# 要重视AI基础理论的创新

## ——记姚期智院士在2020世界人工智能大会上的演讲

◆记者 蔚子 / 文


7月9日，图灵奖获得者、中国科学院院士姚期智在2020世界人工智能大会开幕式上发表主题演讲。他强调：要重视AI基础理论的创新，并指出了人工智能理论研究的三个新方向。

在题为《人工智能理论的新方向》的主题演讲中，他说："AI在现实世界已经有了广泛应用，在这场大会中也能看到AI应用的新进展。但是我想说明的是，所有的这些进展都来自于基础科学。也就是说，AI领域在很多年前就已经打下了理论基础。这给我们的启示是，一定要让理论研究不断发展。"

他指出：当前AI面临的很多问题和挑战，都可以用理论来进行分析。通过理论分析，我们能更清楚地知道我们面临的挑战的本质，以及解决这些挑战的方法。同时，他还指出了"AI是跨学科的行业"这个特点。他说，当前在AI中获得的一些成果，其所处的领域很多是和AI几乎不搭边的学科。

在这次演讲中，姚期智着重阐述了人工智能理论研究三个新方向：一是神经拓扑结构，是神经网络研究的新视角；二是与密码学息息相关的隐私保护学习；三是可控的超级人工智能，要研究如何设计有益的超级智能。

关于"神经拓扑结构"，他追问了"深度学习和神经网络在很多年里都没有进展，后来却意外地在AI应用上获得了成功。可是神经网络的力量到底来源于哪里？"的问题。他说，神经网络如此强大的理论原因直到现在还是个迷。如果能找到原因，这对神经网络的改善、应用无疑是个突破。

他还举例说明了神经网络的应用。他说，气象图数据是"波动"的，现在需要"算法"分析图片中展现的气候形式，即分析图片中的天气是暴风还是正常。人类工程师通常观察气象图的二维或者三维的表现，然后从中找到范式，判断是否符合风暴的特点。而机器学习的标准做法是通过处理高维度的数据，将数据分为不同的数据集类别。这个数据集可能是暴风天气的图像，也可能是正常天气的图像。而高维度数据集正是数学家关注多年的一个方向，拓扑学中贝蒂数的概念就可以

应用于高维度子集。

他说："神经网络的核心问题是，如何设计神经网络的深度和大小，才能够对数据进行分类。把数学和神经网络的能力结合起来，有助于我们了解什么样的数据集更便于神经网络去识别。这也给了我们一个启发，解决AI问题的方法可以从其它学科的角度考虑。"

二是"隐私保护"，这也是当前人工智能理论研究主要的新方向之一。姚期智认为，人工智能+MPC（多方计算）学习的现行技术，让密码学的应用有了实现的可能。"通过多方密码的做法，我们可以将多个当事方的数据结合在一起交给AI来挖掘价值，但过程中又不会暴露数据的来源。这样就能在实现高质量学习的同时保护数据隐私，对金融科技、药物研发非常有用。"

他说，隐私保护与密码学息息相关。例如，他曾在1982年提出的安全多方计算就是相关理论方向。多方计算主要研究在私有信息不被泄露的前提下，多个互不信赖的参与者如何协作进行计算。通过使用多方计算，多个数据库可以联合计算一个函数却不会透露各自的数据。

三是通用的超级人工智能何时到来？这是人工智能理论亟待研究的主要方向之一。

关于这点，姚期智很直接地说："我的答案是不可预知。"他解释说："因为现在的AlphaZero、人脸识别虽然很牛，但仅适用于特定领域。正如1977年，John McCarthy曾经说过'我们需要概念上的突破，1.7爱因斯坦+0.3曼哈顿项目，可能需要5～500年时间'。"

他说，最新的"超级AI理论"的提出，是在2019年。他介绍了伯克利大学Stuart Russell的书。书中提到，虽然超级人工智能不知道什么时候到来，但是我们必须做好准备。在书中他设定了三个原则，每一个原则都要用严格的数学方法来实现。这三个原则分别是：1.利他的——人的利益凌驾于机器利益；2.谦卑的——机器不能自以为是；3.尽心的——机器能学懂人的偏好。此外，Stuart Russell还提出了许多方法论，涉及概率理论和博弈论。

姚期智最后说："我想表达的是，现在的AI应用来自过去的理论研究，AI的一些进步也正来自跨学科领域。今日的理论探索，正为未来的伟大应用奠定良基！"