

人工智能芯片的现状与趋势

——中国科学院物理研究所“人工智能”主题讨论侧记

2019-07-04收到

† email: hxwei@iphy.ac.cn

DOI: 10.7693/wl20190710

“我设想在未来，我们可能就相当于机器人的宠物狗狗，到那时我也会支持机器人的。”

——克劳德·香农

2019年7月1日晚，由科技部引进国外智力管理司、中国科学院科学传播局、北京科学技术委员会支持，中国科学院物理研究所承办的第39期科学咖啡馆活动在物理所M楼咖啡厅举行。科技部引进国外智力管理司邱成利处长主持活动。

科普活动主讲人邀请的是清华大学电子系电路与系统研究所教授、深鉴科技联合创始人汪玉。汪教授长期从事人工智能芯片方面的研究并创办了科技公司，从学术界和产业界的不同角度为大家解读了人工智能芯片的历史与发展。

人工智能芯片的发展之路

人工智能，随着以谷歌的AlphaGo为代表的一系列新闻进入大众的视野，成为了很多人眼中“好

专业”“赚大钱”的方向。人工智能的发展自然与硬件——也就是人们常说的芯片分不开。芯片对于大多数人来说可能既熟悉又陌生。人们每天使用的电脑、手机都离不开它，但是芯片的结构、算法等概念却并未普及大众。汪玉教授的团队主要从事人工智能芯片的深度学习及加速方面的工作。

深度学习的发展主要有两个代表性的时间点。一是2006年在*Science*发表的一篇文章，证明用多层的神经网络可以构造出更好的拟合器。这篇文章告诉我们，通过增加“隐参数层”的数量可以使得问题大大简化。也就是说，从2006年起，深度学习正式出现，神经网络真的变“深”了。

另一时间节点大概是在2009年，斯坦福大学的李飞飞教授集结来自世界各地的一波学生标记了几千万张图。这意味着计算数据集的极速扩大，使计算的准确率达到了72%。2012年的时候又引入了神经网络，准确率一下提升

到了85%。

随着正确率的提高，人们也把视线逐步转向了效率方面。比如战胜李世石的AlphaGo，当时用了几千个CPU与GPU，但最后一次和人类下棋的时候只用了几台服务器和一个带CPU的机器。从“好”到“快”，构成了人工智能芯片发展的另一个维度。

人工智能发展最核心的技术始终是芯片。我们经历了CPU、GPU、手机CPU三个时代，下一个到来的将是人工智能芯片时代，其需求从市场的角度来说是在增容的。PC时代的联网需要网线，然而在不发达国家的局限性使得网络使用不会超过全球人口总数。手机时代由于基站覆盖到了很多不发达国家，数量可以与全球人口相当。而人工智能时代讲求的是“万物互联”，市场规模可能会达到全球人口数量的好几倍，甚至十几倍。因此这是一个可预见的巨大市场。

对于应用来说，速度自然非常重要。目前的加速手段主要有三种：第一种是把晶体管做得越来越小，现在的3—4 nm，已接近了物理极限。第二种是定制加速器，比如GPU就是利用多核定制对图像的



汪玉教授主题报告现场

加速。第三种是通过新的底层物理结构来做计算。比如要计算非线性，原来的方法是利用一堆公式转化为数字电路的通断算出一个曲线，但是光本身就是非线性的，如何利用这个物理机制直接计算，是我们应该思考的问题。

现在不仅传统芯片公司，几乎所有的大公司都在做人工智能芯片。做芯片最难的其实是定义芯片的用途。芯片的通用性和效率往往不可兼得，最通用的CPU能干所有的事，但如果所需要的只是一项特定任务，定制芯片的效率才能达到最高。目前芯片生产已不再是问题，每个大公司都可以根据自己的需求量身定制芯片。

除了芯片的效率，其位置也有讲究。现在很多公司推出了“云端”，也就是云+端的组合，可以解决手机算力不够的问题。相对于云和端，还有一个介于两者之间的“边”。比如一栋楼里的服务器，或者手机基站这些信息集中的地方，其算力虽不如“云”，但是离“端”更近，传输速度更快。如果用“边”来做计算，原则上只要传输速度够快，手机不需要芯片就可以完成各种任务。

对于中国来说，在芯片的原料、制造、系统软件等需要时间积累的方面还差得很多。由于时间所

限，汪玉教授没能深入展开话题，在大家的意犹未尽和掌声中结束了这次主题报告。

中国芯，发展路在何方

在汪玉教授精彩的科普报告之后，现场嘉宾纷纷踊跃提问。科技日报记者刘园园有兴趣详细了解中国芯片发展的现状，汪教授告知中国做得最好的是应用软件，甚至美国人也在学习我们的app。但是芯片的系统做得不够，我们虽然也有华为的鸿蒙系统，但是相比于安卓有很大差距。芯片制造的仪器目前主要从欧洲进口，但对方提供的设备比国外落后一到两代，我们缺乏的还有原材料。最基本的原料和制造是芯片发展最大的问题。

关于产学研合作的问题，汪玉教授谈到他之前看过美国工程院的一个报告，反思了为什么从20世纪60年代开始，美国的芯片制造就一直领先。其核心就是产业、企业和学校之间的合作。从启动到形成产业一般要20年的周期甚至更长，因此学校的研究应该和企业有所区别，企业往往看的是1—3年的事情，学校的研究则应该去看5—10年的事情才有可能做出颠覆性的成果。不过这还涉及到联动的问题，学校和企业之间必须要建立起信任关系，这样企业的钱才能有效支持学校的研究。如何让企业相信你做的东西有用，也是非常重要的一个环节。

中国纺织出版社华长印问到可否根据中国目前发展状况，预测一下中国未来物联网的发展。汪玉表示，中国一定是物联网最先发展起来的地方。自2015年起，

我国已经把科技发展作为基本国策，物联网是一个很重要的、可能超过美国的发展方向，而且中国的市场也足够大，一旦有物联网的应用爆发，行业一定会迅速崛起。

人才培养，道阻且长

来自卡耐基梅隆大学的一位博士生认为，中国目前硬件落后的问题和学生选择专业的风气有关。之前软件行业的工资高、岗位多，大家都一窝蜂地往软件跑，最近看到人工智能赚钱，才有人开始选择人工智能。如果风气不改变，中国很难把自己的短板补上去。对此，汪玉回答，从现在经济水平和物价来看，单凭高工资指望过上不一样的生活是不可能的，但是如果你能做到一个行业的前千分之一，那就有希望。所以大家坚持把一件事情做好才是最有价值的。

中国科学院老专家技术中心赵超同样关注专业的选择问题。汪玉结合自身经历谈到，学生们还是要多见识多交流，保持高质量的朋友圈。对于大趋势的判断不会写在书本上，都在人的脑子里，多交流才能得到答案。另一方面，每5—10年可以考虑转换一个方向，由内而外的求变也是紧跟时代潮流的重要条件。

在嘉宾们的热烈讨论中，咖啡馆活动接近尾声。人工智能是一个很宽泛的话题，芯片的发展又是其中最核心、最基础的部分。虽然我们在芯片的制造上与发达国家仍有差距，但随着国家的大力支持和中国科技的不断发展，相信在不远的将来，我们将会在世界芯片的舞台上拥有自己的一席之地！

(中国科学院物理研究所

姜畅 田春璐 魏红祥 供稿)



科普活动与会嘉宾合影