

智能时代下人机交互的机遇与挑战

——第11期CCF秀湖会议报告

整理: 王党校 喻纯 韩腾 王运涛 王冠云 田丰

背景

人机交互是人与计算设备之间的双向交流,研究和实现适应人类感知-认知-行为规律的交互式计算系统,是机器智能和人类认知综合进步的途径,对人类利用数字生产资料提升生产力具有决定性作用。2021年5月,习近平总书记在两院院士大会上指出,以信息技术、人工智能为代表的新兴科技快速发展,大大拓展了时间、空间和人们的认知范围,人类正在进入一个人机物三元融合的万物智能互联时代。随着手机的创新逐渐放缓,面向人机物三元融合的新一代计算形态正在形成,如同20世纪80年代微软、苹果公司的图形用户界面技术和2007年苹果公司的多指触控技术对个人电脑和手机的推动作用一样,人机交互的突破也将是推动人机物三元融合产业发展和行业转型升级的重要力量。

长期以来,人机交互在交互认知理论和模型、普适界面范式和工具、人机协同操作的交互技术等关键科学和技术问题方面面临挑战。智能时代人机关系快速变革,机器开始承担以往由人脑执行的认知任务,这一变化不仅改变了信息处理过程,影响了交互的过程和机制,还推动了人机协同架构的确立和人机共进系统的发展。人机交互技术是实现人机之间高效对话、协同与共进的桥梁,亟须搭建人机交互技术的研究与开发框架,深刻揭示人类感知-认知-运动机制,对动态交互任务与情境准确建模,提升机器对人的动作、言语等多模态行为内蕴的交

互意图的推理能力,动态生成多通道优化的信息呈现界面。

2024年3月29~31日,第11期CCF秀湖会议在苏州CCF业务总部&学术交流中心举办,来自学术界与工业界的二十余位专家学者围绕“智能时代下人机交互的机遇与挑战”这一主题开展探讨和交流。参会者中既有在人机交互、虚拟/增强现实、视听触融合交互、人机混合智能、大模型等领域开展深度研究的学者,也有深耕教育、健康等多个应用领域的技术专家。他们围绕人机交互新场景与新应用,智能时代下推动交互自然性发展的新技术,人机交互、大模型与具身智能,智能时代下交互理论和范式变迁的新视角四个专题,从人机群组智能、新智能交互终端、多通道交互接口、视听触融合信息呈现以及交互技术应用领域等多个方面分享了自己的观点和看法,既有对前沿技术与工具系统的探索,也有对行业概况与企业实践的经验分享。嘉宾们普遍认为,亟须在智能时代下突破以人为中心的人机交互理论、范式与技术,积极拥抱人工智能技术,探究人机协同与人机共进的技术路径,共同推进中国人机交互领域的持续发展。

会议最后半天,与会嘉宾围绕“智能时代下人机交互”进行了专题研讨,一方面在此前的会议讨论基础上进行总结并形成共识,另一方面就如何共同推动智能人机交互的发展发出倡议。经过热烈的讨论,嘉宾们围绕智能时代下人机交互领域未来发展、人机交互原理与范式、人机交互核心技术栈、

人机交互新场景与应用、高校等科研机构的人才培养模式等方面形成了初步共识。

人机交互应用场景

中国航天科技集团有限公司第五研究院总体设计部副主任设计师朱超重点介绍了空间站中机械臂遥操作的交互场景，强调安全高效为空间站人机交互的宗旨，需要综合考虑空间站特殊环境对人的认知与运动能力的影响，突破多模态的高效交互界面技术、多通道的信息呈现技术，研发可靠安全的交互界面软硬件系统，实现对空间站机械臂等作业设备的天地协同交互。

暨南大学广东智慧教育研究院教授刘子韬围绕人机交互在教育场景下的应用，重点介绍高质量教学、个性化学习以及普惠的成本三大难题，提出智能时代教育领域急需突破人机智能混合的新型教学方法、学生学习场景与知识能力的个性化建模计算技术以及普惠的机器智能能力等关键方法与技术，同时需要综合运用人机交互中的设计理念实现高可靠性的智慧教育产品。

北京理工大学研究员翁冬冬汇报了数字人交互关键技术与应用场景，重点围绕孪生数字人和原生数字人两种形态，介绍了孪生数字人的建模技术、智能体驱动的数字人交互技术等，并重点介绍了数字人在在线直播、展览介绍以及线下场景中的应用。

参会的嘉宾在研讨中介绍了人机交互在智能手机、车载终端、AR/VR头戴设备以及3D显示器等传统和新兴交互终端上人机交互的应用场景，形成了以下主要观点：交互从可控的二维表面扩展到非约束的三维空间，急需人机交互界面范式的革新。

企业嘉宾强调了人机交互是用户终端设备的关键组成，形成了共识：人机交互是决定用户终端可用性的关键，直接决定了终端产业规模。嘉宾均认为人机交互在应用中的核心是以人为本，认为技术的发展应该以人为中心。在人机交互人才培养以及企业人才需求方面，嘉宾认为亟须突破学科壁垒，形成更加成熟且满足企业需求的人

机交互人才培养体系。

人机交互与人工智能（2I关系）

日本工程院院士、日本高知工科大学教授任向实带来了题为“人机共协计算”的特邀报告。他从人机交互技术的哲学思考谈起，探讨了智能时代下计算机技术给人类生活带来的变革，强调了技术是人类内心的外在投射。他还强调了以人为中心的智能技术对实现人机共进的必要性，表示计算机技术需要融入东方哲学思想，以提升人类心智能力为目标。最后他再次强调人机共协计算的宗旨是提升人类心智能力并为人类的可持续发展作出贡献。

西安交通大学教授薛建儒以“人机群组智能技术途径的若干思考”为题作了专题报告。报告深入探讨了自主智能领域，特别关注物理自主智能系统，这些系统当前的性能仍旧落后于人类。报告讨论了莫拉维克悖论（Moravec's Paradox），强调了开发能够适应动态复杂环境系统的挑战。在对未来智能系统生成自主行为以完成复杂任务的展望中，他强调需要改进人机交互。他还提到了通过增强学习和模拟环境训练这些系统的可能性，以及这些技术对提升机器学习效率的潜力。

复旦大学教授黄萱菁以“大模型的能力展现与价值对齐”为题作了专题报告。报告讨论了大型语言模型的发展及其含义，涉及人工智能与人类价值和意图对齐的挑战，人工智能模型的演进，它们理解和生成类似人类的响应的能力，以及围绕人工智能互动的伦理考虑。报告强调创建技术精湛并与社会价值观伦理对齐的模型的重要性。她提出了构建透明、负责任的人工智能系统的必要性，强调了在人工智能研发过程中考虑社会影响和伦理原则的重要性。

北京航空航天大学教授任磊以“工业大模型：打造工业互联网AI新基座”为题作了专题报告。报告涉及人工智能在制造业中的集成，详细介绍了迈向工业4.0的进展以及从数字化向智能化的转变；探讨了人工智能与制造过程结合是如何引领创新生

产方法的，例如个性化制造和数字孪生技术，这些都将重新定义行业的未来格局。他强调智能制造在提升效率、降低成本以及促进可持续发展方面的关键作用，并探讨了人工智能技术在解决复杂制造问题中的应用。

北京科技大学教授班晓娟以“智能交互及其在医学领域中的应用”为题作了专题报告，关注人工智能与医疗保健的交汇点。报告深入探讨了可穿戴传感器和智能交互的进步，阐述这些技术如何转变患者护理和医学研究。探索多模态传感器、个性化健康监控和康复展示了人工智能增强医疗诊断和治疗计划的潜力。她强调了这些技术在促进患者自我管理、提高生活质量以及实现医疗服务的个性化和精准化方面的重要性。

专题报告之后，与会者针对大模型、具身智能与人机交互的深层次联系展开了集中研讨。讨论聚焦于如何通过大模型和具身智能技术，实现更加自然、高效的人机交互，促进虚拟与现实世界之间的无缝连接。与会专家共同探讨了大模型在理解人类语言、情感和社会行为方面的潜力，以及这些技术如何助力创建更加智能、互动性更强的虚拟环境。此外，对于虚拟空间中人机物融合的未来发展趋势，大家一致认为，这不仅要求技术上的突破，还需要对人类社会行为、伦理标准的深刻理解。此次研讨不仅加深了与会者对当前人工智能与人机交互相结合的认识，也为人机交互创新与应用指出了新的方向。

新范式、新理论

清华大学副教授喻纯以“智能时代的人机协同”为主题作引导报告。在大模型的背景下，他重新解读了 Licklider 关于通过人机交互与人工智能实现人机共生的观点，并提出人机交互的发展已从简单的功能访问升级到知识交换的层面，这一升级突显了交互式学习的重要性：在机器与用户的自然交互过程中，使机器能够实现智能的自主进化，这是智能时代下促进人机协同发展的关键。

湖南大学教授王巍以“智能人因交互设计和

触觉审美”为题作引导报告，报告以智能人因交互设计为例指出了工业设计对人机交互发展的重要作用，提出“求知”的人因研究与“求用”的设计创新之间相辅相成的方法范式；进而介绍了触觉体验设计与触觉美学领域当前的研究进展。

中山大学教授武汇岳以“技术中介的社会参与”为主题作引导报告。在智能技术的中介和推动下，会引发包括个体行为、协作策略、社区参与和国际合作等方面的一系列新的科学问题。面对这些新问题，人机交互可能需要重新审视本领域的研究范畴，例如衍生新的定义“人-社区交互”（human-community interaction）或“人-社会交互”（human-social interaction），或者孵化新的研究范式和方法框架。这也会导致传统人机交互的界面设计目标发生改变，并引发对新的更深层次的社会科学问题的思考和探索。

上海交通大学教授程时伟以“眼动交互与脑机交互”为题作引导报告，报告以基于眼动跟踪的跨设备多用户协同交互和基于运动想象的脑机交互为例，从多用户、多设备、多模态、多智能体协同交互的新视角，提出了降低感知、认知负荷和提高交互效率的新机制，强调了交互意图推理和交互路径优化的重要性。

新技术、新设备、新平台

清华大学教授徐迎庆作了题为“AIPC：基于AI智能体的个人移动终端”的引导发言，指出个人移动终端已经成为人们日常生活中不可或缺的一部分，而将其与AI智能体相结合，将会带来一系列的变革和发展。结合AI智能体的个人移动终端将极大地提升用户体验。通过人工智能的智能分析和学习，个人移动终端可以更好地理解用户的需求和行为习惯，从而提供更加个性化、精准的服务。AIPC将推动移动智能设备功能和性能的创新。人工智能技术的引入可以使移动终端更加智能化和灵活化，例如在图像识别、语音识别、自然语言处理等方面，能够为用户提供更加强大、高效的功能。这将促进移动终端硬件和软件技术的持续

进步,进而推动整个行业的发展。AIPC也会对社会生活和生产方式产生深远影响。利用个人移动终端的智能化和连接性,人们可以更加便捷地获取信息、沟通交流、管理生活,从而提高工作效率和生活品质。同时,人工智能技术的普及也将为社会带来新的就业机会和经济增长点,促进经济社会的持续发展。

吉林大学教授孙晓颖作了题为“多媒体终端触觉真实感交互”的引导发言,他首先介绍了触觉反馈对于多通道交互提升真实感体验的重要作用,以及触觉交互在关键领域的应用。特别以振动触觉为例,介绍了宽频线性马达的问世给振动触觉的快速发展带来了机遇。振动触觉可以使用户获得更加身临其境的体验,增强多媒体内容的沉浸感,提升交互的真实性和情感连接。除此之外,空气压膜技术、超声干涉技术等前沿的触觉反馈技术也在快速发展中,有望实现下一代终端的触觉真实感体验。

中国科学院软件研究所研究员韩腾作了题为“视触融合的具身交互技术与装置”的引导发言,指出空间计算和人机协同的发展加速了信息空间与物理世界的深度融合,人与智能信息空间的交互、混合将是未来社会的发展形态。智能信息空间的交互脱离了桌面式图形用户界面的范式,从二维(视觉、听觉)拓展到三维(视觉、听觉、触觉与本体觉)。触觉与本体觉参与的视听触融合呈现对交互的优化调控和认知干预是发展交互生理心理学模型和交互技术的重要突破口。

北京航空航天大学教授王党校作了题为“数字触膜:混合现实环境视触融合交互”的引导发言,从给手机、平板电脑、电动汽车屏幕等触屏设备提供触觉反馈的需求出发,提出了类似手机贴膜的“数字触膜”,即主动电子皮肤的概念,具体原理是通过柔性、超薄、透明的触觉致动器阵列,在屏幕上产生精细材质纹理和局部振动的触觉反馈效果,从而将触屏的“像素”升级为“触素”。该数字触膜不仅能够贴在屏幕表面,还有望贴在自然界的任意实体表面,例如桌面、墙面、人体皮肤表面等,从而为智能家居、智能汽车、VR-AR-MR的穿戴设

备等提供新型触觉反馈界面。他还介绍了通过微流体驱动控制技术和微纳加工制造工艺,实现高透明度的多点触觉驱动薄膜,有望实现屏幕的无视觉遮挡的触觉图案的呈现。

在讨论环节,嘉宾们一致认为AIPC是未来交互终端的形态。任向实提出,发展AIPC是大势所趋,但还未能看到其中的科学挑战,希望相关领域科研人员对此进行进一步提炼。CCF人机交互专委会首任主任、中国科学院软件研究所研究员戴国忠认同触觉交互重要且困难,但认为大家的讨论还不足以解答触觉交互的困境,即触觉交互缺少交互认知理论的支撑,需要从理论上先回答视触听等融合的机制和对交互的影响。同时他提出未来的交互设备一定是触觉装置,但尚未看到成熟的技术。

标准、规范,学术界、产业界与政府的举措

智能时代人机关系快速变革,机器承担的任务更加多元,人与机器的交互更加密切。新范式、新理论、新技术、新设备、新平台的研究与应用,将进一步推进人与智能设备、智能环境交互的常态化,交互更便捷、更高效,将同步改变人们的生活方式和工作方式。同时,人们要意识到人机交互过程必然伴随着机器对人类的逐步影响。在智能时代下,人机交互的机遇与挑战包括隐私保护、道德伦理等方面具体的问题,也包括技术发展和社会影响等宏观层面的思考与探讨,需要从标准、规范等方面寻求人机协作的平衡点,实现科技与人类发展的良性互动。

回首过去,随着人机交互范式、理论、技术、设备、平台的发展(如触摸屏、语音识别和脑机接口等)、应用的丰富(如办公、教育、医疗等),人们工作和生活的效率、便利性得到显著提升,带来了巨大的社会效益。同时,多个国家(中、德、英、法)都意识到了制定相关标准的重要性,确保了交互的质量,保障了用户权益和健康,促进了互操作性,使不同产品和系统能够相互连接和协同工作,推动了技术的发展和革新,也有助于促进国际贸易

和合作。

自2008年开始,德国标准化学会(DIN)先后通过了诸多人机交互相关的标准,例如软件个性化指南(ISO 9241-129-2010)、交互式系统用以人为设计(ISO 9241-210-2010)、电子可视显示器的用户性能试验方法(ISO 9241-304-2009)、软件可访问性指南(ISO 9241-171-2008)、交互式语音应答(IVR)应用程序(ISO 9241-154-2013)、电子直观显示系统的要求(ISO 9241-303:2008)。英国标准协会(BSI)先后通过了人机交互工效学·IVR应用程序(ISO 9241-154-2013)、人机交互的人类工效学·第410部分:物理输入设备·产品设计准则(ISO 9241-410-2008)、人机交互作用的人类工效学·物理输入装置的选择(ISO 9241-420-2011)、人机交互作用的人类工效学·物量投入设备的原则和要求(ISO 9241-400-2007)、系统交互作用的工效学·软件个性化导则(ISO 9241-129-2010)。法国标准化协会(AFNOR)也先后通过了人与系统交互作用的人机工程学·第910部分:触觉和触摸式交互作用的构架(X35-122-910-2011)、人-系统交互工效学:电子视觉显示器术语(X35-122-302-2009)。

2013年,我国成立了全国信息技术标准化委员会用户界面和人机交互技术委员会。2017年,中国科学院软件研究所、中国电子技术标准化研究院、小i机器人共同牵头制定的“信息技术-情感计算用户界面-模型”(ISO/IEC 30150-1:2022 Information technology-Affective computing user interface (AUI) -Part1: Model)国际标准正式发布。2020年,“信息技术 手势交互系统”(GB/Z38665-2020)、“智能制造 人机交互系统 语义库技术要求”(GB/Z 38623-2020)等国家标准发布。

面向未来,计算设备硬件感、显、算分离趋势越发明显,软件智能化程度越来越高,国内外各大公司,如苹果、谷歌、微软、英特尔、英伟达、华为、联想等,均在研发人机交互技术。因此,在软硬件智能化的背景下,推动人机交互标准的建立有利于规范市场和促进行业发展。

人机交互研究和实现适应人类感知-认知-行

为规律的交互式计算系统,人机交互技术是实现人机之间高效对话、协同与共进的桥梁,对人类利用数字生产资料提升生产力具有决定性作用。学术界对人机交互的研究内容涵盖范式、理论、技术、设备、应用平台等多个方面,产业界从应用层面出发,逐步优化计算设备的交互性能与交互体验。学术界和产业界的研究路径与方法各有侧重,因此要重视相互对接的重要性,也要重视科普的重要性,重视统一教材的必要性,增强学术研究与产业厂商之间的相互理解,增强大众对人机交互的认同,培养更多对人机交互领域感兴趣的青年学者。

我国人工智能发展拥有良好的基础,国务院于2017年印发《新一代人工智能发展规划》,提出了面向2030年我国新一代人工智能发展的指导思想、战略目标、重点任务和保障措施,部署构筑我国人工智能发展的先发优势。然而,随着人工智能的智能水平的发展与提升,其不足之处也逐步显现,与此同时,人类在数据处理、模型调优、解释能力、伦理监督等方面发挥的作用越发显著,人的经验、价值判断能力的作用在交互中不断被放大。美国国家科学技术委员会2022年发布的《关键与新兴技术国家战略》(National Strategy for Critical and Emerging Technology)将人机交互与人工智能、量子信息技术共同列为18项战略技术之一,同年,美国国防部发布《竞争时代国防部技术愿景》(Department of Defense Technology Vision for an Era of Competition),也将人机交互列为14项关键技术领域之一。智能时代给人机交互的进一步发展带来了机遇,建议我国科技部、国家自然科学基金委等部门在支持人机交互技术零散项目/课题的同时,酌情对人机交互领域进行全面布局,增强人机交互技术在我国人工智能发展战略中的重要性。

人机交互十大科学技术问题

戴国忠在特邀报告“二八反思:21世纪人机交互的科学与技术问题思考”中,从人机交互学科的发展历程谈起,指出21世纪人机交互发展的5个

理论问题和5个技术问题，进一步分享了人机交互视角下的大模型，并介绍了智能时代人机交互的新环境、新机遇与新挑战，最后他强调要发展人机交互，实现以人为本的信息技术。

围绕戴国忠的特邀报告和四个专题的报告内容，与会专家开展了深入的讨论，凝练出未来人机交互的10个科学技术问题（如图1所示），其中包括人的内在的（4个）：用户体验（自然性）的计算模型、多模态信息感知的融合机制、运动控制的最优策略、心理模型建模；人与智能体交互（2个）：人机协同机制中的任务表征和分配、人类思维链的计算模型；技术层面（4个）：面向情感价值供给的拟人化人机界面、交互意图的感知与理解、交互数字人构建技术、新型人机交互装置。

心理模型建模 人机交互本质上是人机合作关系，人与智能体的交互可以类比为人与人的交互，具有和人与人交互相似的属性和规律。智能体只有做到准确识别人类的心理状态，理解用户的情绪情感、需求和动机，以及能够解释、模拟和预测人类的行为，和谐自然的人机交互才有可能实现。为了实现这一目标，需要针对人机交互这一特定的情境，提出可以准确刻画人类与智能体交互的心理模型，并对这一心理模型进行计算建模。由此可见，未来

人机交互领域需要解决的一个关键科学问题是如何构建可以准确刻画人类与智能体交互的心理模型。构建这样的心理模型需要考虑以下几个问题：第一，未来高度智能化的计算机系统是具有与人相似认知功能的智能体，因此人机交互不是智能体被动地接受人类的指令，而是两个交互主体在各自先验知识和目标的驱动下主动交互的过程；第二，人类与智能体都需要包含信息处理的基本模块，实现多模态交互数据的获取和处理、加工和存储，以及输出和反馈功能；第三，计算机与人类的认知模块之间存在异质同构性，具有相同的功能模块；第四，人机交互是多模态并行、分布式交互的过程；第五，人机交互的心理模型需要考虑到人类作为生物体的非智能方面的生物属性，以及人类作为高度社会化的生物所具有的社会属性。

运动控制的最优策略 交互过程中，用户通常会基于感知到的当前状况，选择收益最大的策略行动，这个过程称为运动控制的最优策略。随着交互系统逐步应用在复杂的环境、特殊用户和上下文情境中，对运动控制过程准确建模并实现最优行为策略面临以下挑战。（1）交互情境和任务的计算性描述：用户运动控制指令的下达依赖于对当前交互情境的理解，包括交互过程的时空特征、状态变化、

认知负荷和决策复杂性等，交互情境和任务的计算性描述是求解运动控制最优策略的基础。（2）量化交互过程的不确定性：目标一致但各不相同是人类运动控制的重要特点，感知噪声、认知差异、运动扰动都会导致用户的运动表达出现不确定性，缺少对不确定性的描述则无法还原和优化用户运动，因此，需要从产生原因、过程和结果等方面对不确定性进行有效的描述和



图1 人机交互十个科学技术问题的框架

量化。(3) 最优化问题的求解：在充满不确定性的复杂交互过程中，用户面临的决策空间往往极为庞大，需要开发出能在多个约束条件下快速收敛的优化算法，以确保模型能够实时响应并生成最优运动控制策略。建立覆盖感知、认知、运动全阶段的用户仿真引擎是研究运动控制优化的重要基础和有效手段，通过用户仿真引擎能快速生成不同交互情境下的用户运动，节省大量用户实验时间，突破极端复杂条件的用户实现限制，为交互运动控制的最优策略研究提供条件。

多模态信息感知的融合机制 人机交互正在由传统的以鼠标/键盘为代表的2D用户界面向以VR/AR为代表的3D空间交互发展。其交互行为不仅仅依赖单一的视觉感官，还依靠更多感官的协同，包括视觉、听觉、触觉、前庭觉和本体觉等。这些感官感知到的多模态信息往往需要在大脑中进行整合，以形成统一、连贯的感知体验。使用认知心理学、神经科学、计算机科学等多学科交叉的研究方法，探究多模态信息感知的融合机制是构建人机交互的底层感知-认知模型的基础。主要的研究思路为，通过主、客观测量相结合的用户实验发现与界定多模态信息感知融合的全新问题，并结合多学科知识进行机理推测与建模。具体的研究方向主要包括探索多感官融合和协同的感知-认知机制，揭示多感官融合一致性的重要作用，研究多感官之间的信息互补效应和冗余信息的整合理论，探究基于跨模态刺激的干预策略，构建多模态信息感知-认知模型，等等。特别是，在复杂的现实环境或虚拟环境中（例如动作/运动重映射等），不同感官的输入可能存在冲突，引发认知困惑，破坏空间感知，带来生理不适感，并干扰动作控制精度。大脑如何处理、解决和适应这些冲突的感知信号尤其值得探索。研究多模态信息感知的融合机制具有重要的理论和实际应用意义，包括提供相关的知识来提升感知精度、增强环境适应能力、改善交互体验并指导感官干预方法设计。

用户体验（交互自然性）的计算模型 该模型旨在研究和优化用户在与系统或设备交互时的自然

性和流畅度。交互自然性是指用户能够将交互意图顺畅地映射为交互行为的过程，这一过程的核心在于决策的有效性和交互行为的经济性。该研究主要涉及三个关键问题。第一，用户在特定情境下表达交互意图时会选择一种其主观上认为信息充分的交互行为。这一选择过程与交互对象的智能程度密切相关。交互对象的智能程度越高，用户需要付出的交互成本就越低。在当前技术背景下，最高级的智能可以参考人类，因此，人机交互的自然性在许多情况下可以借鉴人际交互的标准来定义。第二，用户在选择交互方式时会估计其成本。交互成本可以细分为感知、认知和行为三方面。感知成本涉及用户获取信息的难易程度，认知成本涉及信息处理和决策的复杂性，而行为成本涉及执行交互操作的便捷性。探索基于信息或能量度量成本的方法，是构建自然交互模型的关键一步。第三，为了完成任务，用户会在感知、认知和行为三者之间进行组合和权衡。背后遵循的是最优化控制的基本原理，即用户会选择感知、认知和行为成本最低的交互方式。如何定义和求解这一最优问题，例如利用强化学习等方法，是值得深入研究的方向。总体目标是对用户认为自然的交互决策进行计算建模，提升人机交互的自然性和用户体验。通过研究用户在不同情境下的交互决策过程，可以为设计更智能、更自然的交互系统提供理论和实践依据。

人类思维链的计算模型 该模型旨在研究人类在人机物融合情境下的思维过程和决策机制。人解决问题的智能依赖所掌握的情境信息和任务目标，在不同的精度、时效、成本的要求下，采用不同的应对方法。本问题涉及以下三个关键方面。第一，理解人的思维方式的适应性。人类思维通常可以分为两种基本类型：直观的快速思维和分析性的慢速思维，但更多情况是二者的协同。研究应聚焦用户如何根据情境调整思维模式，以应对不同任务和挑战。第二，对不同情境下的决策方法建模。建立输入（情境）到输出（行动）的决策函数，需要综合优化准确性和计算成本，包括对任务的合理拆解、分步和组合，混合使用基于规则的知识系统和基于

数据驱动的深度模型，以模拟人所具备的不同形式的决策能力。第三，从真实情境中准确构建决策函数及知识库。这涉及从实际情境中的用户行为学习决策模式，也需要在人机交互（例如对话）中更高效地与用户交流和确认，以更快速准确地具体情境中与用户认知状态对齐。整体目标是揭示人类在不同情境下的思维机理，模拟人的情境决策智能，为实现在思维层面与人类协同的人机交互系统提供理论支持和技术基础。

人机协同机制中的任务表征和分配 在人机协同解决复杂问题的背景下，任务表征的建立旨在桥接人与机器的理解与沟通，确保人与机器能够共享并理解任务的目标和状态，从而实现可靠、可信与高效的任务协作。任务表征的建立涉及三个关键问题。第一，用户心理模型和机器内在处理逻辑之间存在鸿沟。一方面，普通用户的心理模型要求任务直观、易于理解，与用户的认知和期望相符；另一方面，为了能够被智能体准确理解和执行，任务模型需要具备可计算特征，涵盖任务的具体步骤和执行逻辑。构建用户可理解、机器可计算的任务模型，为人机互理解提供基础。第二，任务协同过程中人机角色不对等。人是任务的主体，具备创造力、判断力和决策力，负责提供任务目标、定义任务需求；机器则负责执行复杂的逻辑计算和重复性工作。人机协同过程是混合主动的交互式过程：用户改变需求时主动向机器提供信息；机器在缺乏信息时主动向用户发起询问。实现以人为中心、混合主动的任务表征，是实现人机高效沟通、机器协助人类解决复杂多变任务的关键。第三，任务协同知识可积累。任务模型的交互式生成不仅依赖智能体的通用智能，也得益于用户对任务知识的贡献；交互产生的任务模型为智能体进化提供大量数据。如何在协同过程中不断积累任务知识，并通过有效机制实现知识的泛化迁移，是重要的研究方向。总体而言，通过研究任务表征和分配，可以为人机协同完成复杂任务提供可靠、可信、高效的解决方案。

面向情感价值供给的拟人化人机界面 要实现这一目标，核心在于如何将多模态大模型与现有的

情感计算技术相结合，并在虚实融合的交互空间中实现自然且实时的互动。多模态大模型可以整合语音、图像和文本等数据，但真正的挑战在于对这些信息进行实时处理并应用于拟人化界面，使其能够快速响应用户的情感变化。同时，虚实融合的交互空间要求界面能够感知和理解现实环境，并在虚拟世界中创造出与现实无缝衔接的互动体验，从而提升用户的沉浸感，增强情感连接。为此，需要研究多模态情感计算框架的实时优化方法，使其在低延迟环境中实现高效、自然的情感响应；研究面向虚实融合的交互空间智能感知系统，以便更精确地理解用户所处的环境和情感状态；未来的情感计算系统需要实现自我学习和进化，以更好地适应不同用户的个性化需求；还需要研究在超写实数字人的构建过程中，如何对精细化表情、动作、皮肤、毛发进行实时渲染与交互，使其在不同物理互动和社交互动情境下依然保持高度真实感。

交互意图的感知与理解 人机交互输入技术的核心挑战是机器如何根据多模态交互行为数据准确推理用户的交互意图。然而计算机形态和使用情境日益复杂化，交互行为不再受限于特定表面与场景，从离散的确定的桌面二维表面扩展到连续的非结构空间，准确推理交互意图面临全新的挑战。（1）自然行为连续随机：交互场景复杂动态多变，脱离鼠标等规范接口，交互可能随时随地发生，用户表达的有意交互行为与非交互行为混淆性强，具有时空上的高随机性特点；（2）交互意图内隐高噪：受人体运动控制能力约束，人体表达的交互行为具有高随机噪声，外加传感器感知噪声，导致交互意图高度内隐于交互行为的感知数据中，增加了准确推理交互意图的困难；（3）感知数据离散稀疏：用于感知交互行为的传感器分布稀疏，感知带宽窄，无法有效重构用户丰富的交互行为，进一步影响了交互意图的推理。以上问题的核心是交互行为感知信息的高度不充分性，成为限制机器高效感知与理解用户意图的巨大技术挑战。

交互数字人构建技术 该技术旨在深度分析用户行为和环

知识, 创建能够模拟真实用户行为和反应的数字人模型。该研究主要涉及4个关键问题。第一, 确定数字人中的关键知识及其表征方式, 从底层行为数据中抽象出代表用户行为习惯的层级模型。该模型既要参考人的心理模型, 从认知层面理解用户, 还要支持各个场景下机器的推理与计算。第二, 明确数字人中关键知识的适用范围和迁移范围。由于数字人中的知识来源于不同任务背景下用户和智能体的交互历史, 因此需要确认哪些知识可以跨任务迁移, 或者提出多服务背景下的通用用户知识抽取算法。第三, 交互数字人可以实现对输入刺激的预测。面对不同的输入, 需要交互数字人提取相匹配的用户信息来辅助智能体推理, 预测人类用户的反应, 从而支持个性化交互。第四, 设计交互数字人与人类用户之间的交互方式, 例如数字人信息缺失时如何向用户主动询问, 如何向用户呈现数字人信息等。这是为了通过用户的介入, 增强交互数字人的信息准确性和全面性, 提升交互数字人的构建效率。

新型人机交互装置 人机交互装置承载了人机交互的物理介质, 是保证人机交互自然性和交互效果的基础。从主机计算时代的穿孔纸带, 到桌面式计算时代的键盘鼠标, 到移动计算时代的触摸屏, 再到空间计算时代的VR手柄, 交互装置不断推陈出新。未来空间计算时代, 各类新型交互装置例如触觉交互装置、嗅觉交互装置、脑机交互装置等, 必将成为人机交互的核心研究问题。目前的交互装置聚焦在视觉和听觉通道的沉浸式呈现和交互, 缺少对触觉通道的深度开发。触觉在人类与自然界的交互中扮演着基础的、不可缺少的作用。但是, 在人类同数字世界的交互中, 触觉交互依然处于初级阶段。亟待研究视-听-触-力多通道融合自然人机交互新范式, 使用户能通过触觉通道, 采用五指自然交互方式直观地操作数字化信息, 并获得信息内容的触觉和力觉反馈体验, 让计算机呈现的比特世界像原子世界一样真实可触。人工智能推动了材料技术的飞速发展, 利用材料的变形、变刚度、变温度等可编程属性赋能触觉人机交互, 研究基于柔性材料的触觉反馈手套、触觉外骨骼、触觉传感反

馈融合的主动电子皮肤等, 将实现更加逼真自然的触觉交互。这涉及如何设计和制备具有智能交互功能的可编程材料, 使其能够感知环境、与用户进行有效交互, 并根据交互情境做出相应的响应。

未来人机交互的5类刚需应用场景

结合人机交互技术发展前沿和符合中国国情的原则, 与会代表讨论并总结了未来人机交互的5类刚需应用场景。

1. 制造业: 做大做强先进制造业对推进现代化产业体系建设, 加快发展新质生产力具有重要作用。以混合现实、多模态融合、群智协同为代表的人机交互技术可以优化生产效率、提升智能制造的人智融合水平, 进一步激活制造业引擎。以美国为例, 美国国家科学基金会(NSF)“未来与赛博制造业”专项中就包括对工业物联网与可穿戴设备、混合现实与人机增强技术等战略性先进制造业中的应用研究与人才培养计划, 旨在保持未来美国制造业的竞争力。

2. 航空航天: 2018年国际空间探索协调组(ISC-ECG)发布了最新的全球探索路线图, 提出了“近地轨道-地月空间-载人探火”的深空探测发展思路。安全高效的人机联合工作对中国空间站建设与载人登月计划实施至关重要, 需要解决太空及月面环境因素影响、人的能力特性变化、人机交互模式和人机协作等一系列关键人因问题。在未来载人航天领域, 航天员-自动化/机器人协同将是主要模式。近地轨道空间机械臂遥操作、先进机器人在轨验证、星表探测机器人控制等任务场景为人机交互技术发展提供了丰富的真实世界问题。

3. 文化旅游: 中华文明的文化特质为我国人机交互发展提供了独特的人文土壤与思想源泉。在我国人机交互发展历程中, 以汉字激光照排、笔式人机交互为代表的技术创新具有里程碑意义, 构建了相关学科与学术体系, 为国际人机交互发展贡献了中国智慧。面向文化科技深度融合、提高国家文化软实力的发展需求, 文化装备与大众消费、旅游与

现代服务业的发展需要开发和应用新兴人机交互技术，构建沉浸式用户体验，拓展文旅融合新业态、新场景。

4. 医疗健康：面向人民生命健康，人机交互技术在医疗健康领域应用场景广阔，既包括专业医疗领域如智能医疗设备人因研究与产品设计，虚拟仿真技术在临床与培训领域的应用，人工智能辅助诊断中的医智协同，医患平台的社交沟通与情感交互等；也包括未来老龄化社会背景下包容性用户界面、适老化产品设计，以及涉及人民生理心理健康和人文关怀的各种公共性信息服务系统的用户体验设计。

5. 教育：从多媒体教学、远程教育到虚拟现实课堂，人机交互技术应用对教育领域的发展始终具有重要作用。面向未来，人工智能与K-12、职业教育、高等教育等各层面的深度融合将改变知识获取与创造力培养方式。新的“教”与“学”关系为人机交互提出新问题、提供新场景。同时，视听触融合技术、人智共协计算以及下一代自然交互范式也将进一步助力教育变革与发展。

总结

过去60年，人机交互和人工智能的发展呈现出此起彼伏的交替进步趋势。近几年，以ChatGPT为代表的大模型开启了人类智能和机器智能共同发展的新阶段，也开启了人机交互的新时代。人工智能是人类智能的延伸，人机交互是构建未来人机混合智能的必要环节。智能必须通过与用户或其周围环境的交互进行迭代学习，没有交互就没有智能。为实现人机协同共进，人机交互领域应加强模型、范式、平台三大方面的基础性研究。

首先，在模型层面，需要从信息论的角度构建多通道人机交互的信道容量模型，扩展经典的用户信息处理模型（HIP模型），输出信道（motor channel）从手指扩展到全身动作输出乃至生理信号和脑机接口，输入信道（sensory channel）从视听扩展为视听触嗅味多感官融合，从而适应新硬件、新

计算环境的变化。其次，在范式层面，需要研究后WIMP（窗口、图表、菜单、指针）范式时代和后触屏时代，适应VR、AR、MR、机器人等新硬件和泛在、普适、沉浸式计算环境的新型多通道自然人机交互范式。再次，在平台层面，亟须研发适应新范式的自然人机交互硬件平台，例如拟人手写笔、虚实融合环境智能表面交互的数字触摸屏、手势眼动肌电多通道交互装置等；适应新计算环境的软件平台，例如XR操作系统、普适计算泛在操作系统等。 ■

整理：王党校 喻纯 韩腾 王运涛 王冠云
田丰

附：与会专家名单

论坛嘉宾：任向实 宋爱国 王茜莺 王涌天

参会嘉宾（按姓氏拼音排序）：

班晓娟 程时伟 黄进 黄莹菁 刘子韬 任磊
孙伟 孙晓颖 王冠云 王琦 王巍 翁冬冬
武汇岳 谢良 徐迎庆 薛建儒 喻纯 张腾翔
朱超

会议发起人：戴国忠

会议组织者：田丰 王党校

会议记录人：韩腾 王运涛

（本文责任编辑：杨易）