

人机合作中集体能动性的建构

于 雪,初 昊

(大连理工大学 人文与社会科学学部,辽宁 大连 116024)

摘要:基于多能动者的人机合作是当前人机关系的重要表现形式之一,人与机器作为能动者能够实现交流、协同与合作。为了实现更加有效和优质的人机合作,应当建构集体能动性。这种建构是可能的,因为人机合作满足集体行动、集体意向和共同结果这三个建构条件。人机合作中集体能动性的建构包括认知维度和行为维度。认知维度借用“我们一模式”和“共享意向”理论内在解释人机合作中的集体能动性;行为维度借用联合行动模型和随附关系理论观察和模拟外部行动从而建构人机合作中的集体能动性。

关键词:人机合作;集体能动性;认知维度;行为维度;集体意向;人机交互技术;人工智能

[中图分类号]TP18;C912.4 [文献标识码]A [文章编号]1672-934X(2023)02-0062-08

DOI:10.16573/j.cnki.1672-934x.2023.02.007

The Construction of Collective Motility in Human-Machine Cooperation
Yu Xue, Chu Hao

(Division of Humanities and Social Sciences, Dalian University of Technology, Dalian, Liaoning 116024, China)

Abstract: At present, human-machine cooperation based on multiple agents is one of the important forms of human-machine relationship, where human beings and machines as agents are able to communicate, collaborate and cooperate. In order to achieve more effective and higher quality human-machine cooperation, a kind of collective motility should be constructed. The possibility of such a construct is due to that human-machine cooperation meets three conditions, namely, collective action, collective intention and common outcome. The construct of collective motility in their cooperation consists of two dimensions: a cognitive dimension and a behavioral dimension. In their cooperation, respectively, the former draws on the theories of "we-mode" and "shared intention" to intrinsically explain, and the latter observes and simulates the external actions through the joint action model and the dependency theory to construct, the collective motility.

Key words: human-machine cooperation; collective motility; cognitive dimension; behavioral dimension; collective intention; human-machine interaction technology; artificial intelligence

20 世纪中叶以来,人机交互技术得到了广泛的发展,前期的人机交互技术多停留在静态层面,用于辅助用户控制机器,例如人机交互界

面。然而,随着机器行动力的增强,现代智能机器不仅能够处理信息,而且能够像人类能动者一样,参与到越来越复杂的耦合动态系统中,这

收稿日期:2022-08-16

基金项目:国家社会科学基金一般项目(22BZX025)

作者简介:于 雪(1989—),女,副教授,哲学博士,主要从事技术哲学、技术伦理等研究;
初 昊(1993—),男,硕士研究生,研究方向为技术哲学、技术伦理。

就要求人与机器通过合作的方式完成任务。例如,在人与智能家庭控制系统或者军事上的排雷机器进行合作的过程中,人们逐步发现机器的出现使得工作变得更有意义^[1],而且机器表现出越来越类似于“伙伴”的重要性。按照尼基塔·马塔尔(Nikita Mattar)和艾普科·瓦克斯穆特(Ipke Wachsmuth)的观点,人类正在将智能机器的地位从工具提升为活动中的能动者,并在与其进行的共同行动中,将其视为具有高度交互性的“伙伴”^[2]。这表示在共同完成任务的时候,机器也可以在与人类的互动中推动并实施任务。作为“伙伴”的机器,与人类能动者一同构成了行动的主体。当前的研究多倾向于从技术层面探讨人机合作的不同模式,相对缺少从哲学层面思考如何让人机更加有效和优质地合作。集体能动性的切入能够很好地回应这个问题,即为人机合作建构集体能动性能够使人机之间更好地合作。因此,本文尝试为人机合作建构集体能动性。在此之前,本文首先需要论证人机合作如何成为一个集体以及作为集体的人机合作为何需要建构集体能动性。

一、基于多能动者的人机合作

“人机合作”(Human-Machine Cooperation, HMC)指的是人与机器通过合作的方式完成共同的任务。因为这里的“机器”多表现为具有行动力的机器人,因此人机合作也包含了人一机器人合作(Human-Robot Cooperation, HRC)。通常对人机合作的理解是基于双能动者的人机合作,即单个人与单个机器的双能动者合作模式。当前人机合作的多样化表明,人机合作不仅表现为单个人与单个机器的合作,还有可能出现单个人与多个机器的合作以及多个人与单个机器或多个人与多个机器的多能动者合作模式。因此,当前的人机合作是多能动者系统(Multi-Agent Systems, MAS)的一种表现。MAS是指在一个环境中由多个交互的能动者组成的系统。该系统可以是生物系统、社会系统,也可以是人工物系统。MAS的优势是可以

解决单个能动者或中央系统难以解决甚至是不可能解决的问题^[3]。人机合作作为MAS的一种表现形式,其关键在于人和机器通过合作的方式产生共同的结果。因而,对人机合作的理解应当在MAS的框架内,并且应当对人机合作的先决条件及核心概念作出说明。

(一)人机合作的先决条件

为了实现人机合作,至少需要满足以下两个条件。一是任务定义(task definition):操作员(多数情况下是指人类)应该能够以简单的方式定义需执行的任务。任务的定义可以分为任务规范和任务描述,其中任务规范(task specification)是指操作员提供任务执行所必需的参数,例如参数可以是要处理的对象或其位置;任务描述(task description)则是指通过描述该任务,操作员向机器发出有关如何解决特定任务的指令,例如任务描述是通过选择适当的基本操作并在时间空间中对它们进行排序而发生的。二是动态功能分配(dynamic function allocation):指的是根据人类和机器各自的长处和局限性以及环境的变化,将任务进行分解,动态合理地分配给人类和机器,使二者能够取长补短,相互匹配和协调,安全、高效地完成人类或者机器不能单独完成的工作任务。

(二)人机合作的三个核心概念

人机合作可以在两个层面实现:一是符号层面,即网络上的数字信息交换;二是物理层面,即视觉、触觉、听觉。在这两个层面上,人机合作涉及三个核心概念,分别是交流(communication)、协调(coordination)与合作(cooperation)。

首先,交流。人与机器之间的交流指的是物理层面的信息传递和信息解读。信息传递是信息从一个合作伙伴传递到另一个合作伙伴的过程,这一过程通过预先设定的交互模式或协议,使接受者能够理解信息发送者的意图。在物理层面的人机合作中,仅有数字化的数据交流传递是不充分的,还需要有接触力。比如,传感器不仅可以应用于任务执行,还可以应用于

识别交互的情况并接收传输信息。信息解读对于已传递信息而言非常重要,特别是在人机合作过程中能够避免错误的解读,这对防止人类受伤具有重要作用。其次,协调。人与机器在工作中会有重叠的工作空间,为了避免彼此之间相互干扰和冲突,需要在一个行动中规划和协调多个能动者的活动。为了实现人机协调,需要先提取关于可预测的系统运行模式的技术约束条件,例如动力、安全。功能或功能失调的系统分析方法可用于推断(或诱导)处理系统所需的故障排除任务。然后,这些故障排除任务可以根据不同的要求分别分配给人和机器。之后,可以设计自动化处理器来管理未来的自动化任务,以及设计有助于促进未来人类任务的人机界面^[4]。最后,合作。考虑到单一能动者并不具备所有的专业知识、资源或信息,为了更好地实现目标需要彼此合作,其中每个合作伙伴都以其特定的能力为实现最终目标作出贡献。鲁斯·德容(Roos de Jong)提出了一种分层合作机制。例如,自动驾驶汽车涉及两层人机合作机制——“司机—汽车”的合作和“程序员—汽车”的合作,这两层合作有各自的任务目标。同一个机器可能会同时参与到多个人机合作的层级中,在这些合作中大部分工作是由机器承担,而人类承担了发起合作、监督合作和管理合作的任务^[5]。

基于多能动者的人机合作对集体能动性提出了内在的诉求,不同能动者自身的能动性应当通过合作的方式而得到集体的表征,集体能动性的建构将有助于不同能动者之间开展更加有效的合作。

二、人机合作中建构集体能动性的可能性

为了更好地实现有效和优质的人机合作,应当建构人机合作中的集体能动性。基于对集体能动性的条件分析,本文认为在人机合作中建构集体能动性是可能的,因为人机合作满足集体能动性的三个基本条件:集体行动、集体意向和共同结果。

(一)集体能动性

将集体作为主体对其进行阐释是集体本体论的核心要义^[6],而集体本体论关注的两个核心概念就是集体能动者和集体能动性,并且二者紧密相关。在行动哲学的框架中,能动者(agent)指的是能够采取行动的主体,而能动性则是指采取行动的能力。集体能动者(collective agents或group agents)是指将集体作为一个行动主体,而集体能动性(collective agency或group agency)则是指该主体在参与社会行动的过程中表现出的行动能力,即集体行动主体的能动性。换言之,多个个体成员联合起来组成一个集体,通过一些方式来进行彼此间的交流、协作和合作,确立统一的目标以及对自身行为作出承诺,并以集体作为统一的行为主体而采取行动。

但是,并非所有的集体都存在集体能动性,行动哲学框架中满足集体能动性的条件至少有三点:一是该行动应是集体行动(collective actions)^[7]。集体行动也是一种多能动者行动(Multi-Agent Action, MAA),按照雷摩·图梅勒(Raimo Tuomela)的解释,一个行动是多能动者行动当且仅当该行动主体是一个集体且该行动并非意外^[8]。二是该行动应该具有集体意向(collective intention)。集体意向首先表现为集体的目标导向,这既包括一种外在的目标导向,即共同完成某事的目标,也包括一种内在的目标导向,即想要完成该目标的心理表征。其次,集体意向表现为信念、欲望、意向等具体的内在状态,集体的内在状态成为一种联合意向性(joint intentionality)直接影响到集体行动。最后,集体意向还表现为一种情感认同,即作为行动主体的集体在行动时应的情感层面的考量。贝内特·赫尔姆(Bennett Helm)十分重视能动者的情感维度,在他看来,真正的能动者之所以能够参与行动,是因为能动者真正关心这个行动并且认为这个行动是值得追求的^[9]。这对于集体能动者 also 具有重要的借鉴意义,因为这能够为集体能动性提供一个特定的评价视角来理解集

体行动是如何形成的。三是该行动将会产生一个共同结果。基于集体意向的集体行动将会产生一个共同的行动结果,即集体作为主体共同承担由其行动而产生的结果。

(二)人机合作中集体能动性的建构条件

基于对集体能动性的条件分析,本文认为人机合作满足上述三个条件。因此,为其建构集体能动性是可能的。

一是人机合作满足集体行动的条件。合作本身是指两个及以上主体通过多种方式执行任务,以达到共同目的的行动。基于多能动者的人机合作满足集体行动的条件,即该行为是集体成员有目的的行动。随着机器的智能化发展,人机合作表现为多样态的集体行动。在人机合作中,机器作为一个基于算法的行动主体参与其中。而随着诸如大数据分析以及深度学习等算法的不断优化升级,机器越来越能够通过调整自身来配合其他集体成员,一起对共同目标进行调整并最终达成共识,进而互相协调采取统一行动。例如,当下常见的家庭扫地机器人,在经过最初的设置之后,机器本身会根据环境状况进行整体任务分析以及路径规划,并且同人类成员进行功能上的互补。在人类成员对家具布局进行改变的情况下,机器人对环境进行动态分析,重新调整任务内容和路径规划,使得在合作中和人类成员一同动态地对整个家庭进行全面清洁。因此,人机合作满足集体行动的条件。

二是人机合作满足集体意向的条件。自从弗朗兹·布伦塔诺(Franz Clemens Brentano)将意向性这一概念引入哲学领域之后,有关意向及意向性的讨论便日益深入。格特鲁德·安斯康姆(Gertrude E. M. Anscombe)就已经讨论了个体行动是受到个体意向影响的这种紧密关系^[10]。人机合作之所以满足集体意向的条件,是因为人机集体作为合作行动中的行为主体,具备一种集体想要实现目标的内在驱动。由于机器智能化程度的不断提升,集体内的机器可以在不断变化的环境中计算并规划所在集

体的行动方案以及内部的动态配置。但是,人和机器在合作过程中所展现的这种个体成员将所在集体作为第一视角来进行思考的情形,本身就体现出个体是在根据已知其他成员的行为模式,作出的对其他成员的意向性预测和假设。每一个成员的这种意向通过沟通和整合就形成了一种集体意向,进而促进合作实现以及集体行动。

三是人机合作满足共同结果的条件。基于集体意向的人机合作将会产生一个共同的行动结果,例如,在与计算机进行精密化仪器的加工合作中,人机合作作为整个生产行动的主体,最终产出精密化仪器就是二者合作的共同结果。在当前的日常生活中,人们越来越多地使用各类地图导航程序来规划路线,这就是与机器合作完成任务的表现之一。在这一合作过程中,人类会对目的地以及沿路的各种情况进行设定,而机器将会根据动态路线情况变化而不断调整路线,最终实现规划出最优路径这一共同结果。最后,作为集体的人机合作应当承担由其合作行动而产生的结果,如规划不明或行驶偏航导致的路线错误。

鉴于人机合作能够满足集体能动性的基本条件,即人机合作作为一类集体行动,体现集体意向并据此产生共同结果,建构人机合作中的集体能动性是可能的。

三、人机合作中集体能动性的建构路径

在论证人机合作中建构集体能动性可能的基础上,本文转向论证人机合作中建构集体能动性的现实性。从哲学层面上建构集体能动性主要有两种方式:解释和模拟。解释的进路指的是对人机合作中的集体能动性进行内在的解释,这种进路从认知维度出发,主要借用“我们一模式”(we-mode)和“共享意向”(shared intention)的理论解释人机合作中的集体能动性。模拟的进路是指对人机合作中的集体能动性进行外部的观测和模拟,这种进路从行为维度出发,主要借用联合行动模型和随附关系理

论建构人机合作中的集体能动性。

(一)从认知维度建构人机合作中的集体能动性

从认知维度建构人机合作中的集体能动性重点在于解释如何使得作为集体的人机合作的内部成员之间可以进行充分有效的沟通,从而使得这一集体作为行动主体向已设定好的集体目标行动。换言之,如何在作为集体行动的人机合作中建构集体意向。这种建构主要依据“我们一模式”和“共享意向”这两种方式。

1.“我们一模式”的集体意向建构

在人机合作中集体意向的建构方式上,认知维度的方法是对集体的内部思想进行共同的集体概念的建构。这种集体概念的建构,首先是要建构集体意向。集体意向的建构源于约翰·塞尔(John Searle)的“我们一意向”(we-intention)理论模型,在塞尔看来,集体中的每一个个体为了统一目标努力是成员之间的一个共同尝试^[11]。图梅勒在此基础上,对意向内容进行了更深层次的探讨,从而建立了“我们一模式”这个理论模型^[12],也就是参与者从集体角度来产生共同意向。为此,图梅勒对能动者在“我们一模式”和“亲群体的我一模式”(pro-group I-mode)中的表现进行了区分。^①在合作中,每一个个体能动者,其想法如果是“这种情况下,我应该做什么”,则是后者模式的体现;而在“我们一模式”中,每个人的想法更多的是“我们应该做些什么”。也就是说,在“亲群体的我一模式”下能动者有意识参与集体活动是源于其个人意愿。而在“我们一模式”中,能动者所追求的共同目标是由集体的共同的理念所决定的。而正是通过这种思考模式,集体成员在思维上形成了共同的集体意向,确立了集体的主体性地位,为集体意向的建构奠定了基础。

人机合作中集体意向的建构可以采用这种“我们一模式”方法。一方面,采用“我们一模式”方法建构人机合作中的集体意向可以通过建构一个集体理念(ethos)来实现^[13]。集体理念与集体目标不同,合作前人类能动者与机器

能动者都具有仅仅依靠自身是无法完成目标的这种基础的信念,以此为开始形成“需要合作”的集体理念。在这种集体理念的指导下,形成共同的集体目标。同时,建构人机合作的集体理念有助于促进集体目标的完成,因为这种集体理念会使得集体成员具有一种内在的驱动力,从集体的角度出发来思考如何实现集体的目标,并朝着完成这一目标而努力。

另一方面,采用“我们一模式”方法可以为人机合作建构一个集体目标,即为了合作的集体目标。人机合作的集体目标应当是集体中所有能动者的共同决定,因而集体目标不能被简单地还原为人类能动者或机器能动者的单独决定。在行动中,集体能动者才是集体行动的关键,因而每一个集体能动者都需要一种出于逻辑思考所形成的“我们一态度”(we-attitude)来支持合作行动^[14]。为了合作的集体目标是集体中的所有能动者都应当接受的,这种集体目标也由此确保了集体中的能动者会去采取行动的这一共同态度。

2.通过“共享意向”建构集体能动性

在人机合作中,集体能动性的另外一种建构方式是“共享意向”。“共享意向”是米歇尔·布莱特曼(Michael E. Bratman)提出的,指的是每个参与者的相关意向共同组成了“共享意向”,并且在适当的背景下适当地相互关联^[15]。“共享意向”是一种主要由参与者的态度和这些态度之间的相互关系组成的事务状态,与关注个体能动者内在态度或意向不同。换言之,“共享意向”存在于参与者的态度和这些态度之间的关系中,但这些态度本身却并不是“共享意向”。因此,“共享意向”必须要不少于两个人才能形成,并且参与者们需要具有对于同一个目标的共同知识,有意向参加一个合作的行动并根据对方的意向来追求目标,同时愿意妥协和调整各自的子计划。对于集体中的每一个成员来说,“共享意向”模式的建构就是有意向去做一件事,并且通过调整自身的子计划以参与到“共享意向”中。

人机合作中集体能动性的建构可以利用“共享意向”模式来建构集体意向,该模式重点关注集体意向的形成条件。从功能上来说,机器作为“共享意向”的参与者,是可以实现对原有的人类参与者的替代。对于机器而言,由于其强大的信息储备能力,也可以被认为具有了与人类参与者相当的掌握知识的能力。机器在参与合作的行动中,能够接受其他参与者的意向指示,并在追求共同目标时,可以凭借其强大的计算能力,不断地根据其他成员以及合作集体的目标来动态调整自身的子计划。此外,在机器高速计算能力的保证下,这种人机合作中的“共享意向”有助于确立人机集体的行动主体地位,也有助于发挥人与机器的各自优势以提升合作效率。

总体来看,“我们一模式”和“共享意向”的集体意向建构是为人机合作建构集体能动性的重要方式,人机合作所构成的集体可以通过集体协商过程形成集体理念,并最终达成集体目标,提升合作效率。

(二)从行为维度建构人机合作中的集体能动性

从行为维度建构人机合作中的集体能动性,重点在于对人机合作进行外部的观测和模拟,得到对目标完成情况的一个外部评判,这种建构可以借用联合行动模式和随附关系理论来进行论证。

1. 联合行动模式的集体能动性建构

从行为维度上来说,合作是一种相互影响的能动者之间的协调行为。对于第三方观察者来说,合作中的协调行为越是复杂,这种“合作现象”在行为维度上作为连续统一体的等级就越高^[16]。这种多能动者的合作现象,可以被称为“联合行动”(joint action)。

纳塔莉·塞邦斯(Natalie Sebanz)、哈罗德·伯克宁(Harold Bekkering)和巩特尔·诺布利奇(Gunther Knoblich)对联合行动的定义是:两个或两个以上的人在空间和时间上协调行动

以改变环境的任何形式的社会互动^[17]。可见,参与联合行动的能动者的共同目标在于改变环境,这可以通过诸如共同关注、行动观察、任务分享和行动协调等行动来实现。这些行动都有可能在不涉及集体意向的情况下,促成联合行动。因此,联合行动模式对于集体能动性的建构包含三个条件:一是两个(或更多)能动者在空间和时间上协调它们的行为;二是这个行为是从外部观测的;三是这个行为会带来环境的变化。

塞邦斯等人的经典联合行动理论,可以作为一个具有代表性的理论模型来建构人机合作中的集体能动性。根据联合行动模式,人与机器作为联合行动的参与者,可以为其建构改变环境的共同目标。在这一模式中,人类和机器都在同一目标的指引下,在互相无交流的情况下通过各自的行动来改变环境。同时,每一个参与者也会根据其他参与者对环境的变化,而随时调整改变自身的行为,以尽快实现最终目标。因此,基于联合行动模式的集体能动性建构,首先在于促使人和机器在行动上协调彼此的行为;其次要确保该行为是可以被外部观测的;最后确保其观测结果是有意义的。

2. 基于随附关系理论的集体能动性建构

“随附”(supervenience)概念由菲利普·佩蒂特(Philip Pettit)提出,指的是:当我们说,一组事实B,是另一组事实A的“随附”,那么必然地当且仅当事实A确定的时候B也随之确定,而如果A的定域不发生改变,那么B的定域也不会有变化^[18]。在随附关系理论中,集体与个人更像是“输出”与“输入”的关系。集体的结果作为一种“输出”,必然要随附于集体中个人成员的“输入”,并且这种随附关系可以确保集体在行动中作出理性判断。这里的随附关系并不是一种“依附关系”,因为不同的“随附”方式会带来不同的结果,集体决策不能还原为集体内成员个人的决策。这是由于集体的决策受到集体中每一位成员的决策的影响,如果没有众多成员的决策判断,也就无从形成集体的决策判

断。随附关系理论揭示了集体具有一种能够独立于其内部各个成员而展现自身明确目标的能力,这对于建构集体能动性十分重要,且主要体现在建构集体判断上。

一方面,人机合作中的集体判断可以通过多数人的随附(majoritarian supervenience)来建构,即对议程中每个命题的集体判断都是对该命题的绝对多数判断^[18]。对于每一个议程中的每一项指标,都按照多数主义的判断标准来处理结果,就可以得到作为一个独立主体的集体判断。而这个时候,成员个体的判断对集体来说并非充分必要条件,因此集体是具有能动性的。如表 1 所示,甲乙丙组成的采购委员会分别对供应商的价格、质量和运输能力三项进行审核,其中每一位评委都对供应商的某一项表示不满,但对每一项进行综合统计后,却依旧评定为满意,并最终批准引进。由此可以看出,对于这个集体行动主体来说,个体成员的判断结果并不能决定集体判断的结果。

表 1 随附关系下采购委员会的审核结果

| | 价格优势 | 质量水平 | 运输能力 | 是否引进 |
|-----------|------|------|------|------|
| 甲 | 否 | 是 | 是 | |
| 乙 | 是 | 否 | 是 | |
| 丙 | 是 | 是 | 否 | |
| 甲 & 乙 & 丙 | 是 | 是 | 是 | 是 |

另一方面,人机合作中的集体判断可以通过鲁棒的集体理性(robust group rationality)来建构,即通过建构产生的集体判断应当是非常完整的、一致的以及演绎闭合的,及时面对所有可能的个人判断组合时,集体判断都能保持鲁棒性^[19]。在人机合作集体中,每个成员都是一种“输入”的角色,并且依靠严密的逻辑推理运算,最终“输出”集体判断,而这一集体判断是独立于个人判断的集体的结果,由此可以确立其自身的鲁棒性。在建构人机合作集体能动性的过程中,各成员对于整个决策过程的贡献则无所谓同质或异质。因为这里的决策过程是一个

随附的过程,最终的集体决策是符合严格的完整闭合的逻辑判断。因此,人机合作中的各成员只要在决策的过程中作出符合逻辑一致的判断,并且在最终的集体判断呈现之后遵循这一判断行动,就可以从行为维度建构集体能动性。

总体来说,联合行动和随附关系中集体行动的建构也是人机合作中集体能动性建构的重要方式,人机合作所构成的集体产生了可以外部观测和模拟的集体行为,从而形成了集体判断,由此建构了集体能动性。

四、余论:人机合作中建构集体能动性的局限性

人机合作中集体能动性的建构是为应对机器的普及率越来越高、迭代速度越来越快的需求。尽管在人机合作中建构集体能动性具有重要的现实意义,即优化人机合作以实现共赢^[20]。但是,建构集体能动性的局限性也不容忽视,尤其是集体的责任归因问题。作为集体的人机合作将在责任归因问题上面临双重困境:一方面,尚不清楚组成该集体的个体是否应被认为要对集体行动的结果负责;另一方面,如果将责任归因于集体该如何负责,也有待研究。

以自动驾驶汽车为例,由于自动驾驶系统并不能完全应对复杂的交通路况,因此人和自动驾驶系统需要以合作的方式共同完成驾驶任务。人与自动驾驶的合作模式同样会发生交通事故,比如美国特斯拉公司的自动驾驶汽车就曾发生多起事故,而这些事故的责任归因问题非常棘手。这些问题,一方面,包括人类驾驶员和自动驾驶系统作为集体中的个体成员是否需要为这一集体行动产生的事故负责。如果需要,那么自动驾驶系统要如何为其自身的决策和行为承担责任?自动驾驶系统是否具有承担责任的能力?是否可以运用类似人类的奖惩措施针对自动驾驶系统?或者可否通过对自动驾驶程序进行重新设置的方式来承担责任?而另一方面,包括如果认定由集体来承担事故的责

任,那么人机合作作为集体可以通过怎样的方式来承担这一责任,并且内部责任又需进行如何的分配呢^[21]? 人类成员是否需要替代机器“伙伴”来承担集体责任? 这些问题需要在机器的发展过程中,以及机器和人类之间的不断合作融合中,从多个方面进行持续而深入的探讨。

尽管当前有一些尝试对集体进行责任归因以解决这种责任的“多面手”(many hands)问题,如共享责任(shared responsibility)、混合责任(hybrid responsibility)、集体责任(collective responsibility)、责任网络(responsibility network)等。但是,这些尝试几乎都是强调前置责任(forward-looking responsibility),而没有触及后置责任(backward-looking responsibility),即谁来对已经发生的事情的后果承担责任。

综上所述,人机合作中的集体能动性建构是行动哲学所关心的话题,这种建构能够为集体主义的发展提供当代案例。当下,正在飞速发展的机器已经突破了传统的工具属性,并表现出作为人类“伙伴”的重要性,这也为人机合作中建构集体能动性提供了可能。通过认知维度和行为维度的建构方式,本文为人机合作中集体能动性的建构提供了现实性,以促进人机之间更加有效和更加优质的合作。与此同时,人机合作中集体责任的划分、集体道德主体的确立以及人机合作中人的限度等问题也是该研究中不容忽视的问题^[22]。总体来看,人机合作中的集体能动性建构问题是人机关系研究中的重要课题,仍需要从更多角度对此进行更加深入的研究与挖掘。

[注释]

① “亲群体的我一模式”中的“亲群体”(pro-group)意在表现出,虽然只是一种成员产生私人意向,但也是站在群体的角度下来考虑自身的意向和行动。

[参考文献]

[1] Nyholm S, Smids J. Can a robot be a good colleague?

[J]. Science and Engineering Ethics, 2020, 26 (04): 2169-2188.

[2] Mattar N, Wachsmuth I. Strangers and friends [A]// Kurosu M. Human-computer interaction towards intelligent and implicit interaction [C]. Heidelberg: Springer, 2013: 102-111.

[3] Misselhorn C. Collective agency and cooperation in natural and artificial systems [A]// Misselhorn C. Collective agency and cooperation in natural and artificial systems [C]. Cham: Springer, 2015: 3-24.

[4] Hoc J M. Towards a cognitive approach to human-machine cooperation in dynamic situations [J]. International Journal of Human-Computer Studies, 2001, 54(04): 509-540.

[5] Jong R D. The retribution-gap and responsibility-loci related to robots and automated technologies: a reply to nyholm [J]. Science and Engineering Ethics, 2020, 26 (02): 727-735.

[6] 邹桑, 丛杭青. 集体行动者与能动性: 集体本体论的两个核心概念 [J]. 自然辩证法研究, 2017(2): 20-25.

[7] Bardsley N. On collective intentions: collective action in economics and philosophy [J]. Synthese, 2007, 157(02): 141-159.

[8] Tuomela R. A theory of social action [M]. Dordrecht: Springer, 1984: 1-10.

[9] Helm B. Emotional reason; deliberation, motivation, and the nature of value [M]. Cambridge: Cambridge University Press, 2001: 75-81.

[10] Anscombe G E M. Action, intention and ‘Double Effect’ [A]// Geach M, Geach L. Human life, action and ethics: essays by GEM anscombe [C]. Exeter: Imprint Academic, 2005: 196-216.

[11] Searle J. Collective intentions and actions [A]// Cohen P R, Morgan J, Pollack M E. Intentions in communication [C]. Cambridge: MIT Press, 2003: 401-415.

[12] Tuomela R. We-intentions revisited [J]. Philosophical Studies, 2005, 125(03): 327-369.

[13] Tuomela R, Tuomela M. Cooperation and trust in group context [J]. Mind and Society, 2005, 4(01): 49-84.

[14] Tuomela R. Who is afraid of group agents and group minds? [A]// Schmitz M, Kobow B, Schmid H. The background of social reality [C]. Dordrecht: Springer, 2013: 13-35.