

中国科学院学部 科学与技术前沿论坛简报 第 100 次

学部工作局学术与文化处 编报
《中国科学》杂志社

2020 年 4 月 26 日

“语言智能技术未来发展及应用”论坛综述

一、国内外研究现状

信息技术的高速发展已经深入社会生活，不断影响社会经济发展的方向和人们的生活方式。语言智能技术是人工智能领域的重要研究方向，近年来得到了前所未有的重视，取得了令人瞩目的成就。语言智能技术的研究不仅涉及到计算机科学中对语音信号、语言信息和字符信息处理能力的探索，还包括对人脑语音感知、语言认知机理，语言自身的结构、现象、运用规律和演变过程，以及不同语言之间的语义关系等诸多方面，是一个涉及认知科学、信息科学、语言学、数学与逻辑学、心理学等的交叉学科领域，从事这项研究不仅具有重要的理论意义，而且具有巨大的实用价值。

2009 年，继 G. E. Hinton（辛顿）等人将多层神经网络方法成功应用于图象识别，微软实现了基于多层神经网络的语音识别系统，并使识别错误率大幅度下降，深度学习方法得到大规模应用。2014 年，纽约大学 Kyunghyun Cho 和蒙特利尔大学 Yoshua Bengio 等人提出了基于注意机制的编码器-解码器（encoder-decoder）基本框架，对神经

网络结构创新和二次开发，建立了基于神经网络的机器翻译系统，简称神经机器翻译（neural network based MT, NMT）系统。在此基础上，谷歌于 2017 年提出了完全基于注意机制的 Transformer 模型，国内众多公司通过跟踪、完善，实现了自己的神经翻译引擎，为普通用户提供机器翻译服务。语言智能技术的相关产品也已经进入市场，得到积极的反馈。

当前人工智能蓬勃发展，语言智能技术作为人工智能领域的重要研究方向，关系着如何让计算机拥有人类语言智能的问题，是人工智能不断深入、推动人工智能技术进步的关键所在。近 10 年来，随着信息技术的发展，社会信息化水平不断提高，特别是大数据、云计算以及各种智能终端的普及，为语言智能技术提供了良好的应用环境，机器学习、深度学习为语言智能技术带来了新的活力，迫切需要语言智能技术产生新的突破，为国民经济、国家安全和人民生活源源不断地提供创新原动力。另外，随着国家“一带一路”倡议的实施，对语言智能技术的需求进一步增强。如何抓住人工智能发展的机遇，积极融合大数据、云计算等信息领域的创新研究成果，使语言智能技术迈上新台阶，已成为我国乃至全球亟待解决的重大课题。

二、论坛概况

2019 年 11 月 15 日，“语言智能技术未来发展及应用”科学与技术前沿论坛暨中国科学院庆祝建院 70 周年系列学术论坛在北京举办。院士专家们围绕适应国内外对语言智能技术发展、应用的巨大需求等方面展开了广泛而深入的讨论，针对现阶段该研究领域所面临的科学和技术挑战与对策进行了探讨和展望。

本次论坛由中国科学院学部主办，中国科学院信息技术科学部、中国科学院学部学术与出版工作委员会承办，中国科学院声学研究所、中国科学院语言声学与内容理解重点实验室、中国科学院先进水下信息技术重点实验室和《中国科学》杂志社协办。中国科学院院士李启

虎先生担任执行主席，9位院士专家进行了报告演讲，来自清华大学、北京大学、北京师范大学和中国科学院等研究机构、院校及企业的130多位专家学者和师生受邀参加了论坛。

三、报告内容

（一）许嘉璐：重新认识语言文字研究的价值，助力人工智能技术的发展

AI离不开自然语言的自动化处理。自然语言处理离不开传统“小学”，尤其是其中的音韵、训诂之学。民族的语言和文字不仅是表达思想和人际交流的工具，而且是人（民族）的思维方式、把握主客观及其关系的内心实在。HNC（概念层次网络）理论的创建、实践、发展在相当程度上闪烁着中华民族的智慧之光，与西方主流计算语言学学派有着巨大的区别。

现代汉语是华夏古代语言的延续和发展，语言、文字和文化的DNA就蕴涵于现代人说和写的语言和文字之中；察觉之，体认之，运用之，虽然不同学者、不同团队会有视角、方法、技术之异，但却不能无视它、回避它、否认它，否则将捕捉不到中国语言文字的脉搏、神经和感情，难以适合本民族对“人工智能”的需求。

报告举例指出，汉字形、音、义三者之间或并非等边三角形关系，更接近事实的可能是“形”“音”为横（相对稳定）竖（相对易变）二轴，其交叉枢点则为“义”。如此管窥，似可触摸到中华民族思维方式若干特质：形象性、关联性、延展性、系统性、创造性、情理性。语言智能处理的理论、模型、技术，如果更加注意汉语汉字音与义的相关度及其来源与嬗变，或可在“深度语义分析”和人机“自由交互”方面更上一层楼。

（二）李启虎：语言智能技术未来发展及应用的若干思考

语言是思维的物质外壳，也是人们传递信息、思想、情感最自然、最直接、最方便的工具。计算机为处理语言信息提供了技术基础，推

动了语言智能技术的形成和发展。在人工智能概念提出时，自然语言理解也是核心研究内容之一。语言智能技术期望让计算机拥有人类的语言智能，涉及从表层的语音到深层的内容处理。自然语言本身的特点决定了语言智能技术面临巨大的挑战，需要寻找有效方法以获取口语和文本中的真实内容，建立口语、书面语和有效的语义表达之间精确映射的机制，并形成具有学习能力的知识系统，支撑各种语言信息处理的需要。

近年来，语言智能技术的相关产品得到了市场的积极反馈，根据 Gartner 发布的《2018 世界人工智能产业发展蓝皮书》，到 2021 年全球自然语言处理市场的价值预计会达到 160 亿美元。同时，语言智能技术发明专利申请也进入快速发展期，年申请量已经突破 5000 件。中国专利申请表现抢眼，已经成为语言智能技术领域全球排名第一的专利技术来源国和目标国。

报告结合国内外近来语言智能技术和相关人工智能技术发展和应用的情况，概括了语言智能技术发展的前景，提出了若干战略性的应对措施。首先，语言智能技术具有强烈的应用属性，应用场景广泛，也在应用中不断发展完善，应该把语言智能技术的发展融入人工智能的发展大潮之中。其次，知识是智能的基础，对于发展真正的智能必不可少。语言数据资源则是语言计算的基本要素，因此，要立足于自然语言处理的特点，重视语言知识库和语言数据资源建设，服务知识计算与推理，深化技术发展。第三，以国家需求为动力，选取若干语言智能技术应用领域突破关键技术，服务于“一带一路”、维稳/反恐、提高人民生活质量等。最后，需要重视跨学科融合借鉴，触发技术创新思路，积极融合大数据、云计算等信息领域的热门研究成果，使语言智能技术迈上新台阶。

（三）吾守尔·斯拉木：丝绸之路经济带多语言互译平台

丝绸之路经济带沿线国家语言，除了汉语、阿拉伯语外，还涉及

阿尔泰语系约 35 种语言，包括维吾尔语、哈萨克语、柯尔克孜语、土耳其语等，以及印欧语系印度伊朗语族约 60 种，包括塔吉克语、普什图语、波斯语、乌尔都语等。因此，丝绸之路多语言互译平台（以下简称“平台”）有着重要的意义。第一，语言通是民心通、政策通的基础。只有不同国家、不同民族的人们顺畅沟通、相互理解，才能够更好地进行文化、科技、医疗、教育、旅游等方面的交流与合作。第二，是贸易畅通和资金融通的基础。平台将促进中亚、南亚、西亚及整个丝绸之路经济带电子商务、互联网+的应用，加快我国优势电子商务、网络金融、现代物流等企业向“一带一路”拓展。第三，数字文化创作、网络文化传播呼唤平台。实现海量数字文化创意作品的智能机器翻译，将大大提高不同语言文字数字文化作品相互翻译的效率和效益，消除文化传播中的语言障碍，促进丝绸之路经济带沿线国家优秀文化资源的传播与共享。第四，国家稳定安全迫切需要平台。在现代通信工具和互联网上的宗教极端和民族分裂宣传渗透，以及煽动组织暴恐袭击事件的信息等主要使用三大语系（阿尔泰语系、印度伊朗语族、阿拉伯语）的语言文字，极少使用汉语。

平台研究需要大规模采集并构建中亚、南亚、西亚主要国家及民族的自然语言语料库，研究相似语言机器翻译方法，形态复杂语言机器翻译、资源匮乏语言机器翻译、领域自适应机器翻译、枢轴语言机器翻译等理论与方法，重点突破汉语-阿尔泰语系、汉语-印度伊朗语族、汉语-阿拉伯语智能机器翻译核心技术，并应用云计算、大数据技术等。

报告呼吁国内外高校、研究机构和企业加强在人工智能领域的合作，共同努力突破认知计算、混合智能的关键技术，构建新的混合智能平台和智能机器，为推进“一带一路”建设提供有力支撑。

（四）宗成庆：语义表示与解码

神经网络方法在自然语言处理（NLP）中得到了成功应用，NLP

系统的性能得到了大幅度提高，但实现从“处理”到“理解”的跨越，建立真正基于理解的高水平 NLP 系统，仍是该领域面临的核心问题。目前已经有自然语言表示和解码相关的几项研究成果，有助于推进自然语言认知计算的研究。

当前，包括深度学习在内的机器学习方法已经在 NLP 和机器翻译中发挥了重要作用，但是这些方法采取的是一种“赌博”思维，往往对于大概率事件有效，难以做到“举一反三”，从自然语言理解的角度，其性能表现不如三岁孩子，尤其缺乏合理的解释。建立人机交互、人机互助、机器自主进化学习和知识引导下的可解释性 NLP 系统成为未来发展的必然。面向特定应用领域和实际需求构建知识库、语料库等基础资源，充分利用机器学习等最新技术手段是 NLP 系统成败的关键。文本、语音、图像和视频等多源信息综合利用，构建集语音识别、说话人识别、视频分析和图像理解、NLP 与机器翻译等技术于一体的多模态信息处理系统，将成为未来技术研发的必然趋势。

近年来，脑科学研究发现，语义加工在大脑左右半球基本上是对称的，打破了传统研究中认为语言加工左半球偏侧化的情况；不同个体大脑的整个语义网络“看起来特别相似”。

因此，采用表征相似性分析方法和大脑的成分语义表征解释向量表示，已经形成了一种基于联想多通道自编码器的多模态词汇表示方法。该方法在堆叠自编码器的基础上设计了三个模块，分别是：（1）模态联想模块：用于学习不同模态信息之间的关系，从而实现信息的传递。由于在具体实现时只有少部分词汇具有对应的视觉或听觉信息，因此使每个词汇都具有多模态信息，学习模态之间的映射关系是非常必要的；（2）词汇联想模块，用于学习词汇之间的相关性。已有的词汇表示模型没有很好地编码词汇之间的相关性，因此可利用预测相关词汇的目标函数学习词汇之间的相关性；（3）动态门机制模块，用于针对不同类型的词汇对其不同的模态信息学习其重要性的权重。使用

前馈神经网络建模权重计算过程，使用加权拼接的方式得到最终的多模态词汇表示。这可能是一种寻求语义表示和解码的方法，至少提供了一种探索的方法。

（五）周明：机器推理：综合知识和预训练模型更好地理解语言

自然语言处理进入了以深度学习为主流的时代，主要的技术包括词的表示、句子的表示、编码-解码机制、Transformer 等，以及最近流行的预训练模型，形成了基于预训练模型加微调的新框架。但是深度学习仍然面临过分依赖大规模标注数据、不可解释、处理知识和常识的能力薄弱、推理能力欠缺等缺陷，这些领域都期待着更多的研究。报告重点讲述了推理问题，探讨了利用知识和预训练模型进行推理，并利用语义分析、FEVER、CommonsenseQA、GQA 等说明推理的机制，希望引起对推理研究的重视。

报告指出，未来在机器推理方面需要建立知识丰富、可解释和可推断的模型，引入明确的知识作为必不可少的组成部分，把知识和数据结合使用。呼吁建立面向机器推理专用的测试数据集，目前常用的 FEVER, CommonsenseQA 和 GQA 等并不适合对机器推理任务进行评测，不能说明推理正确与否的原因。同时机器推理还应该支持各种现实的需要，例如事实检查和虚假新闻检测等。最后，在机器推理方面，还应当尝试基于其他一些模型的方法，如 GNN、贝叶斯网络和 MLN。

（六）胡国平：从能听会说到能理解会思考

计算智能与感知智能方面已经取得了极大的突破，形成了千亿级市场规模，但在强调知识、推理能力的认知智能方面，机器与人类仍有差距。语言技术是认知智能的重要技术基石，但即便是自然语言处理技术在特定领域已经超过人类，人工智能在认知智能层面依旧面临艰巨挑战。在不断积累拓新的基础上，通过对语言之下智慧本质的思考以及扎扎实实地在行业打磨，科大讯飞努力让人工智能从“能听会说”向“能理解、会思考”进一步迈进。

认知智能的几大研究热点包括：各个行业的认知智能需要独立研发的；基于深度学习的语言智能还可以持续挖潜；人类已有知识的有效利用是研究热点；“半”可解释性语言认知智能将热起来；从“相关关系”到“因果关系”；人机耦合——语言智能现阶段落地的重要路径。

“人工智能+”的时代已切实到来。全球人工智能产业已经进入一个关键的窗口期，谁能够把人工智能应用真正形成规模、应用落地，谁就能在未来智能产业中占领先机。未来得语言智能者得天下！

（七）鄢志杰：阿里语音 AI 的研究与实践

随着人工智能、特别是感知智能在技术上的突破式进展，语音 AI 已经走下科幻电影，成为大众科技。阿里巴巴达摩院语音实验室通过“阿里语音 AI”向集团内的各部门及阿里云上的商业客户提供语音核心技术产品，“阿里语音 AI”已经形成了语音 AI 原子能力和语音 AI 解决方案两种服务模式，取得良好的应用反馈。同时，“阿里语音 AI”最新的技术还会“零时差上云”，使得阿里云客户也可以基于这些能力在自己的行业中构建出具备语音交互能力的产品。

“阿里语音 AI”已经成功形成“浙江电力虚拟调度员”，是全国首个电力虚拟调度员，语音识别、语义理解、知识图谱、智能决策全面结合，提升客户工作效率 4 倍，减少 80% 重复工作，减少生产运维支出上亿元。

（八）吴玺宏：语音识别和对话技术面临的挑战及 PKU 测试研究进展

近年来，随着深度神经网络理论和技术的发展，语音识别的性能取得了长足的进步，基于语音、语言技术的人机对话技术已经成为热点应用。针对语音识别及人机对话技术中存在的问题及面临的挑战，进行了分析和阐释，指出上述领域研究和应用的真正突破在于回归智能理论的本源，在于理论和方法的创新。在 Turing Test 利用语言的交流能力来评价智能水平的基础上，提出了北京大学 PKU Test 研究计

划。这一计划从多方面对智能水平进行评价，具体包括自主动作控制，视听知觉，音系构建和识别，以及语言理解和对话。在语言理解和对话方面，需要基于世界知识和语言知识，在语境约束下实现对消息含义、情感、意图和动机的推测。

（九）颜永红：智能时代 前进中的语言声学

语言声学是研究人类发音听觉相关的学科，随着计算机的到来，以语音识别为代表的语言声学和信息技术日益密不可分。近年来，谷歌推出了语音搜索，苹果公司推出了 Siri，亚马逊推出了 echo，使得语音识别技术再次受到高度关注，掀起了新一轮的产业热潮。可以预见在未来几年内，语音识别技术将得到更广泛的应用。同时需要关注，随着以大数据为特征的智能时代到来，语音交互技术在很多方面得到快速发展，包括端对端的建模技术，边缘计算以及专有推理加速器件的研究快速崛起，复杂环境下语音识别技术研究迭代创新，多模态语音识别技术崭露头角；实现接近人声的语音合成，结合发音特征的神经网络声码器加速落地，个人定制化语音合成。同时语音技术也有很好的应用前景，包括音乐自动分析处理，利用体音检查身体健康，以及工业物联网中的机械故障检测等。

智能时代，以语音为核心的人机交互技术需要进一步提升，构建模拟人类的语音感知和识别能力，更高的识别和合成精度，更强的对语言内容的理解能力。

四、共识与建议

本次论坛围绕语言智能技术研究领域，深入探讨了语言智能技术未来发展及应用的相关问题，促进了语言智能技术领域相关学科发展战略研究更趋系统深入，为推动我国语言智能技术的加速发展做出了努力。与会院士专家认为：

（1）近年来，机器学习、深度学习等智能技术的发展为语言智能技术的进步注入了新的活力，推进语言智能技术不断走向实用，相关

产品表现形式丰富，市场反应积极。同时，国家“一带一路”倡议的实施既为语言智能提供了广阔的应用空间，也对语言智能技术提出了新的要求、新的挑战。

(2) 结合自身的特点，语言智能技术进一步提升，需要在其基础技术层面寻求突破和发展，这些技术包括机器推理，语义知识的表示和利用，以及各类知识库和语言资源的建设。

(3) 在语言智能技术发展探索的进程中，需要重视对我国传统语言学（尤其是音韵、训诂等）研究成果和方法的利用和借鉴。民族的语言不仅是表达思想和人际交流的工具，而且是人（该民族）的思维方式、把握主客观及其关系的内心所在。因此在语言智能处理的理论、模型和技术中，如果更加注意汉语汉字音与义的相关度及其来源和嬗变，有望在“深度语义理解”和人机“自由交互”方面更上一层楼。

(4) 重视跨学科之间的融合借鉴，一方面需要将人工智能的基础技术和方法，以及人工智能其他技术领域的有效方法不断与语言智能技术的具体研究相结合，促进语言智能技术的发展；另一方面，也需要将语言智能技术拓展应用到相关的技术领域，充分利用技术进步带来的效益。

五、论坛总结

与国外语言智能技术同领域研究相比较，国内的研究在具体场景应用技术方面、语言资源的建设和使用方面具有优势，在基础理论模型和技术基础方面还存在差距。

在语言智能技术领域的探索中，我国有学者强调语言智能技术与语言之间的紧密联系，而语言又是一个民族文化的重要载体，因此需要在更大的场景和视野下发现和聚焦语言智能技术可能产生变革的学科领域，为我国的语言智能技术发展抢占先机。

本次论坛围绕语言智能技术研究领域展开，深入探讨了语言智能技术未来发展及应用的相关问题，发挥了论坛促进学术交流、倡导科

学民主的作用，体现了中国科学院学部的学术引领作用。本次论坛产生的观点还将推动“语言智能技术未来发展及应用”咨询项目的研究，服务咨询报告的形成和完善，促进语言智能技术领域相关学科发展战略研究更趋系统和深入。

（作者：李启虎，中国科学院院士，中国科学院声学所研究员；尹力，中国科学院声学所研究员；张全，中国科学院声学所研究员）

联系方式：中国科学院学部工作局学术与文化处，010-59358366