

“火爆”的ChatGPT尚没有“自主心智”

——访人工智能专家、欧洲科学院院士金耀初教授

科技创新世界潮 ②①

◎本报驻德国记者 李山

近日,围绕热门科技话题 ChatGPT,科技日报记者专访了欧洲科学院院士、德国比勒菲尔德大学洪堡人工智能教授金耀初。金教授同时兼任英国萨里大学计算系“计算智能”讲席教授、IEEE 计算智能学会候任理事长,曾任《IEEE 认知与发育系统汇刊》主编等职。

生成式模型的重大进展

谈到 ChatGPT 最核心的东西,金耀初表示,ChatGPT 所代表的技术是第三波人工智能技术取得的一个重大进展。它基于生成式预训练 Transformer 模型,采用了“自注意力机制”。与传统的判别式模型相比,生成式模型的最大特点是直接学习训练数据本身的特点,因而能更有效地学习大容量的样本,特别是海量的语言、视觉信息。

传统的语言处理模型一般只能根据字词的前后相邻词语来找它们之间的关系,因此有很大的局限性。Transformer 引入了自注意力机制后,它就能很灵活地找出不单是某一个句子,而是不同句子,甚至是整篇文章或不同的文章之间词的相互关系,学习能力就增强



德国比勒菲尔德大学的金耀初教授长期从事计算智能、演化计算、机器学习、计算神经科学等交叉学科的理论研究和工程应用,目前已发表学术论文 500 余篇,论文被引用次数达 37400 余次,2019—2022 年连续 4 年被 Web of Science 评为“全球高被引科学家”。

照片由受访者提供

了很多。

横空出世非一日之功

对于 ChatGPT 的横空出世,金耀初强调这仅仅是大众的感受。从科研的角度看,ChatGPT 也是一步一步演进而来的。根据 OpenAI 发表的论文,ChatGPT 是花了很长时间,用了各种学习方法不断训练和微调出来的。生成式模型给出的答案并不是人们事先给它的训练数据,而是模型生成出来的,所以很难保证百分之百正确。ChatGPT 在训练过程中用到非常多的最新的学习技术。例如,无监督学习、有监督学习、多任务学习、小样本学习、自监督学习等。

此外,它还加入了基于人类反馈的强化学习,以获得与人类对齐、共情能力。通过人类评价,即给它生成的答案打分来选择最佳的答案。这个过程实际上是一个强化学习过程,也就是对原来的模型进行微调。其他 AI 模型可能没有像 ChatGPT 这样,用大量的人类反馈来使得模型能产生更加符合人类期待的回答,相应的体验就会有差距。

尚不具备自主的“心智”

金耀初谈到,作为一个从文本到文本的对话系统,ChatGPT 最强的地方是对自然语言的充分学习,生成的文本更



图片来源:视觉中国

加“流畅”并符合人类体验,这是一个重大的技术创新。但值得注意的是,生成式机器学习模型有很多,比如从文本到语音信息,或者文本到图像信息,这些方面 ChatGPT 的表现可能就没那么好了。这是它的一个局限性。

另一方面,目前 ChatGPT 训练的数据截至 2021 年,没有训练过的东西,它是不可能准确回答出来的。换句话说,目前的模型不具备创造新知识的能力。ChatGPT 回答问题用的是推理。它不是那种在数据库中搜索答案的概念。生成式模型是根据问题本身联想出一个答案。这个答案是模型“学习消化”后产生出来的东西,不是原始的数据。严格意义上来说,模型本身甚至不知道这个答案的真正含义,它也不能保证答案肯定是对的。

ChatGPT 不是“万能”的

金耀初认为,ChatGPT 离真正的专业应用还有一定的距离。现在大家感觉很新奇,都在尝试玩一玩。有的人甚至以为 ChatGPT 什么都能干,可以很快用在自动驾驶之类的很多应用上,这其实是误解。自动驾驶的重点不是文本到文本的问题,纯粹的生成式模型是不够的。它需要其他判别式的机器学习模型,特别是实时场景的识别。以现在热议的医疗服务为例,尽管

ChatGPT 在美国执业医师资格考试中取得不错的成绩,但真实情况下人们会用 ChatGPT 来代替医师进行诊疗吗?至少目前还不行,因为无法保证它的答案百分之百是正确的。有人说 ChatGPT 有时会一本正经地胡说八道,这样的描述是比较中肯的,人们不能忽视这种风险。

期待未来更多突破

金耀初表示,这一波人工智能可能不会像第一波和第二波人工智能浪潮那样很快偃旗息鼓。它会继续发展下去,未来三五年,人工智能领域可能还会涌现一些令人吃惊的突破。很多基于 ChatGPT 的新的想法或技术或会涌现,并在此基础上找到真正好的细分领域的应用。

至于人工智能和机器人的结合,金耀初强调,从某种意义上来说,这一直是人工智能研究的重点之一。现在 ChatGPT 与人对话这么流畅,如果把它与数字虚拟人技术结合起来,一般人可能很难分辨出计算机屏幕里回答问题的真人还是一个数字人。当然,未来的发展过程中会涉及很多人工智能的本质问题,包括信任、责任、伦理、法律问题等。乐观地看,ChatGPT 等人工智能应用的影响力急剧扩大,反过来会在一定程度上推动上述问题的解决。

单个电子成像法出现

有望改变新药研发和量子材料表征

科技日报北京 2 月 14 日电(记者刘震)以色列魏茨曼研究所领导的科学家团队在最新一期《应用物理评论》杂志网站发表文章,展示了一种对单个电子成像的新方法。该方法目前仍处于初始阶段,未来有望为各种不同的分子拍摄“特写”,这可能彻底改变新药研发和量子材料的表征。

几十年来,尽管磁共振成像(MRI)

技术在诊断大量疾病方面发挥了重要作用,但仍有一些潜在问题有待解决。例如,MRI 读出效率非常低,需要数千亿个分子的样本量才能发挥作用,这就导致输出被平均化。对于大多数诊断程序,平均值最佳,但与此同时也会丢失一些细节,而这些细节可能隐藏着很重要的生物学过程。

研究人员解释道,为获取单个分子

的细节,需要开发出一种能获取单个分子特写镜头的工具,研究人员可借助这些工具密切观察重要分子的结构,从而为新发现或新药研发开辟道路。

鉴于此,在最新研究中,研究团队开发出了一种可以精确定位电子位置的方法。该方法基于位于一个氮-空位中心附近的旋转磁场,氮-空位是一种特殊的人造金刚石中原子大小的缺陷,一般

被用作量子传感器。由于该氮-空位仅仅原子大小,这种传感器对附近的变化特别敏感,由于其量子性质,它可以区分是单个电子还是多个电子,这使它特别适合精确测量单个电子的位置。

研究团队指出,最新方法朝着精确纳米成像迈出了关键一步,未来科学家将能够使用这项技术来对不同分子进行成像,为其拍摄“特写”。

可吞咽传感器助力监测胃肠道健康

科技日报北京 2 月 14 日电(实习记者张佳欣)据最新发表在《自然·电子》杂志上的一篇文章,美国麻省理工学院和加州理工学院的工程师们展示了一种可吞咽传感器,当通过消化道时,可以监测其位置,从而帮助医生更容易地诊断胃肠道动力障碍,如便秘、胃食管反流和胃癌。

与目前的内窥镜检查等方式相比,新设备的侵入性更小,而且患者可直接

在家中使用。

研究人员的思路是开发一种可吞咽的胶囊,能发出信号以揭示其在胃肠道的哪个位置,使医生能够确定肠道的哪个部分导致了胃肠道蠕动减慢。

研究人员开发了“迷你”磁性传感器。其工作原理是探测位于人体外的电磁线圈产生的磁场。磁场的强度随着距线圈的距离而变化,因此可根据传感器对磁场的测量来计算传感器的

位置。

为了精确定位设备在体内的位置,该系统还包括第二个传感器,该传感器留在身体外部,充当参照点。这个传感器可贴在皮肤上,通过比较它的位置与第一个传感器在体内的位置,研究人员可准确地计算出传感器的位置。

这种可吞咽的传感器还包括一个无线发射器,可将磁场测量结果发

送到附近的计算机或智能手机。当前版本的系统是在收到智能手机的无线触发信号时随时进行测量,但也可对其进行编程,以特定的时间间隔进行测量。

研究人员表示,该系统支持多个设备的本地化而不会影响精度。此外,它还有广泛的检测范围。当前版本的传感器可检测距离为 60 厘米或更短的电

迄今最古老棘龙大脑“艺术照”亮相

科技日报北京 2 月 14 日电(记者刘震)英国南安普顿大学和美国俄亥俄大学的研究人员携手,重建了迄今最古老棘龙的大脑和内耳,以揭示这些大型肉食恐龙如何与周围环境互动。相关研究刊发于最新一期《解剖学杂志》。

棘龙是一种大型兽脚类恐龙,拥有鳄鱼般的长颌和圆锥形牙齿,这帮助它们适应了某种水生生活方式,包括沿着河岸寻找猎物(如大鱼等)。这种生活方式与异特龙和暴龙等兽脚类恐龙截

然不同,这也使棘龙成为最富有争议的恐龙物种之一。

为更好地了解棘龙大脑和感官的进化,研究小组扫描了来自萨里的重爪龙和怀特岛的冥角角龙的化石。根据它们的脑壳材料,这两种迄今已知最古老的棘龙可以追溯到 1.25 亿年前的白垩纪早期。两个恐龙标本的脑壳都保存完好,研究小组对早已腐烂的内部软组织进行了数字化重建。

他们发现,这两种棘龙处理气味的嗅球并不是特别发达,耳朵可能对低频声音很敏感,而大脑中那些负责保持头部稳定和凝视猎物的部分相比后来出现的棘龙更不发达。

主持这项研究的南安普敦大学博士克里斯·巴克说:“尽管这些古老棘龙的生态环境不同寻常,但它们的大脑和感官似乎在许多方面与其它大型兽脚类恐龙拥有共同之处,没有证据表明,它们的半水生生活方式与其大脑的组织方式有关。”

研究人员解释道,这些最新发现表明,棘龙的祖先已经具备了适合半捕食生活方式的大脑和感官适应能力,而棘龙要想成为半水生动物,需要进化出不同寻常的鼻子和牙齿。

研究人员指出,由于化石 CT 成像技术的进步,他们现在能够评估已灭绝动物的认知和感官能力,并探索像棘龙这样的恐龙的大脑是如何进化的。这有望掀起古生物学领域的一场新革命。

创新连线·俄罗斯

数学模型预测流行病发展动态

俄罗斯秋明国立大学创建了一套数学模型,以新冠病毒为例研究传染病。该模型有助于减少未来流行病对经济和社会的负面影响。

数学家、微生物学家、社会学家、经济学家和大数据专家都参与了这项项目的工作。这项工作提供了新的工具,用于预测感染人数的趋势以及由流行病引起的经济指标变化,例如供应和需求变化。

研究人员认为,建立的数学模型

也使评估各种防止传播措施(如口罩模式、隔离措施)的有效性成为可能,还可研究不同人群的目标,例如按对接健康和生命的风险程度或按疫苗接种率指标进行区分。

秋明国立大学生态和农业生物学研究所高级研究员叶夫根尼·布拉克科夫称,由此产生的工具包可用于制定防止任何传染病传播的战略,包括目前未知的传染病,并最大限度地减少其对经济和社会的负面影响。

“变色龙”伪装服可随环境变化

俄罗斯“胸甲”研究所正在研制一种伪装服,它能让战士对热成像仪隐身,并借助像变色龙一样的颜色变化融入环境中。

该所所长弗拉基米尔·科尔穆申说:“我们今天正在完成的任务,是在不同波段内隐藏人,包括热成像波段。这可通过特殊新材料来实现。”

这种伪装不是固定在士兵身上的

干扰工具或电子战工具,而是使用特殊染料的织物来实现“隐身”效果。科尔穆申称,理想情况是制造这样的衣服,使战士在环境背景改变时不被发现。晚上看不到,早上看不到,在阳光下看不到,在田野和森林中看不到。伪装应该自动变化,像变色龙一样。

(本栏目稿件来源:俄罗斯卫星通讯社。整理:本报驻俄罗斯记者董映壁)

新型「触发器」量子比特问世

像电信号一样容易控制

科技日报北京 2 月 14 日电(记者张梦然)澳大利亚研究人员最近展示了一种新型量子比特的操作,称为“触发器”量子比特,它结合了单个原子的精巧量子特性和普通电脑芯片电信号的易控性。研究成果发表在《科学进展》上。

新南威尔士大学研究团队在世界上率先证明,电子的自旋以及硅中单个磷原子的核自旋可用作量子比特。虽然两个量子比特本身都表现得非常好,但它们的运行需要振荡磁场。而磁场很难在单个量子计算机组件的典型纳米尺度上定位。

团队意识到,将量子比特定义为电子和原子核的上/下组合方向,将允许仅使用电场来控制这样的量子位。这个新的量子比特被称为“触发器”,因为它由属于同一原子的两个自旋(电子自旋和核自旋)组成,条件是它们总是指向相反的方向。

该理论预测,通过相对于原子核置换电子,可对触发器量子比特的任意量子态进行编程。新研究以完美的准确性证实了这一预测。

最重要的是,这种电子位移是通过向小金属电极施加电压而不是用振荡磁场照射芯片来获得的。其更类似于传统硅计算机中通常路由的电信号类型芯片。

通过从原子核中置换电子来对触发器量子比特进行电气控制,会伴随一个非常重要的现象:当负电荷(电子)远离正电荷(原子核)时,会形成电偶极子;将两个(或更多)电偶极子彼此靠近放置会在它们之间产生强电耦合。这就可调制执行多量子比特逻辑运算。

这些电偶极子不需要相互接触但又相互影响。理论研究表明,200 纳米是快速和高保真量子操作的最佳距离。研究人员称,这可能是一个改变游戏规则的发展,其足以允许在量子比特间插入各种控制和读出设备,使处理器更容易连接和操作。

近年来,量子通信领域的大进展和小进步不断,只是大多成果仍处在实验室阶段,离百姓生活还有距离。不过,要有耐心:人工智能领域在沉淀多年之后,推出了 ChatGPT 这样大受用户欢迎、震惊整个科技界的里程碑式应用。在不远的将来,量子通信领域会不会也酝酿出类似的“杀手锏”级应用,让看似玄妙难懂、遥不可及的量子比特飞入寻常百姓家?这是很有可能。

