

# 人工智能加快“飞入”寻常百姓家

## 产业界

张辛欣

现场给出关键词,机器就能即兴作诗;点单时对着菜单拍照,便可实现28种语言之间互译……近日,中国科学技术馆与百度公司合作,通过移动大巴方式将人工智能各项应用向社会展示。与此同时,我国也将通过多种形式加快人工智能科普和应用,加速布局产业链,智

能生活正在向人们走来。

作为新一代信息科技的代表,人工智能受到国际社会普遍关注。发达国家纷纷布局于此,试图抢占技术高点和先机,我国也不例外。不久前,国务院印发《新一代人工智能发展规划》,提出到2020年,人工智能总体技术和应用与世界先进水平同步,人工智能产业成为新的重要经济增长点,到2025年,人工智能成为我国产业升级和经济转型的主要动力,智能社会建设取得积极进展。

经过多年积累和攻坚,我国人工智能技术快

速发展,在算法、数据等方面具备领先优势,一些技术甚至带来超乎想象的效果。比如,目前机器语音识别能力已经接近或超过人类。然而,人工智能不只是酷炫,更是要发挥“四两拨千斤”作用,带动社会生活便捷高效,推动产生效益。

百度副总裁王路说,此次和中国科学技术馆达成战略合作框架协议,第一步就是通过移动大巴的形式展示包括图灵测试、语音识别、智能家居等人工智能应用,后续还将把最新技术应用到科普项目中。

此外,百度将把6年来在人工智能领域的创新开放出来,建立共享合作生态体系,加速技术成果应用落地。

工信部信息化和软件服务业司副司长朱皖表示,工信部将积极协调资源推动人工智能技术转化为各类产品和各种应用,并将人工智能产业发展与一系列重大应用工程相结合,加快促进人工智能与各产业领域深度融合,形成数据驱动、人机协同、跨界融合、共创分享的智能经济形态。(据新华社)

## 新鲜事

### 国内最大智能机器人仓库启用



科技日报讯(记者李禾)菜鸟网络高级算法专家胡浩源日前接受科技日报记者采访时表示,由菜鸟网络打造的我国最大机器人仓库已在广东惠阳投入使用。该仓库内有上百台机器人,使天猫超市商品上午下单下午即可送达。

目前国内不少智能仓库内搬运机器人仅十几、二十台。“几十台和上百台机器人一起工作的难度大不相同。”胡浩源介绍说,更多机器人意味着分配任务难度大大增加,合理地将每个任务分配给对应机器人,从而实现整体任务完成效率最优,还要防止机器人之间可能的碰撞、防止部分区域出现机器人拥堵、死锁等。

据菜鸟网络物流专家游有泉介绍,每台机器人能顶起的重量可达500斤,同时还能灵活旋转,可将货架的四面调配到拣货员跟前,方便拣货员工作。“这等于一个货架,四个面都能存储商品,仓库容量被提升了一倍还多。”机器人缺乏电力时会自动归巢充电。

据了解,传统货仓内,消费者下单后,拣货员需跑到货架前将货物拣出。人工拣货通常一个小时拣100多件,而在智能仓内,一小时拣货可达3000件。“这个数值还不是智能仓的最高值。”胡浩源说。

## 大咖说

### 我国机器人企业仍处低端 低水平重复性建设存隐忧

新华社北京8月8日电(记者闫睿 余晓洁)工信部装备工业司副司长罗俊杰在8日举行的2017世界机器人大会新闻发布会上说,中国机器人产业呈现快速发展态势,但不少机器人企业小、散、弱,处在中低端;减速器、控制器、伺服电机等一些关键零部件创新能力薄弱,低水平重复性建设存隐忧,亟待集中力量解决。

机器人领域是当前我国与发达国家之间创新的必争领域。美国、欧盟、日本等均已出台机器人发展战略。按照《中国制造2025》统一部署,我国发布了机器人产业发展规划,对“十三五”期间机器人产业发展进行部署。

“中国已连续3年成为全球最大的工业机器人生产国,占全球25%的市场。庞大的市场为实现领域‘弯道超车’提供了难得的历史机遇,没有理由不加强在这个领域的布局。”中国电子学会副理事长兼秘书长徐晓兰说。

据介绍,针对我国机器人产业发展中存在的问题,工信部会同有关部门出台了相应政策、规章等,通过集中协作突破,打好组合拳,推动中国机器人既创新快速又健康可持续发展。

“一方面,通过进一步整合产业创新资源,推动建立机器人创新中心,重点支持机器人产业关键零部件研发攻关。”罗俊杰说。同时,加大对机器人专业人才培养力度,夯实产业发展基础;开展机器人产业试点示范推广应用,推进机器人在“中国制造2025”十大重点领域及传统制造业应用。另一方面,加强产业发展规划政策的引导作用,促进产业科学化、理性、健康发展。

# 桂冠拿了,“明珠”何时到手



8月4日,中国科学技术馆里小朋友在“旅行号”AI科普大巴上与机器人互动。(新华社记者 马化娟摄)

近日,由斯坦福大学发起的SQuAD(Stanford Question Answering Dataset)挑战赛榜单更新成绩。在这一机器阅读理解领域的国际顶级赛事中,科大讯飞与哈工大联合实验室(HFL)提交的系统模型在来自全球学术界和产业界的研究团队中夺得第一名,这是中国本土研究机构首次摘得该领域的顶级赛事桂冠。

本报记者 吴长锋 李丽云

经常遇到以下两方面的难题:

首先是数据问题,目前基于深度学习模型等统计方法的机器阅读理解,在研究上离不开大量的被人工标注的数据,但是固有的数据集往往存在规模较小、质量不佳等问题。这种情况下,很难基于这些数据做出优秀、有效的模型。

二是算法问题,传统的NLP在做阅读理解或自动问答时,会采用分拆任务的方式将其分成问题分析、篇章分析、关键词抽取等一些步骤,但这种方法容易造成绩效误差的积累,很难得到很好的效果。如,分目标的结果整合最后和终极目标不能完全吻合;局部优化不当可能造成研究进展缓慢等。

“传统自然语言是采用分拆任务的方法将其分成问题分析、篇章分析、关键词抽取等一些步骤,但这种方法容易造成绩效误差的积累,很难得到很好的效果。”科大讯飞AI研究院院长胡国平博士告诉科技日报记者,目前在机器阅读理解领域中,多采用完全端到端的神经网络建模,消除了分步骤产生的级联误差。通过大量的训练数据学习到的泛化的知识表示,对篇章和问题从语义层面上高度抽象化。

“科大讯飞采用的端到端的神经网络模型,把精力更多放在如何能够模拟人类在做阅读理解问题的一些方法。”胡国平博士告诉科技日报记者,科大讯飞提出的基于交互式层叠注意力模型,主要思想是根据给定的问题对篇章进行多次地过滤,同时根据已经被过滤的文章进一步筛选出问题中的关键提问点。这样“交互式”逐步精确答案的范围,最终收获了令人瞩目的成绩。

## 相关链接

### 大数据让机器阅读看到曙光

人们在读完一篇文章之后一般就会在脑海里形成一定的印象,这种对人类而言轻而易举的能力,对计算机来说却并非如此。

据新智元报道,很长一段时间以来,自然语言处理的研究都是基于句子级别的阅读理解。例如给计算机一句话,理解句子中的主谓宾、定状补,谁做了什么等等。但长文本的理解问题一直是研究的一个难点,因为这涉及到句子之间的连贯、上下文和推理等更高级的研究内容。

比如下面这段文本:The Rhine is a European river that begins in the Swiss canton of Graubünden in the southeastern Swiss Alps, forms part of the Swiss-Austrian, Swiss-Liechtenstein border, Swiss-German and then the Franco-German border, then flows through the Rhineland and eventually empties into the North Sea in the Netherlands. The biggest city on the river Rhine is Cologne, Germany with a population of more than 1,050,000 people. It is the second-longest river in Central and Western Europe (after the Danube), at about 1,230 km (760 mi), with an average discharge of about 2,900 m<sup>3</sup>/s (100,000 cu ft/s).(大意:莱茵河是一条位于欧洲的著名河流,始于

瑞士阿尔卑斯山,流经瑞士、奥地利、列支敦士登、法国、德国、荷兰,最终在荷兰注入北海。莱茵河上最大的城市是德国科隆。它是中欧和西欧区域的第二长河,位于多瑙河之后,约1230公里。)

若针对该段内容提问:What river is larger than the Rhine?(什么河比莱茵河大?)人们可以轻易地给出答案:Danube(多瑙河)。但目前即使是最佳的系统模型给出的结果也并不尽人意,它的回答是:科隆。

可见要让计算机真正理解文本内容并像人一样可以对文字进行推理的难度是如此之大。在回答该问题时,计算机除了要处理文中的指代“it”,还需要算法和模型进一步对“larger than”和“after”这两个表述进行推理,从而得知Danube是正确答案。此外,由于文中并没有提到Danube是“river”,所以又加大了系统的推理难度。

但大数据的发展让学者们看到了这一研究方向的曙光。可获取的越来越大的文本数据,加上深度学习的算法以及海量的云计算资源,使得研究者们可以针对长文本做点对点的学习,即对句子、短语、上下文进行建模,这其中就隐藏了一定的推理能力。所以,目前自然语言处理领域就把基于篇章的理解提上研究的议事日程,成为目前该领域的研究焦点之一。

## 聚焦

# 马晓磊:交通大数据的“青年砥柱”

李海峰



### 结缘交通大数据

马晓磊本科就读于北京理工大学通信工程专业,初次接触交通,对马晓磊来说,是一个比较陌生的专业。此后,他来到美国求学,跟随导师王印海教授来到美国华盛顿大学深造,王印海教授老师带领马晓磊走进了交通大数据领域,让他进入了交通科学的世界,认为他应该将学到的理论知识运用到实际当中。

近日,年仅32岁的北京航空航天大学交通科学与工程学院副教授马晓磊荣获了2017年北京市优秀青年人才荣誉称号。对于北京市35岁以下的青年人才来说,这是一个非常重要的荣誉,北京市政府每三年只有26名青年人才入选。

“能赢得这份荣誉我很荣幸,对我也是莫大的动力。”马晓磊倍感振奋。作为一位80后,他和大多数年轻人一样,平日喜欢游泳、健身、打羽毛球。

但是来到实验室和教室里,就会发现这位80后不一样,他自信、谦虚、认真,只有在这个时候,很多人才会懂得,马晓磊获得这份荣誉背后是他过人的天赋和夜以继日、刻苦努力。

马晓磊认为交通工程就是一个这样的专业。通过接触交通仿真、公交信号优先等项目,他觉得交通工程越来越有意思,越来越有收获。因为这个学科能够把自己所学的知识运用到交通实际中,能解决具体问题。在美国华盛顿大学攻读博士期间,马晓磊与导师王印海教授在国际上率先提出E交通科学理念。E交通科学为多源异构数据的融合和挖掘提供了一种非常有效的方法,通过数据挖掘、数据融合和人工智能等技术对原始数据进

行质量控制,利用云计算和地理信息数据库将交通综合信息有效提取出来,实现了大数据科学在交通科学领域的开创性应用。研究成果被国外多家媒体报道,并被美国交通部邀请到华盛顿特区总部参加“Transportation Data Palooza”会议并作报告。

回到中国后,身为80后,马晓磊不负众望,在科研和教学方面挑起了大梁。马晓磊在国外的時候,就单独教过一门本科生的课程——交通管理控制,受到了同学们的欢迎和认可,学生们纷纷反映收获颇丰,最后课程评分时候他的课程远远高于平均分。回到中国后,身为80后,马晓磊不负众望,在科研和教学方面挑起了大梁。马晓磊非常感谢他目前所在的北航团队,尤其是长江学者、北航副校长王云鹏教授,王云鹏教授

在马晓磊在北航工作期间,无论是经费上还是学生培养上,给予了很大的支持。

回国后,马晓磊更是把在国外的教学经验带到了国内。如今,他教授的两门课程——智能交通系统和交通大数据分析,在这两门课程上,依托自身的国际化视野,马晓磊用了大量国外智能交通案例与同学互动,并进行双语教学,为同学们打开了眼界。

在教学中,马晓磊特别注重学科之间的碰撞,他总能给学生带来很多惊喜和启发。“教书育人,最重要的是要和同学之间互动。”在实际教学中,马晓磊还把典型交通管理项目方案,当成课程作业让学生们进行分组讨论,并鼓励同学们进行双语演讲,这样的教学,让同学们收获颇丰。

马晓磊认为,青年科技人才,尤其是海归人才,要高速高质量成长,一定要多合作,合力提升在国际上的影响力。“青年科技人才应该主动积极地参加一些国内外的学术会议,多和国内外同行建立学术合作。”马晓磊总是

鼓励自己的学生,多向国际前沿看齐。

### 科研应用成果丰

迄今为止,马晓磊有两篇论文具有很重要的影响力。

一篇是2013年,马晓磊发表在Transportation Research Part C 交通领域顶级期刊上的论文——《Mining Smart Card Data for Transit Riders' Travel Patterns》。这篇文章针对城市公交居民出行起讫点估计问题,针对上车刷卡下车不刷卡的难点,提出了一系列人工神经网络模型中在长时序数据预测时无法收敛的问题,可使预测精度提升近2倍。目前,这篇文章的研究成果已经和中国一汽等企业进行了深度合作。

“下一阶段,我计划继续推进,将科研成果产业化,并计划在中国的一些城市进行试点,让交通大数据为更多人智慧、便捷出行服务。”马晓磊说。目前,马晓磊团队已经与成都、杭州、济南等城市开展了“互联网+公交”项目合作,通过融合移动互联网数据与传统的公交IC卡,优化传统公交网络,开设新型定制公交线路,提升公交服务水平,吸引更多居民乘坐公共交通,最终达到缓解拥堵的目的。智能、智慧交通的未来,便是马晓磊期望的未来,这个未来,在马晓磊这样的青年科学家努力下,形成了结实的中流砥柱,为智能、智慧交通的未来打下了坚实的基础,也为这样的蓝图描绘了更确切的实现日期。