



人工智能：让机器拥有学习能力

——访日本通讯与网络专家吴剑明博士

本报驻日本记者 葛进

■新视野

最近,国内外掀起了一阵讨论人工智能的热潮,其背后主因是以谷歌和脸书为代表的高新技术企业广泛应用机器学习以及深度学习技术开始显现出了良好的效果。特别是在语音识别和图像识别两个领域,这几年的提升非常快。谷歌在2014年将语音识别的精度从2012年的84%提升到98%,而脸书的人脸识别精度在短短几年里甚至超过人类肉眼的水平。

与此同时,各大企业开始了一场抢夺人工智能人才的大战,比如谷歌设立了“谷歌大脑计划”,并在2014年初斥资5亿美元收购了人工智能创业公司DeepMind。国内的搜索巨头百度也不甘示弱,高薪从谷歌挖来了斯坦福大学人工智能专家吴恩达负责“百度大脑计划”。那么,人工智能到底是什么?其经过了什么样的发展历程?未来的人工智能到底会是什么样子?就此,科技日报记者采访了日本KDDI研究所研究员、通讯与网络专家吴剑明博士。

人工智能是如何出现的?

吴剑明博士称,人工智能的概念早在大约60年前就出现了。在1956年美国达特茅斯大学召开的世界上第一次人工智能会议上,由当时就职于麻省理工学院的约翰·麦卡锡教授提议并推动后,人工智能概念就成了一个热门话题。不过在这60年里,人工智能的发展几度起起伏伏,有过红红火火的日子,也有过被打入冷宫受尽白眼的岁月。

那么到底什么是人工智能呢?为什么人工智能之路会如此的沧桑坎坷呢?吴剑明说,顾名思义,人工智能就是让机器拥有学习能力,甚至能够像人类一样去思考工作。在人工智能的早期,有些科学家非常乐观地认为,随着计算机的普及和CPU计算能力的提高,人工智能的实现指日可待。不少早年的科幻电影描述,到了2000年,机器人无所不能。但事实证明,人工智能的发展没有预期的那么美好和顺利,因为人类对脑机制的理解还是有道迈不过去的坎儿。

一般人看来,人工智能属于计算机科学的范畴,但实际上人工智能覆盖了计算机科学、数学、心理学、哲学和语言学等几乎所有自然科学和社会科学学科,要想有所突破,仅仅靠精通计算机的专家学者和技术人员是远远不够的。幸运的是,强攻不果,迂回之战却有了突破,近年来随着机器学习以及深度学习技术的成熟,人类离人工智能的梦想实现还是实实在在地迈出了一大步。

发展历程:从经典数理逻辑到专家系统

人工智能从出现发展到现在,经历了一个比较曲折的过程。

吴剑明指出,人工智能从1956年概念建立至今,最初的30年是逻辑学派占主导地位,主要是因为逻辑推理和启发式搜索在智能模拟中避开了当时知之甚少的大脑思维规律中深层次的复杂问题,利用了比较成熟有效的逻辑学方法。这在当时计算机性能不高、普及率也很低的条件下显得简单灵活、见效快,是一个捷径。通过计算机实现的人工智能很快在定理证明、问题求解、模式识别等关键领域取得了重大突破,崭露头角。于是早年的科学家们乐观地认为,人工智能与传统计算机程序的本质差别在于它能够进行逻辑推理,依靠逻辑推理定理的完善和计算机的高性能CPU以及大容量存储设备的普及,可以在不久的将来完全解决智能模拟问题。

吴剑明说,逻辑学派撇开大脑的微观结构和智能的进化过程,单纯利用程序或逻辑学对问题求解的过程来模拟人类的思维过程,所以也被分类为弱人工智能。这种方法专注于建立被解问题的数学模型,即找到该问题输入和输出之间的数量关系,把它转化为一个数学问题,然后找到用计算机实现该数学问题的解决算法。然而经过对经典数理逻辑理论解决智能模拟问题进行深入研究后,科学家们才发现这条路是走不通的。主要原因在于,人工智能中的推理和搜索存在组合爆炸问题。也就是说,计算时间与问题的复杂度成几何级数正比,绝大部分人类的思维过程仅仅靠计算机的高速计算能力是无法模拟和解决的。吴剑明举了个通俗的例子来解释组合爆炸的严重性:一张纸折叠50次的厚度是多少?很多人直觉会认为是黄页电话簿的厚度。错了,答案是地球到太阳之间距离!这就是数学上几何级数的恐怖之处。此外,人类思维中的绝大部分问题都无法转化为一个数学问题,原因在于人类思维过程中充满了不确定性、

矛盾和演化。而科学家们长期的实验也证明,人类在解决问题时并没有使用数理逻辑运算,人类思考的过程是无法用经典数理逻辑理论进行描述的。

吴剑明称,事实证明,通过经典数理逻辑的方法是实现不了真正的人工智能的,科学家需要找到其他办法来解决所遇到的难题。

他说,在此之后的10多年,也就是80年代开始,人工智能进入了专家系统发展的黄金时代。科学家们发现,人类之所以能快速高效地解决各种复杂问题,不仅是由于人有逻辑推理能力,更由于人具有知识,特别是有关领域的专门知识。这时期尽管也属于前面提到的弱人工智能时代,但确定了基于知识的逻辑推理在智能模拟中的重要地位,人工智能开始从实验室走向实际应用。

但在这个时期,专家系统的瓶颈问题也显现了,那就是知识获取的途径一直没有得到良好的解决,主要原因在于不像现在有互联网,有云计算,有无处不在的智能手机,那个时代专家知识库的构建常常是没有完备性和可靠性保证的经验知识,专家学者和技术人员不得不依靠各种经验性的非精确推理模型。而且,人类思维面临的实际问题中,只有很少一部分是可以确切定义的确定性问题的,大部分是带有不确定性的问题。所以当知识工程深入到这些问题

机器学习思想并不复杂,它模拟人类在生活中学习成长的过程,从数据中自动分析获得规律,并利用规律对未知数据进行预测。因为机器学习算法中涉及了大量的统计学理论,所以也被称为统计学习理论。

换句话说,机器学习的本质就是将人的操作/思维过程的输入与输出记录下来,然后统计(又叫做训练)出一个模型用来对新的数据进行预测,使得这个模型对输入输出达到和人类相似的表现,这种方式也慢慢成了现代强人工智能最基本的核心理念。

大家直观地想象一下人和机器的区别在哪里?其实不在于弱人工智能强调的计算能力、存储能力,或是推理能力,任何人和机器最大的区别在于,人类在成长、生活过程中积累了很多的历史与经验,人类会定期对这些经验进行“归纳”,获得生活的“规律”。当遇到未知问题时,人类会使用这些“规律”对未来进行“推测”,从而指导自己的生活和工作的,甚至去创造新的东西出来。我们老祖宗说得对,“以史为鉴,可以知兴替;以人为鉴,可以明得失”,这与机器学习的思想是极为接近的。机器学习中的“训练”与“预测”过程,也可以分别精确地一一对应到人类的“归纳”和“推测”过程。

吴剑明说,随着现代互联网和个人计算机的普

工的专业知识和经验,有的时候能不能选好甚至需要一定的运气。由于这部分人工操作对最终算法的准确性起到非常关键的作用,不但非常消耗时间和精力,且如混入一些模棱两可或是错误的的数据,那么很可能会前功尽弃,事倍功半。

既然手工选取特征不太好,人类也不可避免的有主观偏差,那么能不能自动地学习一些特征呢?吴剑明指出,深度学习就是用来干这个事情的,它的一个别名Unsupervised Feature Learning,顾名思义,Unsupervised的意思就是不要人参与特征的选取过程。

吴剑明介绍说,这类研究最早起步的契机是,研究瞳孔与大脑皮层神经元的对应关系的科学家们发现了一些有趣的现象,人眼和大脑合作识别看到物体时,通过神经元互相合作很可能有一个分层次识别过程。具体的讲,从原始信号摄入开始(瞳孔摄入像素Pixels),接着做初步处理(大脑皮层某些细胞发现像素色块间边缘的局部变化特征),然后下一步抽象(大脑皮层判定眼前的物体的形状、颜色、质地等),最后再进一步抽象(从周围的场景、物体和物体间的空间位置关系等等)确定识别的物体。

深度学习正是运用了类似的分层次抽象思想,更高层次的概念从低层次的概念学习得到,而每一层都

幅度降低了多层神经网络训练上的计算量,减少了训练偏差,和传统机器学习相比,优势明显。从此,神经网络重新成为了机器学习界中的主流学习技术。神经网络改头换面,为深度学习开启了学术界和工业界的新浪潮。

吴剑明指出,深度学习的实质,是通过构建具有很多隐层的机器学习模型和海量的训练数据,来学习更有用的特征,从而最终提升分类或预测的准确性。区别于传统的机器学习,深度学习的不同在于:首先,强调了模型结构的深度,通常有5层、6层甚至10多层的隐层节点;其次,明确突出了无监督特征学习的重要性,也就是说,通过逐层特征变换,将样本在原空间的特征表示变换到一个新特征空间,从而使分类或预测更加容易。与机器学习所需要的依靠人工建立特征的方法相比,利用大数据来自动提取学习特征,是我们朝着真正的人工智能迈进的又一步。

在另一方面,和传统人工神经网络的迭代训练需要过于复杂的计算量不同,深度学习并不同时训练所有层,辛顿教授提出了更为有效的降低训练上的计算量和减少训练偏差的方法。简单的说,就是自底向上每次只训练一层网络,通过非监督学习“逐层初始化”(layer-wise pre-training)网络,当所有层训练完之后,再自顶向下反向调教(back propagation)优化。打个比方,人类的大脑发育初期,大脑每一部分的职责分工并不是明确的,我们对外界事物的理解由浅到深也是出自于本能或是实践,而去了学校学习后则可以通过后期的教育来纠正自己错误的认识,进而对事物有更为体系和深入的理解。

即使这样,深度学习也是需要很大的计算量的,好在近年来,得益于计算机速度的提升、大规模集群技术的兴起、GPU的应用以及众多优化算法的出现,耗时数月的训练过程可缩短为天数甚至数小时,深度学习终于在实践中有用了武之地。

未来发展:通过仿生学思路实现突破

吴剑明指出,目前世界各科技发达国家人工智能的发展路线不尽相同。和其他科技领域有些相像,在人工智能行业,一直是欧美在带头创造新理论,而日本则在改良和一直上工夫。相对来说,日本在硬件方面,比如机器人的精密机械制造、机器人仿人运动方面有优势,但在关键的人工智能理论方面,原创或是原创的理论很少,所以不管在学术界还是在产业界,目前看来还不如欧美,特别是美国那么活跃。此外,可能和日本经济的长期不景气有关,近年来一些日本大公司对于研发投入的决心和长期眼光还不够。比如在语音识别领域,由于谷歌走在了前面,一些日本大公司干脆就放弃了自家的语音识别技术转而使用谷歌的技术。但这样一来,人工智能的关键人口和背后的数据就被谷歌夺走了,这也导致在人工智能领域很难再翻身超越。

至于哪个路线的未来性更大,吴剑明表示,尽管机器学习和深度学习在语音、图像、文本识别上有了长足的进步,也让计算机变得聪明智能了很多,但和人类所具有的智能相比,仍然有本质区别。比如计算机能精确识别人脸、物体,但是却无法识别如张三闯红灯、李四喝醉了这样更为抽象的场景。

人类具有丰富的联想能力、理解能力、创造能力,要实现这些能力而又不过于建立人类类似的机制,就会绕很大的弯路,几乎是不可能的。要真正实现人工智能,必须借鉴人脑先进结构和学习思维的机制,再通过深度学习这样的方法进行规模、结构和机理上的模拟,通过仿生学思路实现人工智能的突破。

当然,人类是从低等生物经历几十亿年,在地球生态圈这么庞大的空间中进化而来,要获得人类这样充足的进化时空环境几乎是不可能的。在没有完全弄清大脑原理之前,通过模仿部分人脑原理来逐步渐进,可能是比较现实的办法。比如谷歌在收购DeepMind之后明确表示,不会首先将其在收购人部门,而是先从基础的语义识别开始。而百度也是将深度学习技术应用在具体的用户服务方面,比如说提高中文语音识别率,完善图像识别能力。所谓循序渐进,按部就班,就像人类一样有五感才会有思考,把人工智能神经网络底层的学习水平给完善了,才会有更抽象的高层的学习水平的突破。从这点来看,应该对深度学习未来进一步的理论发展充满希望。

吴剑明表示,深度学习运用在各个单独领域比如声音识别、图像识别时已经得到了很好的效果,当计算机不需要被事先告知明确数据含义,下一步能够融合吸收和理解各个部分的数据,起到1+1>2的作用时,人工智能也许将会被真正兑现。



时,经典数理逻辑的局限性不可避免地暴露出来了。尽管弱人工智能时代使人工智能理论有了长足的发展和进步,但离实用还有比较大的距离,一直到2000年后机器学习和深度学习的出现,科学家们才发现终于找对了方向。

机器学习:其实是“旧瓶装新酒”

吴剑明指出,如果2000年前是弱人工智能时代,那么2000年之后就可以称为强人工智能时代。

吴剑明说,与弱人工智能相比,强人工智能在最近的十多年里慢慢成为主流。强人工智能又称仿生方法,这个方法认为,人工智能应该专注于模拟人脑的结构机制,也就是说通过计算机模拟人的思维过程,应该通过研究神经元的互相合作机制,而不是逻辑学运算过程。这一学派重视对人脑模型的研究,谷歌就是强人工智能技术的推土机之一。

吴剑明说,在历史上,如果我们想让计算机工作或是按照弱人工智能的方式运作,我们会给它编好一段段的指令,然后计算机就会遵照这个指令忠实地一步步执行下去,或是按照事先制定好的知识逻辑公式推导下去。有前因才会有后果。但这样的方式仍然只属于机器的范畴而不是人工智能。强人工智能和弱人工智能的区别就在于,它们不是接受事先安排好的指令或是逻辑推理,而是从输入的数据里自己发现事物的规律。

吴剑明所说的推动强人工智能迅速发展的机器学习技术。

他认为,近年来大放异彩的机器学习其实是“旧瓶装新酒”。早在1956年人工智能概念出现后不久,就有了对机器学习的研究,但之后迟迟没有进展。和前面提到的专家系统类似,原因在于那个时代知识或是数据获得的途径非常少,难度以及成本又非常大。

及,以及智能手机的崛起,海量数据和知识的获得变得非常容易和低成本了,这也直接促进了机器学习的飞速发展和实用性的迅速提高。当训练数据的样本足够大,学习算法方向对的话,就有望接近极限,达到甚至超过人类的能力。比如前几年“谷歌大脑”通过从网络上的一千万张图片学习建立各种各样的猫的品种、颜色、姿势和拍摄角度等特征量,然后对于任意一张图片,它就可以从中把“猫”准确地识别出来。

深度学习:技术进步使之终有用武之地

吴剑明指出,现在除了机器学习,人工智能还出现了一个叫“深度学习”的概念。深度学习是机器学习研究中的一个新的领域,和机器学习相比,它让人工智能又前进了一步。深度学习在机器学习的基础上进一步深入模拟人脑进行分析学习的神经网络,它模仿人脑的机制来训练和预测数据,例如图像、声音和文本。

2012年6月,《纽约时报》披露了应用深度学习的“谷歌大脑计划”,吸引了公众的广泛关注。这个项目的主导之一就是斯坦福大学人工智能专家吴恩达。这一项目用16000个CPU Core的并行计算平台,训练一种称为“深度神经网络”(DNN)的机器学习模型(内部共有10亿个节点),在语音识别和图像识别领域获得了巨大的成功。

吴剑明介绍,和大多数机器学习技术需要先给训练数据提取特征做标注不同,深度学习直接把海量数据投放到算法中,系统会自动从数据中学习。比如“谷歌大脑”识别猫的算法,在训练数据的时候不用告诉机器“这是一只猫”,深度学习系统自己找到了什么是“猫”这个分类。

机器学习所需要的提取特征做标注,其实需要人

自底向上,对没有人工标注的数据进行学习,最后再用人工监督自顶向下反向进行调优。这一点也为深度学习赢得了重要的优势。

吴剑明表示,现在欧美、日本,包括我们中国的学术界都对深度学习非常关注,深度学习的威力目前在语音识别和图像识别上得到了很好的验证。不过在自然对话、自我进化机器人等人工智能更高深的领域里,它的效果还有待进一步考察。有意思的是,深度学习也是“旧瓶装新酒”,它的思想其实来自于上世纪80年代成熟的神经网络技术(ANN)。神经网络同样包含输入层、输出层以及中间的若干隐层,每层都有若干结点及连接这些结点的边,在训练数据集上会学习区分超平面,从而建立模型。但后来科学家们发现,当时科学家们发现的人工神经网络实用性很差,究其原因,是由于巨大的计算量使得人工神经网络只能包含少许隐层,从而限制了性能。花费了大量人力物力之后,科学家们发现,只有少数几个特殊场景可以成功应用。所以到上世纪90年代开始,神经网络失去了关注和经费,成为了食之无味、弃之可惜的鸡肋行业。

但为什么一个上世纪90年代被放弃的技术又重新回到了万众瞩目的地位呢?因为80年代虽然理论基础完备,但真正到达实用所需要的数据和计算能力都不具备。近年来,随着互联网的发展,计算机硬件的价格下降,以及谷歌这样“怪物级”高科技公司的存在,以前高不可及的困难也终于有了解决的可能性。另外,功夫不负有心人,深度学习领域最重要的科学家,多伦多大学的辛顿教授带领的团队一直没有放弃对神经网络技术的研究,2006年他在《科学》上发表了一篇文章,解决了神经网络在计算上的难题,同时也说明了深层神经网络在学习上的优异性。辛顿教授提出的新理论大