

“可能情感”的可能性 与“数字人文”研究的新动向

谭光辉

[提要]可能情感主要有四种:未命名情感、未体验情感、他物种情感、人造情感。人造情感技术首先要解决情感的数字化问题,这是人工智能工程的难点,难的原因在于电脑无法像人类意识那样追求意义。人工智能是科学,科学是按因果律的思维方式建构起来的,追求确定性,人类意识的时间性本质决定了它按目的论的方式存在,追求可能性,且不能吻合现代科学的证伪原则,因此人类情感不可能被完全数字化。但是人类情感和智能可能被部分数字化,人工情感和人工智能可能在部分方面弥补或超越人类情感或智能。数字技术和人文学科联合才能互动发展,因果论和目的论两种思维方式一道被重视才能为人工智能的发展做出更大贡献。

[关键词]人造情感;人工智能;数字人文;目的论;因果论;情感研究

中图分类号: I206.7

文献标识码: A

文章编号: 1004—3926(2018)09—0167—07

基金项目: 国家社会科学基金西部项目“文学作品情感表现与接受的符号现象学研究”(16XZW007)阶段性成果。

作者简介: 谭光辉,四川师范大学文学院教授,文学博士,研究方向:中国现当代文学、符号学、叙述学。四川成都 610068

一、什么是可能情感?

在现代汉语中,“可能情感”一词有多个意思。第一个意思是某事可能引发某人的某种情感,比如记者在做新闻报道的时候就会设想该故事可能引发读者什么样的情感反应。第二个意思是指包含在同一个叙述或语言中的多种情感可能,例如可以通过情感识别软件,“计算情感转换概率,从而得到嵌入对话中的任何一种可能的情感”^{[1] (P.59)}。第三个意思是用来指某种“非现实指称叙述”引发的情感。很多人分不清虚构世界和可能世界的界限,统统将之称为可能世界,认为:“作者时时处处暗示着你,有意无意地指点着你,通过具体的情境感染着你,使你不由自主地按照作者的价值观念来分析评断这可能世界的一切,从而按照作者的喜怒哀乐而喜怒哀乐”^[2]。这里所谓的可能世界的情感,实际上就被用来指由作者通过虚构故事引发的读者的情感反应,但是这些情感反应并没有超出人类的常规情感范围。本文将要讨论的可能情感,不同于这几个意思。本文用此术语指称那些人类可能感知到但还没有被知晓的情感类型。这类情感或者可以通过一种全新的叙述引发,或者可以通过人工情感合成的方式获得。阿来曾经说过一句意味深长的话“在我

的小说中,只有不可能的情感,而没有不可能的事情。”^{[3] (P.235)}他这里所说的“不可能的情感”,大约就是指那种由一种新鲜的、不常见的故事引发的,普通人未曾体验过的情感。对于没有体验过又可能通过某种渠道体验的人的潜在情感,就是本文所说的“可能情感”。可能情感不是“不可解情感”,可能情感必然在某个时期或某种条件下“可能”被人类理解。

常识告诉我们,人类的情感是异常丰富的,然而被命名的情感数量却是有限的。那些未被命名而又被人类体验过的情感,是第一类可能情感。人们常常习惯性地 will 将体验到的情感划分入某种已经被命名过的、近似的情感类型之中,这容易造成接受者在理解这种情感的时候忽视这些情感的内部差异性。例如喜,就有很多种不同的喜,但是用来描述喜的词语却很有限,人们不得不借助更多的形容词来限定,但后来发现这些形容词也不能够描述得准确,就不得不借助不同的故事将其叙述出来,然而在称呼这些不同的喜的时候,一概称之为“喜”、“高兴”、“快乐”等。

第二类可能情感,是对于个体来说尚未体验过的情感。例如,小孩子开始是不懂幽默的,随着年龄的增长,他逐渐懂得了幽默。幽默产生的情感体验,对于尚未体验过该情感的小孩子而言就

是可能情感。对每个个体而言,没有体验过的情感是很多的。人类之所以愿意不断地阅读新的故事和文学作品,就是因为需要体验一些新的情感,“就是要寻找心灵的慰藉,并接受高尚情感的洗礼”^{[4] (P.127)}。这就说明,人的某些情感是可以经历从无到有的过程的。现在没有体验过的情感,未见得未来就不可能体验到。对任何人而言,都存在一些自己尚未体验过而别人体验过的情感,甚至有些情感是所有人类都还没有体验过的,但随着人类情感认知能力的提升,未来可能成为一种常见情感。这种可能情感还包括一些尚未被激活的情感,比如多数遵纪守法的人就没有体验过吸食毒品或服用其他神经兴奋类药物之后唤起的那种情感。这既不被法律允许,也很少有人愿意去尝试。又如革命先烈们饱受酷刑时体验的情感,我们现在就只能通过文学作品的文字描述与想象。想象出来的情感,对我们而言也是一种“可能情感”。

第三种可能情感,是他物种情感。对人类而言,我们大概永远不可能体验到变成一只猴子或一只狗时将会获得的情感体验,但人类却喜欢在文学作品中去描述这种可能的情感,比如卡夫卡的《变形记》、帕穆克的《我的名字叫红》、莫言的《生死疲劳》等等。可惜的是,人类想象的他物种情感,其实还是人的情感,作家只是借用了—个动物的名字表达人的情感而已。真正体验他物种情感,绝不是简单地凭借文学想象就可以完成的。大卫·铃木(David Suzuki)和阿曼达·麦康纳(Amanda McConnell)说:

从演化过程来看,人类可能是在漫长的时间里,发展出一种深植于基因、渴望其他物种相伴的需求。威尔森创造了一个新词“亲生命性”(biophilia)来形容这种需求,意指“专注于生命与生命互动进程的先天倾向性”,会产生“人与他物种的情感连结,这种多样化的情感反应交织为象征,而成为文化的一部分。”^{[5] (P.180)}

然而即使人类产生了“与他物种的情感连接”的渴望,也未见得就能够设身处地地体验到他物种的情感。这恰如康德哲学告诉我们的道理一样,人只能感知人所能感知到的“表象”,当然人也只能体验到人类的心灵所能体验到的情感,他物种的情感状态,是“物自体”,是不可知的,人类除了想象之外,别无他途。然而人类从动物的各种表现中直观地感觉到,动物是有情感的,甚至有不

少科学家专心地研究动物情感,例如查尔斯·达尔文(Charles Darwin)的《人与动物的情感》(四川人民出版社1999)、马克·贝科夫(Bekoff, M.)的《动物的情感世界》(科学出版社2008)等等。这样,我们就陷入一个很难理解的悖论,一方面,科学研究结论和直观感受使我们不得不承认动物也是有情感的,另一方面,我们又不得不承认康德的说法有道理,我们不可能真正体验到动物或其他生物的情感。这样,他物种情感对人而言就只能是一种“可能情感”。

第四种可能情感,叫做“人工情感”或“人造情感”,这是本文讨论的主要情感。人工情感的主要目标是“模拟、延伸和扩展人的‘自然情感’”,其扩展意义包括四个方面的内容:(1)技术性地帮助理解他人的情感”;(2)技术性地辅助人的情感表达”;(3)技术性地创造情感——人造情感”;(4)技术性地克服不良情感”。^{[6] (P.39-41)}就现代技术正在努力的方向而言,第(1)(2)(4)方面都要依靠对人的常规情感的理解,唯有第(3)方面是一种情感创造和情感生成。科学家们正在努力研究如何在计算机或机器人上模拟或生成情感模式,并进一步借鉴机器人系统中的情感生成模式,寻找技术手段在人脑中建构能够产生我们所需要的情感状态的生理或物理化学状态,“我们需要什么,人工情感技术就可以提供并在我们的身上产生出这种情感”。^{[6] (P.40-41)}当然,如果计算机程序出错,就可能产生出更多种类型的我们意想不到的情感,这些情感都可以叫做“可能情感”。科学家们认为,人工情感是在人工智能理论框架下的一个质的进步。虽然笔者对人造情感是否可能暂时持保留态度,但是从理论上说,通过物理化学的手段让人获得某种虚假的情感体验也并非完全不可能。

人造情感的基本问题,是把人文问题变成科学问题,把精神问题变成身体问题,这是典型的科学主义的思维方式。照此思路发展下去,必然带来一些新的问题:如果一切精神问题最终都可以通过技术手段变成身体问题的话,那么人文学科还有存在的必要吗?人类是否已有足够的心理准备接受这一难以接受的可能现实?

二、情感的数字化问题

人工情感,也就是情感的数字化问题。21世纪以来的科学主题之一就是人工智能,而人工智能的难点和瓶颈正是情感的数字化,这一进程似

乎变得越来越趋向可能。人工智能飞速发展在给人类的工作和生活带来便利的同时,也引发了人们深深的忧虑。人们担心,当人的一切都被数字化之后,人类的存在还有什么意义?

肖峰认为,人工情感可能引发的人文问题主要有四个方面:第一是情感的真实性问题,即人工情感是不是真正意义上的情感的问题;第二是情感的属人性问题,即人工情感能否被“人化”,或者说人造的情感是否属于人的一部分的问题;第三是情感的个体性与丰富性问题,即人工情感能否保证人类情感无限多样的个体差异的问题;第四是情感的复杂性问题,即能否确定计算机在获得某种情感的同时也获得了引发该情感的内容和原因的问题。^{[7] (P. 253-255)} 这几个问题确实值得思考,但是这几个问题都有一个共同的假设前提:人工情感与人的自然情感的产生机制是不一样的。由于产生机制不一样,所以人工情感总是一个外在于人的东西,当它变成人的一部分的时候,总是不能与人形成一个和谐的整体。其中,第一、二个问题,有点类似“特修斯之船”的哲学悖论,这些问题同样可以用于器官移植、器官修复等医学伦理领域。第三、四个问题是技术问题,是在质疑人工情感技术是否有能力做到与人的自然情感产生原理相同,是否能够让人工情感足够丰富、复杂如人类的自然情感。总的来说,这几个问题讨论的,是关于人造情感与人的自然情感的差异、相融程度等问题,不包括人造情感与自然情感的本质差异到底在哪里、情感的数字化到底是否可能等问题。

仇德辉认为,虽然人工情感常被当作永远无法实现的空想,但是世界上没有不能认识的事物,“既然我们能够实现一般思维的数字化,同样可以实现情感的数字化,既然能够实施人工智能,就必然能够实施人工情感,既然能够制作智能机器人,就必然能够制作情感机器人”^{[8] (P. 218)}。然而,实际操作过程中的情况并不如此令人乐观。胡伯特·杜瑞夫斯(Hubert L. Dreyfus)长期研究人工智能的进展,在20世纪90年代得出了这样一个令人失望的结论:至少暂时,基于这样一个假设的研究计划已经走进了死胡同——人类运用事实和规则产生智能——而且没有任何理由认为这样的研究会成功。^{[9] (P. iv)} 他的总结源于20世纪60年代以来人类发展人工智能的多种努力而终归失败的结果。科学家们在上世纪末就已经发现,人类心智活动并非完全按照因果逻辑推理的方式进行,而在这之前

有关人工智能的科学技术手段全都还要依赖这种基本思维方式。本文认同这个说法,人之所以为人的一个根本原因,在于人对意义不停地追寻,而人工智能的技术手段至今无法解决这一难题。

熊培云曾经说过一句很有道理的话“人是意义动物,我们不是为了追求真理而来到这个世界的,但没有谁不追求意义。”^{[10] (P. 402)} 卿文光在总结人类早期思想中的“万物皆流”思想时问道:为什么希腊、中国和印度最早自觉到万物皆流的先贤皆是厌世遁世?答案就是“人是意义的动物,人必然会藉着精神而去赋予他所感知到的东西以及一切他的生活赖以维系的东西以某种意义。”^{[11] (P. 126)} 自卡西尔把人定义为“符号的动物”引发广泛思考之后,符号学家们进一步认识到人是意义的动物。格雷马斯认为,“对人而言,人类世界从本质上来说大概就是意义的世界。”^{[12] (P. 1)} 这是因为,没有意义,符号也就不成立,意义与符号互为存在的原因。赵毅衡认为“符号学即意义学”^[13]。李思屈也认为“当我们说‘人是符号的动物’时,我们实际是在说‘人是意义的动物’。”^[14] 从索绪尔语言学的角度看,符号与意义是一体两面,所以符号和意义都不可能离开对方而单独存在。不论从哪个角度看,我们都可以得出一个共同的结论:人之所以为人,是因为人生而具有追求意义的本能,“人赋予世界以意义”,“人在世界中发现意义”,“人是为意义而活着的”。^{[15] (P. 27)}

我们不禁会问,什么是意义?对这个问题,符号学家常常陷入尴尬,只能进入“符号”与“意义”的循环解释,赵毅衡给“符号”下的定义,按他自己所说,“事实上也的确是一个解释循环”。^{[16] (P. 3)} 那么,如何不进入循环解释而给人所追求的“意义”一个更为本源性的解释呢?这可能会涉及人类智能(特别是情感智能)与人工智能的一个核心区别。这个核心区别将会告诉我们,人类的情感是不可能被完全数字化的,人类的很多其他智能也是不可能被完全数字化的。数字人文,必须明白自己的限度,人类也完全没有必要对数字人文的高速发展过分担忧,至少会像周昌乐所说的那样:“机器的情感最终是不会超过人类本身的,因为人类总可以发展安全装置来阻止此类事情的发生。”^{[17] (P. 115)}

三、意义的意义:人类情感为什么不能被完全数字化

意义是什么呢?这个问题回答起来很绕,名

家们各执一词。英国学者奥格登(C. K. Ogden)和瑞恰慈(I. A. Richards, 一译理查兹)在《意义之意义》一书中提出了 16 条实际上 22 种意义的定义。^[18] 英国语言学家杰弗里·利奇(Geoffrey Leech)在《语义学》中给“意义”分成七种不同类型: 理性意义、内涵意义、社会意义、情感意义、反映意义、搭配意义和主题意义。^[19] (P. 13-33) 前者主要梳理了“意义”一词的种种不同用法, 后者主要指出“意义”主要被用在什么场合, 实际上并未从本质上给出意义到底是什么的答案。格雷马斯(A. J. Greimas)说“意义, 作为意义之形式, 可以被定义为意义变化的可能性”, “在哲学家的语言里, 它还指一个意向或一个目标”, 他还指出“在最近这段时间里, 符号学的进步主要就是扩展了它的演练场, 进一步探索了获取意义的各种可能策略。对意义的本质我们依然一无所知, 但我们却更深入地了解到它在何处显现, 如何演变。”^[20] (P. 12-14) 即使经过数代哲学家和符号学家的努力, 我们最终也无法给“意义”一个本质性的说明, 这既说明了本质主义立场的破产, 又说明了“意义”这个词本身的飘忽不定。赵毅衡在仔细考察了奥格登和瑞恰慈的 22 条定义之后得出结论说“只有意义中包含的主客观互相建构、把主体变成意义世界里的存在者, 才是意义的‘本质’。”^[18] 当我们不得不把“主体”、“主观”等因素考虑进对“意义”的理解中去的时候, “意义”就不得不成为一个不具有确定性的概念。

我们会经常常识性地认为, 数学追求确定性, 而且是所有学科中最具确定性的, 数学需要逻辑严格的证明。然而深究下去, 恐怕并非完全如此。康德的数学哲学告诉我们, 数学的最初依据仍然是直观和假定, 具有主观性的特征。包向飞在研究康德的数学哲学之后说“作为数学重要组成部分的‘理想化的规定’总是与‘注意力的方向’联系在一起, 而‘注意力的方向’却总是受到历史和文化的牵引, 数学中的‘理想化规定’从来就不是一次性完成的。”^[21] (P. 192) 归根结底, 数学看似追求确定性, 但其作为出发点的基础并不一定可以被严格证明有确定性的。比如, 等式 $1 + 1 = 2$ 是代数的基础, 其他运算最终都是从这里出发推算出来的。但是如果我问为什么 $1 + 1 = 2$, 怎么证明? 数学家就没有办法了。又比如, 整个欧几里德几何学, 都是由欧几里德的 5 个公设和 5 个公理推导出来的, 而这 5 个公设本身, 是无法被证

明的。而对第 5 条公设的置疑, 最终导致了罗氏几何和黎曼几何的诞生。任何数学定理的推导, 都要给出一个“理想化的规定”条件, 这些都说明数学的确定性本身是有限度的。这正如威廉·韩思(William Haines)所说, 那些可以举一反三而又不容置疑的前提通常很罕见, 这些前提可能来自常识, 或生活经验和观察到的普通事实, 或来自其他学科及科学领域的知识, “有时我们不过是相信了一些对我们而言似乎是对的东西, 甚至相信那些明明证明是无根据的东西。”^[22] (P. 15) 数学与其他学科在本质上最终是类似的, 思维出发的基点, 是意识努力的方向。用并非绝对确定的数字方法, 去描述一个不具确定性的“意义生物”的工作方式, 当然就不可能得出确定性的结论。但是, 就努力方向而言, 数学追求确定性, 而意义无法追求确定性。这个结论可以用一个简单的例子予以说明: 在高考的试卷中, 以数学为代表的理科命题, 大多是给出条件(因)让学生求出唯一答案(果), 或者给出条件(因)也给出答案(果)让考生回答推导过程(证明); 但是以语文为代表的文科命题不可能这样, 只能给出一个现象(果)让考生去回答为什么会这样(因), 或者给出一个表述(符号)让考生去推测作者是什么意思(意义)。这个道理很简单, 如果数学命题像语文那样从果追因, 不断追下去就只能得到那几个强制规定的公设, 如果语文像数学命题那样从因追果, 追下去就永远没有办法使答案唯一。

意义为什么无法追求确定性? 或许我们换个思路来思考会让答案更清晰。在思考“意义”的意义的时候, 人们习惯性地意义看作一种相对于符号而言的共时性存在, 而这可能是问题的根本症结所在。共时性的两个存在, 在常识中就被理解为确定的, 比如我们不会置疑眼之所见的世界本身的组合关系的确定性。索绪尔最重要的贡献之一, 就是把语言学从历时性研究变成了共时性研究。这就造成一个根深蒂固的理解, 总是认为符号与意义之间存在一种共时性的、固定的关系。然而我们知道, 人的存在并非共时性的存在, 而是历时性的存在, 人存在于时间之中, 这也是海德格尔哲学的基本观念。奥伊肯认为, “人存在于时间之中并经受着不断的考验和变化”, “有一种新的、独特的历史在人的经验中展现, 使人类不同于纯粹的自然之物。”^[23] (P. 102) 周宪的说法来自海德格尔“说历史是人的本质规定, 也就意味

着时间乃是人的本质规定。”^{[24] (P.252)} 从康德的理念来看,时间是人的先验知识,所以人认识自己只能在时间之中展开。

当我们在讨论“符号”和“意义”的关系时,我们却忽略了二者之间的时间关系,而忽略空间关系可能还发生在时间被忽略之前。“符号”和“意义”既无空间关系,又无时间关系,那么人的意识是如何将二者联系起来的?仔细思考我们会发现,意识是自动将二者纳入时间关系中,然后抹去了时间。同时,意识也抹去了符号和意义之间的空间关系。或者说,符号与意义,我们既无法想象二者之间存在时间关系,也无法想象二者之间存在空间关系。既无时间差异,又无空间差异的两个东西,为什么意识不将它们看作同一个东西呢?这显然违背了欧几里德在《几何原本》中提出的“彼此能重合的物体是全等的”这一公设。可能的答案是:意识本身的性质决定了它将二者处理为不同。那么意识本身到底是时间性的还是空间性的呢?本文给出的答案是:由于我们没有办法想象出意识的空间性,所以意识只能是时间性的。

艾克哈特·托尔(Eckhart Tolle)曾说“意识本身是无时间性的,因此不会进化,它从未诞生,也不会灭亡。”^{[25] (P.253)} 这个说法显然带有太强的宗教色彩,不可取信。胡塞尔(Husserl)在建构他的现象学理论时,非常清晰地阐明了意识的时间本质“如果我们撇开一切超越不谈,那么,对于感知,就其一切现象学的构造项而言保留了它的现象学的时间性,此时间性属于它的不可扬弃的本质。”^{[26] (P.255)} 海德格尔(Heidegger)进一步发挥:“胡塞尔还称之为时间意识的东西,即关于时间的意识,恰恰是原始意义上的时间本身。”^{[27] (P.285)} 海德格尔大约是综合了康德哲学和胡塞尔现象学的基本观念,得出的结论也非常令人深思。关于时间的意识就是时间本身,所以意识就只能是时间性的。进而,被意识处理的一切也就只能是时间性的。由此可知,在意识中进行辨析的符号与意义的关系,也就只可能是时间性的关系——或者先有符号,或者先有意义。存在于意识中的意义,也只能是时间性的。赵毅衡对此观念的表述非常清晰“意义世界却是人类的意识与事物交汇而开拓出来的,反过来,意义也造就了意识,意识是人类存在的根本原因和根本方式。”^[28] 因此,意识和意义,互为存在前提,且都是时间性的。

此观念也可以置换为关于人的存在意义的问

题,或者认为“人的存在”在先,或者认为“人存在的意义”在先。用通俗的话来讲,前者思考“人活着的目的是什么”,后者思考“人活着的原因是什么”。前者是目的论的思考方式,后者是因果论的思考方式。因果论的追求是唯一的、确定的、必然世界的,目的论的追求是发散的、不确定的、可能世界的。我们非常坚定地相信,人有自由意志,人不是按因果论的方式存在,而是按目的论的方式存在。目的永远是不确定的,所以人只有在完成了存在的过程之后,才能确定其本质,存在先于本质。

现在我们用此观念来观察情感的数字化问题,就会得出比较清晰的答案。数学以及以数学为基础建立的自然科学的追求是确定性,从公设出发的整个数学的大厦完全按照因果论的逻辑推导建设而成,那么以此为基础制造的人工智能、人工情感也必然努力地追求确定性、唯一性、必然性。以时间性为本质而存在的意识按目的论的方式存在于时间之中,其本质与数学的追求相反。原因显而易见:人工智能是科学,人的情感是人文。现代科学的基本原则之一是波普尔提出的证伪原则,要点有二:一是理论必须是严格的普遍陈述,二是理论要有一个或几个可被证伪的关键点。证伪原则认为科学就是大胆猜测,然后通过证伪不断剔除不合理或虚假的部分。然而人的行为和情感理论,能否成为严格普遍的陈述尚不好说,关键是无法证伪。你永远无法准确预测一个人下一刻将要做什么、怎么想,我们没有足够的信心提出大胆的预测。二是情感理论没有办法提供证伪的途径,情感不可重复实验,即使重复实验也不能说明什么,没有意义。沈劲认为“人的情感世界可以说是一个神秘的‘黑箱’,它的许多方面是既无法证实也无法证伪的”^{[29] (P.147)}。爱德华兹则用了半本书的篇幅论证,关于宗教情感有十二种现象是既无法证实也无法证伪的。^{[30] (P.29-84)} 连情感的合理性也是无法证伪的,因为“不管情感自身是否合理,但是它的善良动机及其体现的道德性都能够为合理性提供辩护”^{[31] (P.223)}。情感及人存在于时间之中,而时间无法逆转,所以无法提供证伪的途径。人、情感乃至人文,从本质上说就不能吻合现代科学思想的基本原则。

但这并不能说明有关符号和意义、人和情感的问题就不能被部分数字化。在意识之中,一切时间中的关系,向前看是目的,回头看就是因果。

意识总结因果关系的时候总是循着目的论的路径回去才会显得更合理,所以对人和情感问题的原因解释,永远都只能说哪一个更合理,而不能说哪一个才正确。对世界的理解,用目的论思考的时候,就被理解为可能;用因果关系思考的时候,就被理解为必然。按因果论的方式去理解人的世界是很可疑的,按目的论的方式去理解人的世界可能更趋合理。未来可能的世界,随着时间推移到此刻,像量子被观察时一样,坍缩为现实世界。此刻我回看过往,便会将现在的世界理解为果,将过去的世界理解为因。时间不断向前,永劫不复。但是科学努力按对过去世界的因果理解,部分模拟过去的世界。虽然永远无法追上时间的脚步,永远无法像意识那样永无止息地追求目的,但是至少可以模仿意识曾经留下过的足迹。

四、人造可能情感的可能性及科学与人文的互动发展

虽然如上文所论,人工智能和情感模拟不可能完全模拟或取代人的意识,但是人工智能却可能在技术上做到将人的某一方面的能力无限放大,特别是对人的理性思考能力的放大,比如数字计算、逻辑推理、记忆能力等,电脑早已超过人脑。现在智能计算机发展的难点和热点是机器情感,科学家们认为:只有那些“既有理智又有情感的系统才能成为名副其实的智能机器”,人工智能创始人之一 Minsky 教授指出“问题不在于智能机器能否有任何情感,而在于机器实现智能时不能没有情感”。^{[32] (P. 236)}情感不但决定了逻辑推理内容选择的倾向性,而且决定了是否选择逻辑进行推理的倾向性。只有让计算机具有主动选择的能力,人工智能才是真正可能的。

对机器情感的研究目前面临巨大的困难。科学家们认为,如何认知人类感性以及如何使计算机产生感性被看作是研究智力产生的突破点。虽然情感计算机在一些领域取得了一些突破,但如何自动检测出人(比如进行会话中的人)的感性仍存在着极大的困难。^{[33] (P. 284)}之所以如此困难,原因之一恐怕是人对自我是如何感知的都不甚清楚,而这需要通过人艰苦的自我反思来完成,这就给人文学科留下了艰巨任务。机器情感在检测、识别感性方面尚且难于突破,模拟和创生感性当然也就暂时无从谈起。机器情感的研究分为两种,外在情感研究主要解决人机交互的问题,内部情感研究主要解决计算机内部如何用情感来对其

决策做出贡献。^{[34] (P. 17)}外在情感研究是一个比较单纯的技术问题,内部情感研究涉及情感的生成机制等理论问题,甚至涉及哲学问题。本文讨论的核心,是内部情感研究的问题。情感的生成必然有一定的物质基础,而探明这些物质基础的转换现象,从技术上讲是可能的。一旦这些现象被探明,再根据需要合成或生成情感也就是可能的。

通过技术手段实现人造情感,可能会推动人文学科的更快发展。肖峰认为人造情感可以“弥补自然情感的不足,从而丰富人的情感世界,克服情感匮乏等人文问题”,该技术“拓展了技术与人文交融的新境界,成为促进人的全面发展的新手段”。^[35]人造情感如果实现,对人文学科的影响当然会无比巨大,甚至可能彻底改写人文学科的知识结构。但另一方面,如果人文学科自身裹足不前,不向意识深处挺进,实现高级的人造情感恐怕也将遥遥无期。因此,人造情感技术与人文学科的发展是相互促进的,二者合力作用,才能不断向意识深处挺进,揭开人类情感和智能的秘密。对人造情感的探索,可以从对其他可能情感的探索入手进行。对未命名情感的思考,可以打破语言造成的思维局限;对尚未体验过的情感思考,可以探索情感的形成过程;对他物种情感的研究,将有助于情感的物质因素分析。如上文所述,数字技术与人文的差别主要是思考方向的差别,单方向的思考难免都有自己的缺陷,两个方向的思考就可以做到互相补充、互相验证。

数字技术与人文学科联手才能真正实现人工智能,也才能推进两种学科的进一步发展。这不仅涉及人工智能,也涉及科学和人文研究的方方面面,这必然导致科学和人文的双双向转。一方面,人文学科必然面临一个数字转向的问题。李点指出“数字媒介和传统人文的联姻已成事实”^[36]，“数字人文所代表的方法和思想首先是对我们所处的数字化时代的理性认可,预示了人文学科在后理论时代的某种‘数字转向’。”^[37]韩晗认为“在大数据时代,人文研究必然会发生新的转向。人文研究要与大数据相结合,而大数据又依赖于社会学、统计学而存在。”^[38]因为人文学科只有利用大数据提供的数字,才能避免流于空泛。另一方面,数字人文的发展也必然导致科学思考的人文转向。徐晓风和刘英指出,20世纪下半叶以来,西方科学哲学出现了文化方面的转向,科学与人文互相融合,“一方面,只有通过人文精神的

引领,科学技术的发展才不会脱离人类的向度,才不会变成一种毫无人性的功利性探究;另一方面,科学精神则应当不断内化为人文精神的一部分,只有如此,人文精神才能更富科学性、清晰性、准确性,也才能更好地指导实践。”^[39]

数字人文研究是一个新领域,也是一个学术研究的回归。人类的学问最初是不分文理的,到亚里士多德的时候就分开了。但是亚里士多德认为其区别主要在于研究对象不同,研究事物具体形态变化的被称为“物理学”,不属于前一系列的讨论叫“物理学之后”(metaphysics)即形而上学,这也就是后来自然科学和人文学科分野的雏形。《易传》中也有“形而上者谓之道”“形而下者谓之器”的区分,与亚氏的分法是一致的。但是古时候的区分仅仅是指研究对象的区分,对二者研究路径的根本区别一直没有讨论清楚。究其原因,是因为科学研究始终暗含了目的,而人文研究始终试图找到因果,二者一直混杂,人们没有想清楚二者在两种学科中的思路实际上早已分道扬镳。后来随着目的论的衰落和因果论的兴起,科学和人文都在强调因果,而作为人文精神核心的目的论也就逐渐被人文学者遗忘,后来竟然难于被接受。可喜的变化始于现代科学的兴起,“目的论又成为科学家和哲学家们热议的话题”^{[40] (p.24)}。目的论被重新讨论与被认可,最终必将与因果论一道,为今后的科学研究,特别是为情感、人工智能等复杂问题的研究做出贡献。

参考文献:

- [1]石琳. 智能虚拟环境中的人工情感研究[M]. 武汉: 武汉大学出版社, 2015.
- [2]刘大为. 论语言的情感表现[J]. 华东师范大学学报(哲学社会科学版), 1985(6).
- [3]阿来. 落不定的尘埃: 阿来藏地随笔[M]. 武汉: 长江文艺出版社, 2016.
- [4]倪文锦. 挑战与应答: 语文教学科学化、艺术化的追求[M]. 济南: 山东教育出版社, 2012.
- [5][美]大卫·铃木, [美]阿曼达·麦康纳. 神圣的平衡: 重寻人类的自然定位[M]. 何颖怡译. 汕头: 汕头大学出版社, 2003.
- [6]魏斌, 陈寒, 张妍, 黄素蓉. 人工情感原理及其应用[M]. 武汉: 华中科技大学出版社, 2017.
- [7]肖峰. 信息技术哲学[M]. 广州: 华南理工大学出版社, 2016.
- [8]仇德辉. 数理情感学[M]. 长沙: 湖南人民出版社, 2008.
- [9]Dreyfus H L. *What Computers Still Can't Do: a Critique of Artificial Reason*. MIT press, 1992.
- [10]熊培云. 西风东土: 两个世界的挫折[M]. 北京: 新星出版社, 2016.
- [11]卿文光. 思辨的希腊哲学史(一): 前智者派哲学[M].

- 北京: 人民日报出版社, 2015.
- [12][法]格雷马斯. 结构语文学: 方法研究[M]. 吴泓缈译. 北京: 生活·读书·新知三联书店, 1999.
- [13]赵毅衡. 关于认知符号学的思考: 人文还是科学? [J]. 符号与传媒, 2015(2).
- [14]李思屈. 精神符号学导论[J]. 中外文化与文论, 2015(3).
- [15]聂焱. 广义同义修辞学[M]. 北京: 中国社会科学出版社, 2009.
- [16]赵毅衡. 符号学[M]. 南京: 南京大学出版社, 2012.
- [17]周昌乐. 智能科学技术导论[M]. 北京: 机械工业出版社, 2015.
- [18]赵毅衡. 意义的意义之意义: 论符号学与现象学的结合部[J]. 学习与探索, 2015(1).
- [19][英]杰弗里·利奇. 语文学[M]. 李瑞华等译. 上海: 上海外语教育出版社, 1987.
- [20][法]格雷马斯. 论意义: 符号学论文集[M]. 吴泓缈, 冯学俊译. 天津: 百花文艺出版社, 2011.
- [21]包向飞. 康德的数学哲学[M]. 武汉: 武汉大学出版社, 2013.
- [22][美]威廉·韩思. 伦理学: 美国治学法[M]. 孟悦译. 北京: 社会科学文献出版社, 1994.
- [23][德]R. 奥伊肯. 人生之意义与价值: 奥伊肯卷[M]. 张蕾译. 北京: 北京联合出版公司, 2015.
- [24]周宪. 现代性的张力[M]. 北京: 首都师范大学出版社, 2001.
- [25][德]艾克哈特·托尔. 新世界: 灵性的觉醒(修订版)[M]. 张德芬译. 成都: 四川文艺出版社, 2016.
- [26][德]埃德蒙德·胡塞尔. 逻辑学与认识论导论(1906-1907年讲座)[M]. 郑辟瑞译. 北京: 商务印书馆, 2016.
- [27][德]马丁·海德格尔. 从莱布尼茨出发的逻辑学的形而上学始基[M]. 赵卫国译. 西安: 西北大学出版社, 2015.
- [28]赵毅衡. 意义理论: 符号现象学, 哲学符号学[J]. 符号与传媒, 2017(2).
- [29]沈勃. 十字路口的缪斯[M]. 南京: 江苏人民出版社, 1992.
- [30][美]乔纳森·爱德华兹. 宗教情感[M]. 杨基译. 北京: 生活·读书·新知三联书店, 2013.
- [31]黄璇. 情感与现代政治: 卢梭政治哲学研究[M]. 北京: 商务印书馆, 2016.
- [32]李蕾, 王小捷. 机器智能[M]. 北京: 清华大学出版社, 2016.
- [33][日]任福继. 人类感性认知与机器情感创生[A]//张玉台. 中国科协第五届青年学术年会文集: 科技工程与经济社会协调发展[C]. 北京: 中国科学技术出版社, 2004.
- [34]谷学静, 王志良, 郭宇承. 人机交互中的情感虚拟人技术[M]. 北京: 机械工业出版社, 2015.
- [35]肖峰. 人工情感: 技术与人文的新融合[J]. 马克思主义与现实, 2012(1).
- [36][美]李点. 面对数字人文的幽灵[J]. 燕山大学学报(哲学社会科学版), 2017(1).
- [37][美]李点. 人文学科在后理论时代的“数字转向”[N]. 社会科学报, 2017-08-10.
- [38]韩晗. 论“大数据”与人文研究的转向[J]. 晋阳学刊, 2014(3).
- [39]徐晓风, 刘英. 科学解释与人文理解: 重思科学哲学的“文化转向”[J]. 北方论丛, 2016(5).
- [40]董岩. 国家应对气候变化立法研究: 以立法目的多元论为视角[M]. 北京: 中国政法大学出版社, 2015.

收稿日期 2018-05-10 责任编辑 申 燕