

专家谈 AI 前沿风向

竞速新赛道，让 AI 赋能千行百业

复旦大学计算机科学技术学院教授邱锡鹏

发挥上海自身优势，建设国际算法创新基地

■本报记者 徐品卉

如果用一个词来概括 2023 世界人工智能大会(WAIC2023)的热点,那么大模型当之无愧。百度文心、阿里通义、华为盘古、讯飞星火、商汤日日新、复旦MOSS……国内外 30 余款大模型将同台竞技,背后的核心开发团队也将悉数亮相。

这是一条“狂飙”的新赛道。去年底,对话式大型语言模型 ChatGPT 火爆出圈,激起 AI 领域“千层浪”,也带动大模型这一重要的底层基座变得炙手可热。今年以来,国内“百模大战”趋于白热化,那么下一步是重技术还是重应用?自主研发大模型,对于上海科创中心建设有着怎样的意义?2023 世界人工智能大会开幕之际,记者采访了复旦大学计算机科学技术学院教授邱锡鹏。

出生5个月,MOSS更强了

作为学界顶尖研究者,邱锡鹏还有一个更让人关注的身份——复旦大学 MOSS 系统负责人。今年 2 月,邱锡鹏团队发布 MOSS,成为国内首个发布的类 ChatGPT 模型。

复旦大学 MOSS 系统亮相 WAIC2023,将是该大模型首次现身大型展会。邱锡鹏告诉记者,过去 5 个月来,MOSS 系统日夜成长,不断迭代优化,比如:4 月 20 日,第三代迭代版本 MOSS003 大模型上线,成为国内首个插件增强的开源对话语言模型。“相比 2 月‘出生’时,如今的 MOSS 系统性能提升不少,主要表现在逻辑能力、无害性、有用性等方面。”邱锡鹏总结道,大模型靠大量语料来训练,一些能力已超越大部分成年人。

“围绕大模型,我们做的不仅仅是模型本身,还有很多挑战要解决。”邱锡鹏举例说,其中一个就是普惠化,即将大模型巨大算力成本降下来,“过去优化一个百亿参数量级的大模型,需要数千 A100 芯片做硬件支撑,一台机器就要 100 多万元,并非普通研究机构承担得起。”5 个月来,这也是邱锡鹏的主要研究方向之一。就在上个月,他带领团队发布了低内存优化技术(LOMO),可将大模型训练内存使用量降低到之前的 10.8%。新方法能够在一台消费级显卡的机器上,对 650 亿参数大模型进行全参数微调,大大降低了使用门槛。

技术与应用,两手都要抓

随着 ChatGPT 火遍全球,国内“大模型之战”也已打响。中国科学技术信息研究所发布的《中国人工智能大模型地图研究报告》显示,截至 5 月底,国内 10 亿级参数规模以上的大模型已发布 79 个,互联网大厂、人工智能公司、行业头部企业、高校及科研机构、大数据及云计算供应商都参与到大模型的比拼中。如果算上垂直应用方向,说“千模大战”并不过分。

那么,走出实验室阶段的大模型,未来发展应该更重技术还是更偏应用?在邱锡鹏看来,技术与应用两手都要抓。作为一线科研工作者,他对于当下国内的大模型技术并不满意,“虽然是‘百模大战’‘千模大战’,但国内竞逐者的大模型基座比较雷同,从技术到模型架构,再到能力,没有让人特别眼前一亮的产品。竞争者未来一定要结合自主研发,做出更多有创新性的产品。”

邱锡鹏也提到了应用驱动的重要性,正如中国在移动互联网应用领域领跑全球,应用是我们的优势项。“任何技术走出实验室,在下落落地应用的过程中,都会遇到很多挑战,以应用驱动技术发展是一种很好的方式。”但他指出,目前国内缺少的是支撑高级应用的功能,比如一些高级的插件以及推理能力,特别是在有用性方面,缺少能真正帮助人们提升效率的杀手级应用。

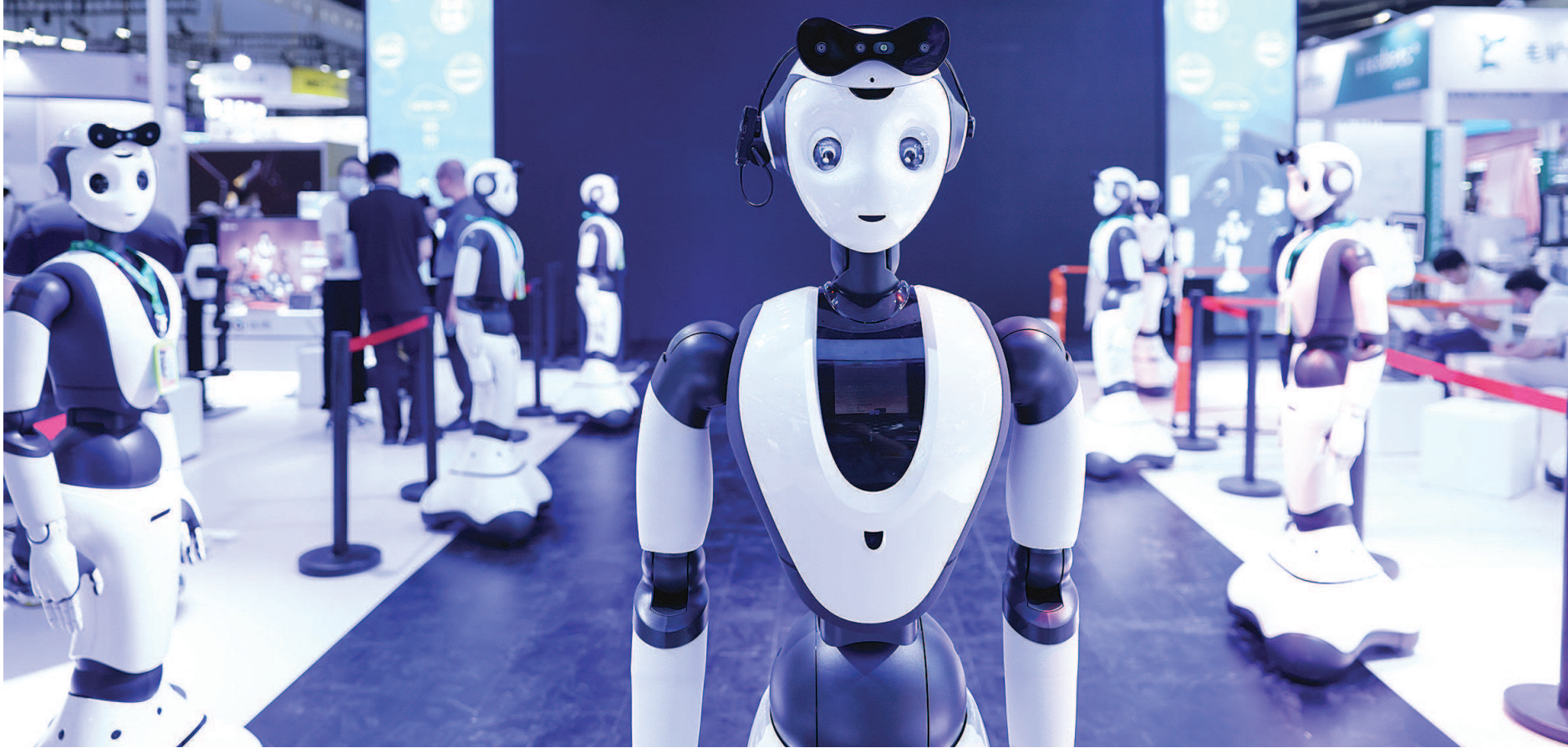
他透露,目前复旦大学 MOSS 系统在开源社区里的下载量排名靠前,“截至目前,我们已收到几百家企业的使用申请,覆盖智能信息服务、金融、医疗、教育等行业,不少企业在使用后也给出反馈,绝大部分希望在垂直领域有进一步优化。”

做好技术基座助力科创中心建设

作为生成式人工智能的技术底座,大模型无疑是兵家必争之地。从《报告》数据来看,国内大模型的“出产地”集中在北京、上海、广东和浙江等省市,这 4 个地方也是近 3 年人工智能服务器采购数量最高的地区。

在邱锡鹏看来,争夺大模型的技术底座话语权,对于上海科创中心建设的意义重大。相比于国外技术实力,目前国内大模型还有一定差距,体现在语料清洗、工程、算法等方面,对自主研发的要求很高。再往生态圈延伸,国内大模型需要从头到尾有一套自主研发的技术,适配国内的配套硬件,保障在任何环节不被卡脖子。邱锡鹏认为,上海在前沿科技领域有独特的优势,体现在人工智能企业数量多、能级高,从底层技术到大模型、算力,从系统设计到应用,全链条企业都有涉及,更容易从生态上做规划。

就在近日,《上海市推动制造业高质量发展三年行动计划(2023-2025 年)》出炉,其中对大模型、算力等多方面作出指引,提出“瞄准人工智能技术前沿,构建通用大模型,面向垂直领域发展产业生态,建设国际算法创新基地”。邱锡鹏认为,上海应发挥自身优势,在 AI 大模型领域走出有中国特色的创新路。



达闼机器人展区,搭载大模型对话能力的双足机器人正在进行“彩排”测试。

本报记者 袁婧摄

上海大学终身教授、上海市智能制造及机器人重点实验室名誉主任方明伦——

国产机器人:正在从“高原”向“高峰”攀登

■本报记者 周渊

机器人被誉为“制造业皇冠顶端的明珠”,其研发、制造、应用是衡量一个国家科技创新和高端制造业水平的重要标志。“上世纪 80 年代,我曾提出机器人是人类生产的工具、生活的助手,如今看来,机器人已全方位服务于国民经济和社会生产生活,机器人产业快速发展。作为一项庞大的系统工程,可以说,机器人的竞争也是未来国家综合实力竞争的一大热点。”著名机器人专家、上海大学终身教授、上海市智能制造及机器人重点实验室名誉主任方明伦认为,如今,我国机器人发展有了更有利的环境,最有可能突围的机会在于新兴的服务机器人。

机器人也是今年世界人工智能大会关注的十大前沿风向之一,在展会现场,达闼搭载大模型对话能力的双足机器人、特斯拉人形机器人“擎天柱”、上理工小贝 4.0 等 20 余款智能机器人集中亮相,共同探索未来机器人产业的无限可能。

在方明伦看来,近年来,中国机器人产业发展呈现“量大面广”态势,但仍处在从“高原”向“高峰”攀登的过程中。“抓好顶层

设计和基础研究两端,持之以恒地推进,未来中国一定会诞生顶尖的机器人企业。”

服务机器人是上海参与全球竞争的机遇

全球范围内,机器人应用最广泛的是工业领域。方明伦告诉记者,在我国,工业中低端领域已大量应用国产机器人,但高端领域依然依赖进口。

上海是中国最早引进和开始研发工业机器人的主要城市之一,上世纪 80 年代就研制出 5 个型号的工业机器人,方明伦主持研发了“上海 2 号”和“上海 5 号”工业机器人。“40 多年过去了,我国机器人技术有了很大进步,产业有了很大发展,但国际机器人四巨头日本发那科、安川电机、瑞士 ABB 和德国库卡依然占据着主要市场,我想中国应该有几家这样的企业,上海至少也应该有一家。”研究了 40 多年机器人,这位年过八旬的老教授不无期待地说。

突围的机遇或许在于服务机器人。方明伦谈到,随着波士顿动力人形机器人“阿特拉斯”、特斯拉人形机器人“擎天柱”相继

出圈,机器人产业已然发生变化,产业链长、应用更为广泛的服务机器人将是未来竞争的主战场。“服务机器人在国际上尚未形成‘四巨头’这样的企业,在国内,服务机器人已在医疗、养老等民生领域试水应用,细分领域涌现了一些优秀企业,我有信心,国产机器人一定会领先世界。”他说。

方明伦特别提到,上海有着先发优势,作为国产机器人研究和生产应用的高地,市政府出台多项政策,高校和研究机构位居全国前列,并在张江等地形成了产业集聚效应。

进一步抓好基础研究和人才培养

对于服务机器人的发展方向,方明伦总结为更智慧、更低碳、更低成本,而这背后都指向了更为扎实的技术底座。

“国内机器人产品目前竞争不过进口机器人,不在于不懂原理,不擅长组装,而是被卡在了元器件上。因为基础材料和工艺达不到要求,使得国产元器件的性能无法与进口产品相比。”方明伦与记者分享了一个小故事,曾经他和一位朋友去日本出差,回国

时朋友特意带回来几个螺栓。“螺栓的生产非常简单,只涉及材料和热处理,但当时国内没有让他满意的产品。这也意味着,我们要发展机器人产业,必须重视材料和工艺等基础研究。”

机器人是技术密集型行业,机器人产业的发展需要多学科交叉融合人才。方明伦表示,波士顿动力的机器人能翻跟斗、跑步,可以看出他们的科研人员具备强大的数学、力学功底。我们的人才培养模式也必须与时俱进,打好数学基础,形成良好的思维方法,结合机器人项目培养学生,让教育最终回到产业中去。

眼下,方明伦所在的上海市智能制造及机器人重点实验室将研究重点放在了智能康复机器人、人形仿生机器人、农业机器人等领域。其中,智能康复辅助机器人已在华山医院、岳阳医院以及上海多个社区卫生服务中心投入应用。

方明伦认为,产品的好坏最终要接受市场的检验。“政产学研用”应当同步进行,对于还在迭代的产品,市场需要更加包容,政府应当在产业应用中加大引导力度,让国产机器人有机会被应用、有机会迭代发展。

脑机接口专家陶虎——

大模型赋能,脑机可望开启“大脑元宇宙”

■本报记者 张懿

融合“电脑”与人脑,BT(生命科学)与 IT(信息科技),类脑智能就处在这样一个充满魅力的交叉点上。今年世界人工智能大会将类脑智能以及脑机接口设为重点探讨的十大前沿方向之一。在此背景下,记者专访了国内脑机接口研究的领军人物,中国科学院上海微系统与信息技术研究所副所长、脑虎科技创始人兼首席科学家陶虎,请他对这一领域的最新成果进行解读,对其未来发展作展望。

一流团队瞄准超级脑力

某种程度上,人工智能就是模仿人类最强大的器官——大脑,而脑机接口则是借助 IT 技术,与大脑进行信息交互的重要手段。通过脑机接口,我们可以让计算机了解大脑是如何工作的,而后去模拟其思维方式。

科学技术部发布的“中国脑计划”重大专项,总体架构被定为“一体两翼”。“一体”就是“认识脑”即推进前沿脑科学研究,“两翼”则是“保护脑”“模拟脑”,前者聚焦重大脑疾病诊治,后者则融合脑机接口与类脑智能相关研究。

作为“中国脑计划”模拟脑领域责任专家,陶虎已率团队跻身国际脑机接口的第一梯队。依托中国科学院上海微系统与信息技术研究所

高水平的底层技术平台,他们形成了脑机接口核心器件、芯片的自主研发和制造能力。同时,还推出独具特色的基于蚕丝蛋白的柔性脑机接口技术,相较于其他团队的植入式方案,该柔性技术无论在手术本身的创伤,还是植入以后长期在体的安全性、稳定性上都更具优势。

陶虎告诉记者,大脑就像一台超级计算机,一方面功能强大,拥有 800 多亿个神经元,每一个都能与周边成百上千的神经元进行物质和信息交互,形成强大网络;另一方面,这台“超算”能耗极低,平均不到 30 瓦,而按照冯·诺依曼架构设计的顶级超算,虽然算力极强,但电力消耗却抵得上几万户家庭。

要让电脑具备人脑的能力和“性能功耗比”,目前有两条可能的路径:一是通过类脑研究,开发类脑芯片、类脑计算架构,模拟大脑运行;二是借助脑机接口与人脑交互,跳过五官、四肢,直接利用大脑的核心能力。

脑机接口赋能 AI 大模型

当前,由陶虎创立的脑虎科技已搭建了一个包括物理学、化学、生物学、机器人技术、材料科学、临床医学等在内的多学科交叉团队,而且团队从一开始就紧扣临床需求

开展工作。过去一年,他们针对一些脑疾病开展临床试验,并取得积极进展。本次世界人工智能大会上,他们将在“脑机智能与数字生命”主题论坛上发布并解读脑机接口技术阶段性的科研成果。

同时,对于生成式人工智能(AIGC),陶虎也是高度关注。他认为,由 ChatGPT 所引发的 AIGC 热潮可能是人工智能领域多年来最重要的技术进展,这让很多人都看到 AI 大模型可以在某些领域表现出接近人脑的能力,甚至有时比人更出色。这一突破,为脑机接口进一步迭代提供了重要契机。

陶虎表示,算力、算法和数据是推动 AI 发展的三大要素,其中最关键的可能是数据。通过脑机接口,未来他们团队或许有机会把采集到的大量高质量、标准化大脑信号输入给 AI,从而推动开发“脑机接口大模型”。这个大模型不仅能模仿人的智能,更可以模仿心灵。假如“脑机接口大模型”可以学习某些精神异常人群的脑活动特征,然后叠加艺术类 AI 大模型对梵高作品的理解,那么, AI 对梵高的模拟完全有机会触及更本质层面,从而创作出更神似梵高风格的画作。

再进一步,假如这样的 AIGC 大模型能不断演进,变得更强大、更成熟,那就有机会开发出所谓“大脑元宇宙”。陶虎认为,当进入这样一种“赛博空间”后,人们不仅可以在物理层面与虚拟场景互动,更可以和其他角色在意念层面进行交流,从而营造出远比一

般元宇宙更深刻的沉浸感。

未来一年脑病治疗或迎突破

当前,脑机接口研究已进入临床阶段,并有望带来实质性改变。陶虎透露,包括他在内,国际上有多多个脑机接口团队都将关注重点放在重大脑疾病治疗上。未来一年,脑机接口很有希望在渐冻症辅助治疗方面获得临床的功能性验证。

渐冻症是一种因为运动神经元快速凋亡导致人不断失去语言和肢体功能的疾病。在借助脑机接口读取病人的大脑信号后,研究者可以在 3 个层面治疗带来帮助。据陶虎介绍,第一步是通过脑机接口,让病人能用意念操控机械臂,重塑部分运动功能,或是借助计算机语音合成技术,重获语言表达能力。第二步,脑机接口或许能帮助医生发现神经元凋亡的规律,从而加速相关药物和疗法的开发。第三步,也是最高阶段,人们或许有机会用脑机接口激活和调控大脑中冗余的“备份神经元”,让它们取代那些凋亡的神经元,使病人恢复健康。陶虎认为,这“三步走”的过程有望在未来 8 至 10 年完成。

除了面临床床需求,让病人变成正常人外,脑机接口更广阔的应用场景也许是为普通人赋能——用陶虎的话说,就是要让正常人变成“超人”。他表示,人的五官和四肢,其实并没有充分利用大脑的潜力。未来,我们应该有机会让脑机接口,与那些比人体器官更强大的传感器、执行器相连,从而让人获得超级视觉、超级听觉,以及超级体能。这样,脑机接口就会变成一个强大的赋能平台,“我希望在未来 15 至 30 年内,这一幕能变成现实。”

智联世界 生成未来