

器人的机会很容易得到，期刊可能会发现完全由聊天机器人撰写的伪造论文巨增，这可能确实是一个问题。这样你就会发现聊天机器人写论文，聊天机器人读论文，完全绕过了人类。就像任何事情一样，如果滥用到一定程度，那就是一个大问题。

学术出版社要求作者记录他们如何使用聊天机器人的政策，你认为这个主意好吗？

我们从不声明计算机如何进行计算，那么如果你有一款帮助你写作的工具，为什么要在意呢？如果同行评审过程正常运作，应该能够识别出完全由人工智能编写的文献综述等内容。如果一位认真负责的科学家在论文中声明使用聊天机器人来改进他们的引言，我不知道读者在阅读作者声明时应该怎么办，除了“好吧，我很高兴它提高了论文的可读性”。

对于母语不是英语的人来说，聊天机器人会有很大的帮助作用吗？

绝对是的。这无疑是个好事，

我希望它能够改善与来自缺乏良好英语教育国家的科学家之间的交流。它也可能帮助母语为英语但表达能力差的人。毕竟，某人拥有一个出色的想法，但没有得到妥善表达是令人非常惋惜的。我认为在学术写作中，当然也在物理学中，论文的可读性比研究的质量差很多。

聊天机器人也可以帮助跨学科研究吗？

可以。如果我和与我有共同研究兴趣的 chemist 讨论我的工作，但我使用了他们不理解的术语，可能会使他们理解我的工作非常吃力。因此，我可以对聊天机器人说：“我为一个物理学会议写了这个报告摘要，你能不能用 chemist 容易理解的语言改写一下？”但是，我始终要保持谨慎，因为这可能会改变我所写内容的含义。

聊天机器人可能存在多样性问题吗？

原则上是的，但我们必须谨慎对待。这取决于训练机器人模型的数据以及该数据集是否具有潜在的

偏见。但这并不意味着它没有潜在用途。例如，你在写作中没有使用包容性语言，你可以对聊天机器人说“把稿子改得更具有包容性”。我总是保持警惕，避免完全依赖聊天机器人来完成这项工作。我们都有潜意识的偏见，不应该忽视它们而依赖于人工智能来解决这个问题，因为聊天机器人并不总是能帮你解决这个问题。

你对聊天机器人的未来持乐观态度吗？

是的。我认为最大的优点是提高学术写作的清晰度。你可以编写实验室指南并将其交给聊天机器人，询问指南是否清晰或需要改进。这有点像你在交给学生之前，向 100 人或 1000 人询问意见。通过这样做，你希望可以使指南更加易于理解，从而使学生可以专注于物理学习。同样重要的是，我们现在只处于这些聊天机器人的早期阶段，更复杂的聊天机器人一定会出现。

## 在遗忘之前，记忆变得模糊和混乱

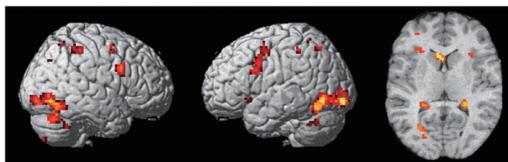
(北京大学 徐仁新 编译自 Philip Ball. *Physics*, January 27, 2023)

大脑信息存储模型揭示了记忆如何随着年龄增长而衰退。

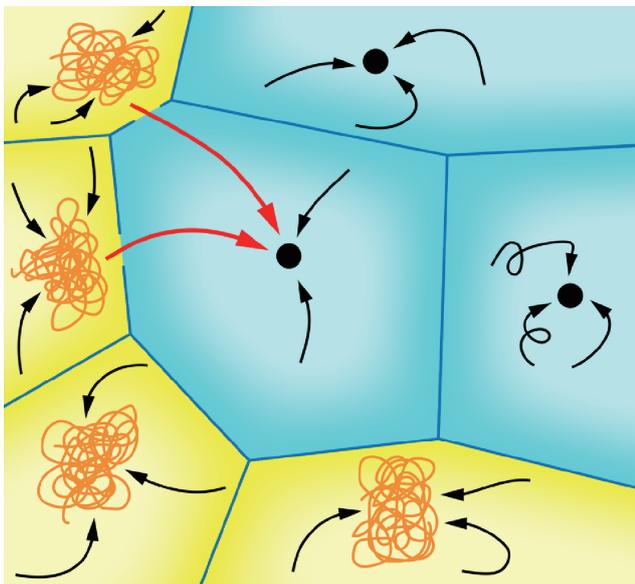
理论上构建的“吸引子网络”(attractor networks)可用于模拟大脑的记忆。基于该网络的一项研究再现了记忆存储并被最终遗忘的过程。计算模拟表明，随着年龄的增长，保留记忆的神经活动图案变得混乱，无法预测，最终成为随机噪声。虽然大脑中是否会真的如此尚

不清楚，但学者建议通过监测记忆过程中的神经活动随时间的演变来研究这一行为。

信号在网络节点(神经元)之间传递会表现出不同图案，是记忆存储和检索的依据，人工神经网络和



从三个不同方向叠加而成的大脑磁共振图像。它显示了在短暂记忆测试过程中大脑的激活区域。大脑模型的网络结构模拟表明：记忆随时间推移而消失，表征记忆的脑图案会变得混乱



吸引子网络动力学示意图。由边界分开的区域代表吸引子状态。刚编码的记忆(蓝色区域)对应于稳定的吸引子,其动力学由图中的“点”所示意。旧记忆变得混乱:那里的动力学并不完全重复(黄色区域)。一些混沌吸引子随着记忆老化而逐渐消失。部分老化的混沌吸引子可能会影响相邻的稳定吸引子(红色箭头),以致不同的回忆线索引发完全不同的记忆

生物神经网络均如此。对于人工神经网络而言,任一时刻某个节点的输出值取决于与其相连的其他节点的输入。类似地,生物神经元放电的概率及频率也取决于对其输入。另一相似之处是,代表突触的节点间链接被“加权”,可以放大或衰减所传递的信号。链接权重取决于两个节点的同步程度,新的记忆可改变权重。

在吸引子网络模型中,节点间交换的信号反映了实际神经元的放电率,并作为输入确定接收神经元的响应。该网络中,这样的信号持续且不断变化地传播。为了在网络中刻上“记忆”,人们将一串二进制数(代表某项记忆)分配给若干节点,然后观察网络的活动性如何随着权重的调整而演变。节点之间传递的信号最终形成一种重复图案,

称为“吸引子态”,也就是编了码的记忆。

若将与记忆有数学关联的新二进制数作用于节点,则可检索记忆,让网络的活动性再次呈现为相应的吸引子态。通常,一个吸引子网络可保存若干记忆,每个记忆对应于不同的吸引子态。因此,网络的活动性在这些状态之间来回变换。

此前的生物神经网络研究表明,网络的活动性比预期的稳定吸引子态更随机。此外,吸引子网络的研究显示可能出现“灾难性遗忘”:如果刻录的记忆状态太多,则无法检索。

看来记忆态并非永恒。纽约大学神经科学家 Ulises Pereira Obilinovic 等研究记忆是如何改变的。他们调整信号权重,让刻上记忆时的初始权重逐渐减小,记忆将最终消失。模拟给出了两类记忆状态。最新的记忆对应于“定点”吸引子,具有清晰而持久的图案,就像围绕太阳的行星轨道。但随着记忆的老化,将转变为第二种类型:混沌吸引子,其活动不精确重复,有点类似天气模式。

随着更多记忆的获得,混沌吸引子的随机性明显增加,直至吸引

子状态消失为噪声。此时不再能检索记忆,记忆完全被遗忘。该研究认为,“遗忘”始于神经网络的活动从规则到混沌的转变,直至一定时间内成为噪声。看来,人们的记忆并非忠实可靠。该模型中没有灾难性遗忘:旧的记忆将自动消失,不存在过载的可能。

吸引子网络理论在大脑存储的研究方面影响很大。在 Obilinovic 等发表的论文(*Phys. Rev. X*, 2023, 13: 011009)中,作者研究了一个稀疏连接的吸引子网络,那里的记忆通过所谓的“Hebbian 突触可塑规则”来训练:新的记忆被不断学习,旧的记忆会被遗忘。他们还构建了一个动态平均场理论来描述记忆随时间的演化,并将该理论与网络模拟作对比,从而给出记忆最大化情形的遗忘时间。可见,在吸引子网络理论中,以记忆的遗忘为代价而不断学习新知识。该研究可以较好地解释大脑皮层中记忆状态的多样性,特别是皮层活动的快速变化。

鉴于旧的记忆存储于较混杂的状态中,研究人员推测,检索这类记忆时神经元放电时间的波动会更大。这一看法可以通过监测记忆后不同时期的神经活动性来检验。

柏林工业大学神经科学家 Tilo Schwalger 认为,这些预言确实可以检验,且相关成果可应用于动物神经网络。葡萄牙学者 Francesca Mastrogiuseppe 对此表示赞同,并补充说该研究“处于理论神经科学两条主要方向的交叉点:一个跟记忆有关,另一个跟大脑中不规则的神经活动有关”。她进一步指出,研究表明这两者有联系。