

二十一世纪计算 | 有意识的图灵机：一种基于认知神经科学的计算机架构

原创：微软亚洲研究院 微软研究院AI头条 前天

▲ 点击蓝字关注微软研究院AI头条



Microsoft Research
微软亚洲研究院

编者按：在1979年之前，对意识的科学研究还是一个禁忌，直到1987年一位诺贝尔奖得主开始研究意识，1990年磁共振成像(fMRI)的发明使人们可以看到脑内血液流动的情况，意识研究才真正成为科学。

意识研究，有哲学、神经科学、结构学等多种不同的范式。对于计算机科学可以对意识理论产生什么样的贡献，卡内基梅隆大学计算机科学系杰出教授Lenore Blum提出了一个可被验证的模型，通过全局分析模式研究大脑的结构，以剧院的类比来描述意识，并通过对大脑结构的研究来理解意识。



Lenore Blum教授在“二十一世纪的计算”大会暨微软教育峰会上的演讲视频

哲学家和认知科学家David Chalmers将理解人类意识的问题分为“简单问题”和“困难问题”。而对于人工智能而言，简单问题是要制造一个能够模拟人的情感的机器人，这类问题可以用认知科学的标准方法来衡量；而困难问题则是制造一个能够感受人的情绪的机器人，这类问题属于经验的问题。

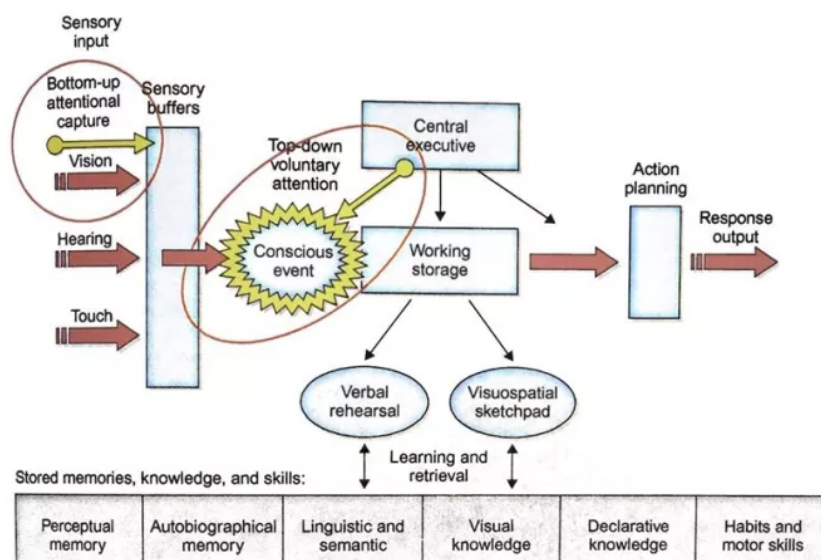
我们的研究提出了一个大脑意识的处理模型，帮助我们区分模拟和真实的体验。



卡内基梅隆大学计算机科学系杰出教授Lenore Blum

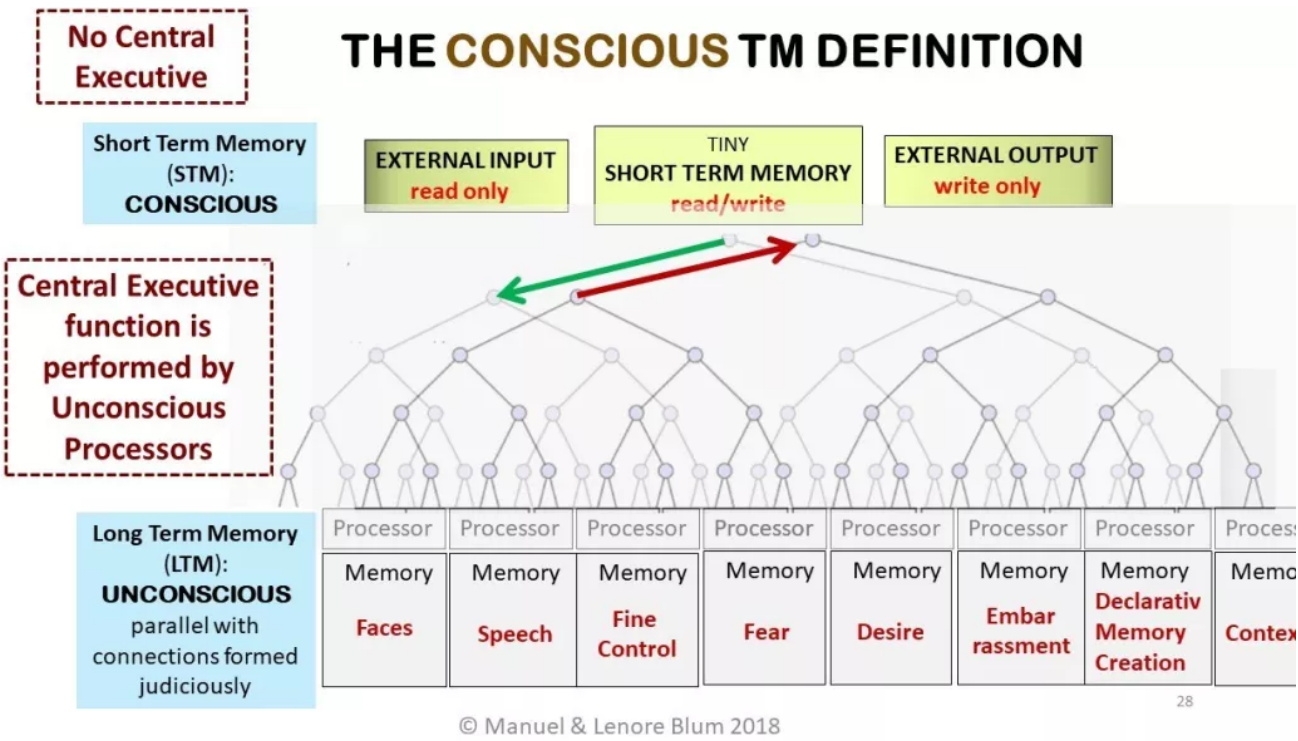
根据神经科学家Bernard Baars对意识的剧场比喻，意识或短时记忆，就像是在剧场舞台上演员的表演，而储存在大脑中的长期记忆则像是坐在舞台下面黑暗中的无意识的观众。在Baars教授的意识模型中，感官从外部接受到的信息由无意识部分进行处理，最终以语言或其它有意识的形态呈现出来。

Bernard Baars' Model of Consciousness



在这个模型的基础之上，我们对模型中的“意识”进行了重新定义，进而构建了一个从意识到无意识的树形结构。

这个模型中没有有一个统一的中央处理器，取而代之的是一个从意识到无意识的二进制树型结构：根节点上是短期记忆的意识处理器，子节点上是大量长期记忆的潜意识的处理器，这些处理器行使中央处理器的功能。不同的处理器有不同的信息传递路径，而作为一种信息的传播方式，它们会通过最短的方式做出最快速的信息传递。



不同感受有不同的传递模式，信息传递的优先级也不尽相同，感受间的协调机制使得某种感受能够最终被人们感知。假设每次只有一个感受涌现，一旦它浮出水面，就可以提前等待其它感受逐渐显现。

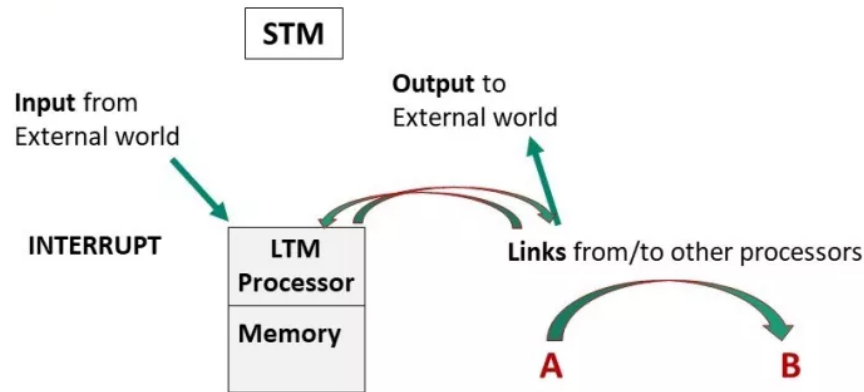
对于感受的变化过程，处理器会选取它接受到信息的积极或负面的权重（疼痛的权重为负，开心的权重为正），不同感受的权重不同，对权重进行叠加会重新形成一个结点，权重值最大的感受信息最先浮出水面。感受信息与不同的感官联系在一起，不同的感受有不同的地址，在输出终端你将会知道这个感受来自于哪里，感受的强度如何。

意识和潜意识之间存在着某种互动，从意识转化到潜意识或从潜意识转化到意识。在长期记忆的处理器中，从外部世界获取的输入经过处理反馈给短期记忆，短期记忆再把信息传递回来，如果传递的过程被打断，就根据权重做分配，最终对外部世界做出反应。

子节点上的潜意识中有很多处理器，数万亿个小神经元都是处理器，它们之间相互连接。如果潜意识处理器A向短期记忆的意识处理器发出一个查询请求，意识处理器将查询请求发布给所有潜意识处理器，潜意识处理器B做出了回答，A和B之间就建立了关联。随着这样的互动地增加，两个处理器之间的连接也会不断增强，有意识的行为也可以慢慢变成长期记忆中的无意识行为。

THE CONSCIOUS TURING MACHINE

Details of the Dynamics: The LTM Processors



- LTM processor **A** links up to **B** when **A** answers **B's** call.
Linking enables conscious processing to become unconscious.
“Neurons that fire together generally wire together.”

但是在现实的一些情况下，即使潜意识的和短期记忆之间的连接被破坏，人们仍然能采取行动。例如盲视现象（Blindsight）：人们由于脑部的损伤不能有意识地看到一些东西，但是即使看不见，人们也会通过其它感受产生动作。比如，潜意识告诉你你饿了，即使你看不到食物也可以找到它。这时，潜意识部分没有通过视觉，而是直接就跳到短期记忆，让它产生行动。

另一类与此相似的现象是“非注意盲视”（Inattention Blindness），这是指人们有时会无意识地忽视某些内容。它的产生是因为一些潜意识处理器对它接收到的信息赋予的权重过高，而另一些处理器对信息赋予的权重过低，导致一些感受就被人们忽视了。

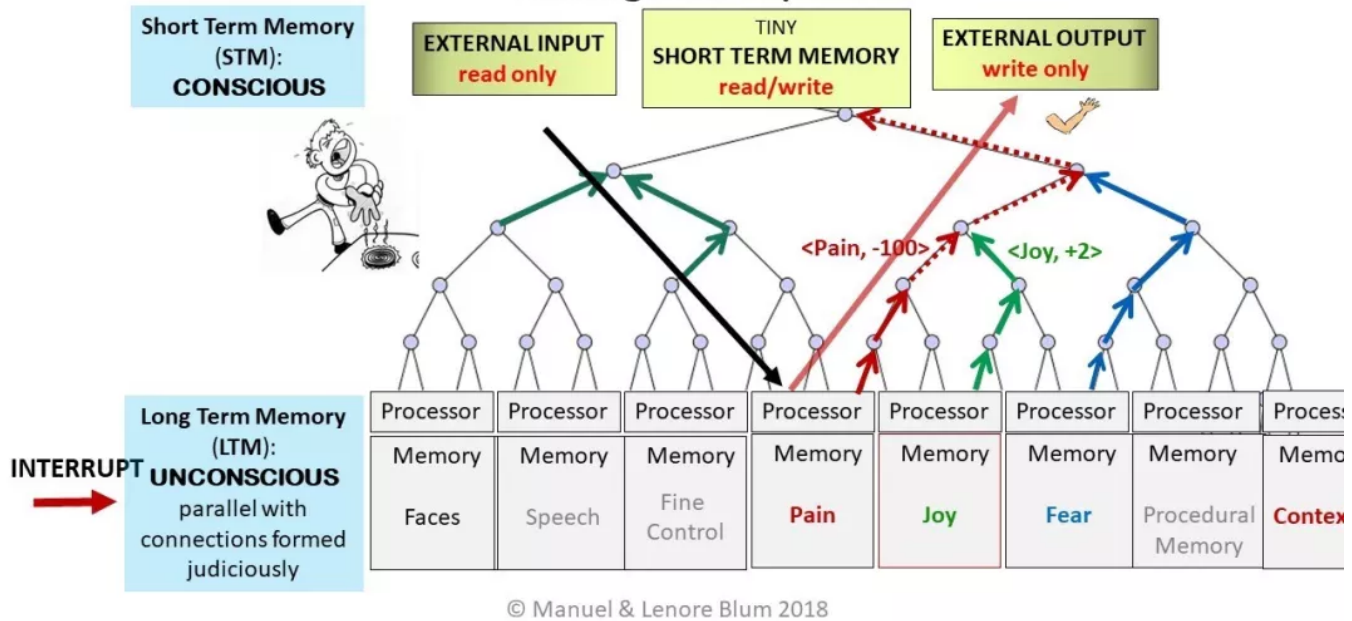
想要解决暂时性盲点问题，机器学习就发挥了重要作用。我们可以在机器学习的算法中加入来自经验的反馈，根据处理器提供信息数量的多少，不断调整不同感知的权重，防止因为处理器提供的反馈信息的多寡而导致的感知偏差，得到一个更加优化的算法。

涌现式的疼痛和快乐

在有意识的图灵机的模型中，疼痛和快乐的感受的出现都是涌现式的，而非通过编程实现的。

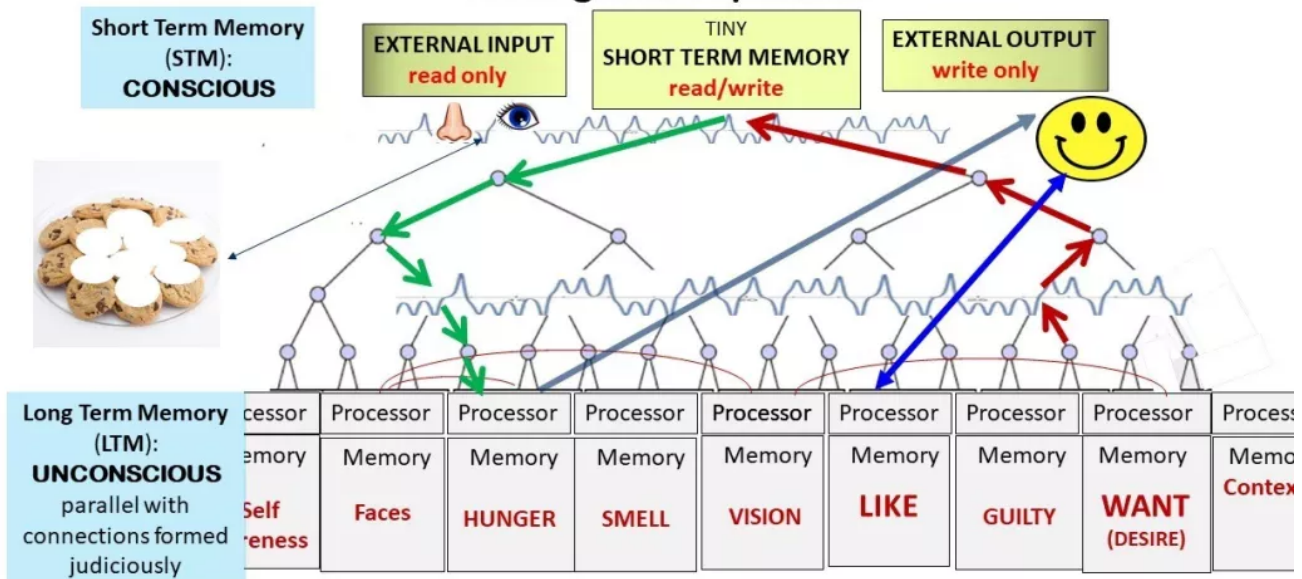
在这个模型中，因为疼痛的权重超过了其它感受，所以和其它因子相比可能会以最快的速度传递到短时记忆中。有时候在极端疼痛情况下，人们只能感受到疼痛。但在其它情况下，除了疼痛之外，人们还可以进行其它行为。就像如果将手放在热炉子上，会瞬间感受到疼痛，但是把手拿开，这样的感受不会进入到人的短时记忆。

THE CONSCIOUS TURING MACHINE Emergent Properties



与疼痛的处理器不同，快乐来自于一种自由和灵活性。有学者将快乐定义为欲望、动机和喜爱的集合，现有的对快乐的研究认为，在人的大脑中有一个愉悦刺激中心，可以通过对某个开关的操控使得人类的快乐感更强，但这并不是快乐本身，而是容易成瘾的动机和欲望。但无论是哪种快乐，调用的神经元都是类似的，而且即使有些人的脑干受到破坏，他仍然能够感受到快乐。

THE CONSCIOUS TURING MACHINE Emergent Properties



但是快乐的感受并不是一成不变的，它与负罪感之间可能存在着某种动态循环。就像如果人们闻到饼干的味道，就会产生想吃它的欲望，吃饼干能产生愉悦的感受，并且这种感受随着食用饼干不断循环。但是当饼干吃完的时候，又可能会因为食用过量而产生一种负罪感。

计算机科学对意识研究的贡献在于对脑部结构的模拟，这不仅能够提高疼痛测量的精确度，还将能提供一种比传统的镜子测试更具普适性的测试意识方法。因为不同感受刺激的大脑部位不同，所以在测试中风或患有幽闭恐惧症的人是否存在意识时，就可以让患者想象不同的场景，观察大脑中被激活的部位，做出更精准的判断。

未来，随着对人类认知了解的不断深入，为人工智能注入认知的研究将在情感、注意力研究、医学、人机协作等领域得到更多的应用。

你也许还想看：

- 二十一世纪计算 | 人工智能如何在行业中落地？
- 二十一世纪计算 | Yoshua Bengio：深度学习通往人类水平AI的挑战
- 二十一世纪计算 | 萨提亚·纳德拉，沈向洋，洪小文，Raj Reddy，丛京生都参与的学术峰会，到底讨论了什么？

感谢你关注“微软研究院 AI 头条”，我们期待你的留言和投稿，共建交流平台。来稿请寄：
msraai@microsoft.com。

Microsoft Research

微软最前沿的科技信息



长按指纹
识别二维码，
一键关注黑科技！