



北京信息科学与技术国家研究中心

Beijing National Research Center for Information Science and Technology

简 报

办公室编印

2023 年 4 月刊

2023 年 4 月 30 日

本期导读

- 第三届清华信息前沿交叉论坛举行
- 北京信息科学与技术国家研究中心系列交叉论坛（第五十八期）举办
- 北京信息科学与技术国家研究中心系列交叉论坛（第五十九期）举办
- 国家重点研发计划“网络信息论的关键数学基础”项目启动暨实施方案论证会召开
- 清华大学信息国家研究中心在神经网络解释方法领域取得进展

◆ 焦点要闻

第三届清华信息前沿交叉论坛举行

4 月 23 日，值清华大学 112 周年校庆之际，由清华大学信息科学技术学院和北京信息科学与技术国家研究中心主办的第三届清华信息前沿交叉论坛在清华大学信息楼（FIT 楼）多功能厅举行。清华大学副校长、中国科学院院士姜培学，信息学院院长和信息国家研究中心主任、中国工程院院士戴琼海等出席论坛。论坛由信息国家研究中心副主任、中国工程院院士罗毅主持，校内师生 130 余人参加论坛。



姜培学致辞

姜培学发表致辞，对论坛的举办表示祝贺，肯定了信息国家研究中心的定位、交叉创新群体的建设举措和已经取得的一系列重要科研成果。姜培学指出，信息国家研究中心是清华国重体系建设的重要一员，以本论坛为代表的多维度学术活动进一步推动了多学科、校内外的交流与合作。他鼓励中心继续发挥学科交叉“催



化剂”作用，加速带动产学研融合发展，以创新为引领、以改革激发内在活力，坚持“顶天、立地、树人”的科研宗旨，更好地服务国家重大战略需求。

中国科学院院士王怀民以“群智范式：软件开发范式的新变革”为题作报告，指出当今时代软件已经成为信息基础设施的重要组成部分，并从“工程范式”“开源范式”和“群智范式”三个方面阐述了软件开发范式的发展进程及特点，将自然科学与哲学高度融合，概括出“软件开发范式的变革首先是世界观的变革”。

中国工程院院士苏东林以“电磁兼容传导敏感性时域测试系统”为题作报告，从解决电磁兼容问题出发，阐述了研制时域电磁敏感性科学仪器面临的关键科学问题、重要科学发现、成果推广应用等，强调了该研究领域需要瞄准的国家需求和卡脖子难题。

清华大学教授朱文武作题为“跨空间大数据关联表征学习理论与方法”的报告，阐述了跨空间大数据关联表征学习项目的研究背景、科学问题、研究思路及重要科学发现，介绍了研究产生的国际影响，以及面临的新挑战。

清华大学教授黄翊东作题为“光电子微纳结构关键制备工艺的研究”的报告，介绍了团队多年来在微纳结构光电子芯片的制备工艺方面的研究历程，以及微纳光电子制备平台的搭建，指出要以核心关键制备工艺的突破，推动光电子芯片领域的创新研究及产业化进程。

在院士圆桌论坛环节，针对“如何同时实现探索国际前沿和面向国家重大需求的科研”“如何更好地从事学科交叉研究”等问题，院士们分享了自己的独到见解。

中国工程院院士赵沁平结合多年工作经历分享了自己对于“学科”的理解，指出信息领域具有学科跨度大、交叉多的特点，鼓励青年学者关注原始创新，多阅读、多交流和多实践，挖掘真正的科学问题。中国工程院院士刘永坚认为，探索国际前沿与解决国家重大需求之间没有矛盾，无论从事基础研究还是应用研究，都需要瞄准国家重大需求，不断凝炼和识别关键技术及科学问题。王怀民认为，随着时代的发展，当代中国的科技需求已经与国际前沿发展同步，年轻人要积极面向国家重大需求解决国际前沿问题。苏东林分享了自己打造多学科、紧耦合团队以及创办电磁学领域国际期刊的经历，她强调，有国家需求牵引，从制度和组织上发挥大团队内多学科人员各自优势，学科间长久相互助益、协同创新是解决复杂科学问题、完成重大科学任务的关键。罗毅以我国 LED 技术突破应用于半导体照明为例，提出做科学研究应该始终以需求为导向，要坚持从需求中凝练重大科学问题，鼓励青年学者进主流，解决真问题，“把文章写在祖国大地上”。

戴琼海作总结讲话。戴琼海表示，本次论坛对科研工作与重大需求的深入讨论既可以上升到哲学思想的感悟，同时又对实际工作具有重要的指导意义，希望



与会师生深刻理解，把参加论坛的收获更好地运用到实际科研工作中。

清华信息前沿交叉论坛是清华大学信息科学技术学院和北京信息科学与技术国家研究中心每年在清华大学校庆期间举办的重要活动之一，首届论坛于 2021 年清华大学 110 周年校庆期间举行。每年同期在信息楼举行“清华信息创新成果展”，汇集展示清华信息学科最新的科研创新成果。

北京信息科学与技术国家研究中心系列交叉论坛（第五十八期）举办

4 月 13 日晚，北京信息科学与技术国家研究中心系列交叉论坛（第五十八期）通过线上会议和直播的形式举办，本次论坛邀请了香港科技大学首席副校长、英国皇家工程院院士、欧洲科学院院士郭毅可作题为“人有力的用处——拥抱人机共生同创的新时代”的报告。



郭毅可作报告

论坛由清华大学信息学院院长、信息国家研究中心主任戴琼海院士和信息国家研究中心副主任丁贵广教授共同主持。信息国家研究中心扩大会成员、团队负责人以及校内外师生 370 余人通过腾讯会议在线参加论坛，累计约 34 万人次通过上直播、新浪、百度等直播平台在线观看。

郭毅可在报告中提出了“拥抱人机共生同创”的理念，他认为随着人工智能技术的发展和普及，机器可以辅助人们解决很多问题。同时，人类的创造力和智慧也是机器所无法替代的，因此人与机器的合作可以实现更多的创新和进步。报告最后，郭毅可呼吁人们要拥抱人机共生的新时代，并积极参与其中，共同创造更美好的未来。

问答环节，郭毅可的精彩报告也引来了大家的思考和热烈讨论。例如：大模型背景下是否会有颠覆性技术的存在？未来嵌入认知推理和约束的语言模型会不会是自然语言的一个发展方向，以及未来的发展思路？对马斯克等千人呼吁暂停 AI 过度开发的观点？郭毅可针对与会人员的提问一一解答，大家就各领域的研究内容展开了深入讨论与交流。

北京信息科学与技术国家研究中心系列交叉论坛（第五十九期）举办

4 月 27 日晚，北京信息科学与技术国家研究中心系列交叉论坛（第五十九期）通过线上会议和直播的形式举办，本次论坛邀请了加拿大工程院院士、加拿大英属哥伦比亚大学电子与计算机工程系终身教授王真（Z. Jane Wang）作主题报告。论坛由清华大学信息学院院长、信息国家研究中心主任戴琼海院士和信息

学院副院长任天令教授共同主持。信息国家研究中心扩大会成员、团队负责人以及校内外师生 90 余人通过腾讯会议在线参加论坛,累计约 26 万人次通过上直播、新浪、百度等直播平台在线观看。

人工智能(AI),尤其是深度学习模型,在许多领域的突破和革命为我们带来了巨大的机遇和挑战。然而,由于当前深度学习模型的“黑盒”本质,缺乏可解释性;数字媒体数据和深度学习模型容易受到操纵和攻击的影响,使得我们很难信任数字媒体(例如来源、完整性



王真(Z. Jane Wang)作报告

真实性),这是生物医学和数字媒体安全应用面临的严峻挑战。因此,我们建议探索传统信号/图像处理(SP/IP)和深度学习之间的交叉点,通过利用领域知识和深度学习的学习能力来缓解传统 SP/IP 和黑箱深度学习方法的缺陷。报告中,王真首先提供了一个概述,然后重点介绍了在生物医学和媒体安全和取证领域的研究。最后,她以集思广益的方式总结了未来的研究方向。

提问交流环节,王真同与会人员就如何降低大模型算力开销,以及人工智能和深度学习能否在医疗健康方面提供更好的支持等问题展开了深入讨论与交流。

◆ 科学研究

国家重点研发计划“网络信息论的关键数学基础”项目启动暨实施方案论证会召开

4月8日,由清华大学承担的“十四五”国家重点研发计划“数学和应用研究”重点专项“网络信息论的关键数学基础”项目启动暨实施方案论证会在北京西郊宾馆召开。科技部高技术中心基础一处处长朱庆平、清华大学科研院科研项目部主任朱付元、清华大学信息国家研究中心副主任朱文武教授出席会议并讲话。项目评审专家及项目组成员共 40 余人参加会议。

启动会上,朱付元代表项目承担单位清华大学致辞,对科技部和各位专家表示衷心感谢,并对后续工作开展提出了组织实施、项目管理方面的指导意见,期待项目的顺利推进并圆满结题。

朱文武代表信息国家研究中心致辞,对以殷柳国研究员为代表的项目组成员表示祝贺,并表示信息国家研究中心高度重视科技部重点研发计划项目,将会全力配合项目的有序实施。

朱庆平代表科技部高技术中心对项目启动表示祝贺,指出该项目的顺利立项

与实施，对于推动我国信息论与编码领域的研究具有重要意义，希望项目组围绕项目目标积极探索，做出原创性成果。

本次实施方案论证会邀请了西安电子科技大学白宝明教授、天津大学宋占杰教授、西南交通大学唐小虎教授、北京交通大学艾渤教授、北京航空航天大学张筱教授、中国科学院计算技术研究所周一青研究员、清华大学葛宁教授、北京邮电大学牛凯教授、北京航空航天大学黄勤教授等国内数学、信息论和通信领域的 9 位知名专家到会指导，成立了项目实施方案评审专家组

会上，项目负责人及各分课题负责人分别就项目及课题实施方案进行汇报。在听取汇报后，项目专家组对项目和课题实施方案进行了充分质询，并就项目研究难点、课题间关联性、项目标志性成果及测试验证方案等进行了深入交流，提出了建设性意见。经讨论，专家组认为项目及课题实施方案目标清晰，研究内容全面、技术路线具有创新性，任务分工明确、计划进度安排合理，技术指标考核方式科学合理，一致同意通过实施方案论证。

该项目由清华大学、上海交通大学、西南交通大学、南开大学、香港中文大学（深圳）共同承担。本次会议的召开标志着项目研究工作全面启动，项目组将聚焦通信网络技术发展面临的困境，探索从新的视角构建面向信息闭环的无线网络通信理论模型，力争实现理论方法源头创新。

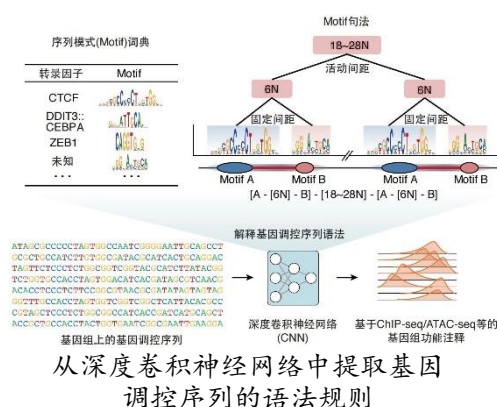
◆ 重点成果介绍

清华大学信息国家研究中心在神经网络解释方法领域取得进展

清华大学信息国家研究中心数基生命系统交叉创新群体成员、自动化系汪小我教授与美国斯坦福大学统计系王永雄教授团队合作，开发了神经网络解释算法 NeuronMotif，实现了从神经元中自动归纳和提取基因调控序列编码规则。研究成果以“NeuronMotif: 通过深度神经网络的逐层解耦破译基因顺式调控编码

(NeuronMotif: Deciphering cis-regulatory codes by layer-wise demixing of deep neural networks)”为题，于 2023 年 4 月 6 日发表于《美国科学院院刊》(PNAS) 期刊上。

近年来，大数据驱动的深度神经网络模型已在多个领域取得了重要进展。例如，在蛋白质结构和基因调控功能预测等方面实现了高精度的预测。然而，这些



模型通常是不透明的黑盒，难以理解。透明化神经网络黑盒可以帮助我们获取模型从数据中学习到的知识，从“只知其然”转变为“知其所以然”，对于促进神经网络理论发展和增进相关领域的认知具有重要意义。例如，基因调控序列是基因组上能够调控基因表达的 DNA 序列，决定了细胞的功能和状态；当前，卷积神经网络（CNN）等深度神经网络模型在基因调控序列的功能预测问题上能够取得很好的预测效果，但由于神经网络解释方法的缺失，使得研究者难以从这些深度模型中发现基因调控序列的深层编码规则，严重制约了基因调控规律解析和重构的研究进程。

研究团队发现，CNN 解释困难的一个主要原因在于深层神经元大多都是“多面神经元”（multifaceted neuron），这种神经元能够同时被多种不同序列模式激活，直接对这些混合模式进行可视化往往只能获得难以被人理解的结果，而 CNN 中的最大池化结构是导致深层神经元识别多种序列模式的关键原因。针对多面神经元问题及其形成机制，研究团队提出了 NeuronMotif 方法，该方法首先基于蒙特卡洛采样和遗传算法得到大量能充分激活神经元的序列集合，然后计算这些序列在神经网络各层的特征图，并通过反向逐层聚类的方式分离不同序列模式所对应的特征图，将具有不同模式的序列划分到不同的子集中，最终通过对每个子集的可视化获得易于理解的序列模式特征。在此基础上，该方法构建了基于结构化语法树的自动化知识提取方法，从数据中自动归纳转录因子结合位点序列模式、组合模式、间距、次序等调控序列语法规则。

NeuronMotif 归纳提取的调控序列语法规则能够得到文献和多种生物组学数据的支撑和验证。通过对人类基因组数据的学习，NeuronMotif 发现了数百种序列模式，与数据库中的已知基因调控序列模式高度匹配；通过 NeuronMotif 获取的深层神经元语法规则与多种细胞/组织中的染色质开放区域模式高度匹配，并得到了基因表达数据的支持；序列模式的组合和排布等规律在跨物种基因组序列上具有显著的序列保守性。

NeuronMotif 为我们提供了一种解读深度神经网络模型中深层神经元识别模式的新方法。其解释结果还可用于人工神经网络的诊断和改进，有助于降低神经网络调参的困难。借助该方法，我们可以通过神经网络的训练和解释，从数据中获取可供人类专家理解的知识，帮助我们更加深入地理解胚胎发育、疾病发生等生物过程中的基因调控规律，并为基因治疗等应用中定制化逆向构造人工基因调控序列提供支撑。

清华大学自动化系博士后魏征为本文的第一作者，清华大学汪小我教授和美国斯坦福大学王永雄教授为本文的共同通讯作者。清华大学李衍达教授、张学工教授、江瑞教授、魏磊助理研究员、花奎博士，斯坦福大学博士后马士宁也对本



文做出了重要贡献。该研究得到了国家自然科学基金、国家重点研发计划、北京信息科学与技术国家研究中心交叉创新群体、清华大学国强研究院、北京智源人工智能研究院等的资助。

报：清华大学党政领导、信息国家研究中心建设运行管理委员会成员、信息国家研究中心学术委员会成员、信息学院党政联席会成员、信息国家研究中心党政联席会成员

送：相关院系、部处负责人

发：信息国家研究中心各部门负责人

编辑：李琳

审核：丁贵广

联系电话：62792099

E-mail: bnrist@tsinghua.edu.cn