻辑学、心灵哲学与认知科学哲学。目前正在主持国家社科基金重大招标项目“认知科学对哲学的挑战”。

## 刘晓力:人类智慧真的要被机器替代了吗

陈也喆

说起人工智能,很多人可能觉得这个主题与生活离得很远,其实我们当下的生活里,处处都是人工智能。

会端盘子的机器人,会下棋的“阿尔法狗”,VR(虚拟现实技术)电影……那些在科幻电影里出现的画面,正一步步真实地走入我们的生活。

为什么自然界会进化出像人一样具有意识的动物?人工智能真的能取代人类吗?为什么斯蒂芬·霍金如此“害怕”人工智能?

4月8日,刘晓力做客天一讲堂,主讲《人的自我和机器自我》,给我们分享了一些高科技的话题。

### 人的意识是怎样逐步形成的

这是刘晓力第一次来宁波,她非常喜欢这个城市,山清水秀,人杰地灵,历史与文化底蕴深厚,气候宜人,非常适合居住。

刘晓力是学数学出身,后来学了逻辑,然后去做了科学哲学的研究。“人工智能的发展极端迅速,我不一定能跟上它的前沿,只是做一番哲学的思考。”她谦逊地说。

刘晓力一直在思考这样一些问题:人的意识和自我认知究竟是怎样进化的?为什么自然界会进化出像人一样具有意识的动物?

机器是人造的,人跟其他动物最大的区别,就是可以使用抽象符号,能

制造机器为我们所用。可是自然界为什么可以进化出人的物种,也可以进化出其他低等动物,就是没有进化出机器呢?

那么,有没有可能机器有自己的进化方式,造出来的东西,以后就不听人类的指挥了?这个问题已经有人开始思考了。

我们知道,虽然很多小动物都有情感,但没有人那么复杂,也不像人类一样有自我意识。

其实,人类的自我认知也是一个进化的产物。人类最早也只是为了生存,最初只是找找虫子,找找吃食,为了便利,开始发明工具、运用工具,慢

慢地认识世界、改造世界,一步步进化到我们现在这样的智能。

从“认知科学哲学”范畴的角度提出“自我”这个概念的,是哈佛大学的心理学家威廉·詹姆斯。威廉·詹姆斯觉得“自我”包含三个方面:第一个方面,是基因遗传,是一个身体的自我,我这个身体的DNA跟别人是不一样的,是独立的个体;第二个方面,是文化遗传,我们在社会中跟人交流,能够分清你我他,是一种人际关系中的自我;第三个方面,是心理的自我,每个人的心理世界是很丰富的,小时候怎么样,长大了有什么爱恨情仇,有什么思想体验,这是一种心理积淀。

### 手机成了人类心灵的延展

刘晓力说,最近几年的人工智能的发展有革命性的趋势,人们很担心机器会不会真的跟人一样,可以有意识地跟人自然交流。甚至有一天,机器人自我觉醒,操纵人类,毁灭人类。这在现在看来并不是危言耸听,也不是天方夜谭。

我们现在的医疗科技也很先进。一个人耳背,可以给他一个人工耳蜗,改善他的听力;一个人的视力有问题,

可以植入芯片;心脏不好,可以安装一个起搏机;大脑某个部分受损,可以安装一个电子设备。

我们可以一层一层地武装,直到全身都变成机械。但是,这样全副武装的人,还是他自己吗?

刘晓力说,未来的世界,可能我们都是一个sample(标本)。其实我们现在就有被机器操控的趋势。

举个最简单的例子,我们现在的

生活都离不开手机。如果一天没带手机,感觉就像丢了魂一样。事实上,手机已经成了人类心灵的延展。

试想一下,让一个年轻人断网断电,放下手机,会是怎样的情景?是不是他自己认知世界的能力,以及与外界交流的能力都发生了巨大的变化?这些都是当下值得我们思考的问题。

### 斯蒂芬·霍金“害怕”人工智能

刘晓力说,著名物理学家斯蒂芬·霍金十分“害怕”人工智能,他曾发表过一些公开的演讲,告诉大家不要去探索外星人,如果外星人有一天真的来了,会统治全人类。

霍金如此恐慌,是因为他几乎是最靠近人工智能的一个人。他患有肌肉萎缩性侧索硬化症(卢伽雷氏症),全身瘫痪,不能言语,手部只有三根手指可以活动。

他的身上安装了很多有助于身体功能增强的技术设备,他靠眨眼睛和动嘴唇来控制电脑屏幕。这个屏幕上有很多特制的键盘,他用眼睛盯住键盘上某一个键,这个键就会跳出来一

条命令,或者是一个小句子,或者是一个短语。他还可以用声音控制前面的键盘,然后把他的想法表达出来。他的助手会通过他表达的内容,整理出一篇论文。

斯蒂芬·霍金的担心不是多余的。事实上现在的人工智能日新月异,已经非常先进,只要再攻破一些难点,超越人类已是指日可待。

人类有一个非常强的能力,就是“同情共感”。我们对别人的痛苦,会怀有同情和慈悲,机器有吗?目前还没有,但是机器有一种镜像功能。

十多年前,科学家发现,人类的大脑神经元,有一种功能叫镜像神经元

功能。举个简单的例子,大家可能都看过一个视频,一个3岁的小男孩,常跟着奶奶去跳广场舞,结果时间一长,跳得比大妈们都好。大家会说这个小男孩的学习和模仿能力很强,其实这种模仿能力就是大脑的镜像神经元所控制的。所以只要能够从具象神经元的原理去建构,机器人也会有这样一个人际关系中的自我,它有模仿能力和学习能力,就有可能更像人的自我。

刘晓力说,2015年,伦斯勒理工学院声称制造了有自我认知能力的机器人,他们的机器人其实就是能在镜像实验中辨识自己。

### 刘晓力推荐书目

#### 《情感机器:日常思维、人工智能和人类的未来》

作者:(美)马文·明斯基

这是一部讨论人类自我思维、情感与人工智能的著作,作者被誉为人工智能之父,是人工智能领域首位图灵奖获得者。大脑如何产生思维?情感、情绪、信念、欲望这些又是如何形成的?如果机器有一天能产生自我意识和情感,那是不是意味着真正人工智能的诞生?在书中,明斯基讨论了人类思维的本质,并论证了自我意识、情感和欲望只是人类的思维方式,而人工智能技术或许在未来可以创造出真正的情感机器。

#### 《神经漫游者》

作者:(美)威廉·吉布森

这是第一本同时获得“雨果奖”、“星云奖”与“菲利普·狄克奖”三大科幻小说大奖的著作。本书于1984年完成,最大的成就是预示了20世纪以后的电脑网络世界,并成为很多电影拍摄主题的思想来源。在书中,影响最为深远的三个概念是赛博空间、赛博人和赛博洛克。作者在书中描画了一个离奇、疯狂、似是而非的未来世界。这些意象的批判性隐喻,意味着未来世界错综复杂,耐人寻味。这部小说在创作上的突破,开阔了科幻审视社会的视野,给科幻小说带来了新的活力。

#### 《你的第一本哲学书》

作者:(美)托马斯·内格尔

作者用鲜活的语言,严密的论证分析,以及深刻的洞察力向读者展现了什么才是真正的哲学思考、哲学的真正趣味在哪里。书中的话题很深刻,同时也很亲切,就好像发生在你身边一样,比如我们怎么知道外部世界的存在,怎么去判断一个行为的对错,怎么面对死亡的恐惧,又如何寻找到生活的意义。这本书的目的不是要给出答案,而是以一种非常平易的方式把人们引向哲学问题。

天一讲堂  
宁波市图书馆

投稿E-mail:ljz@cnnb.com.cn