

揭秘天机芯

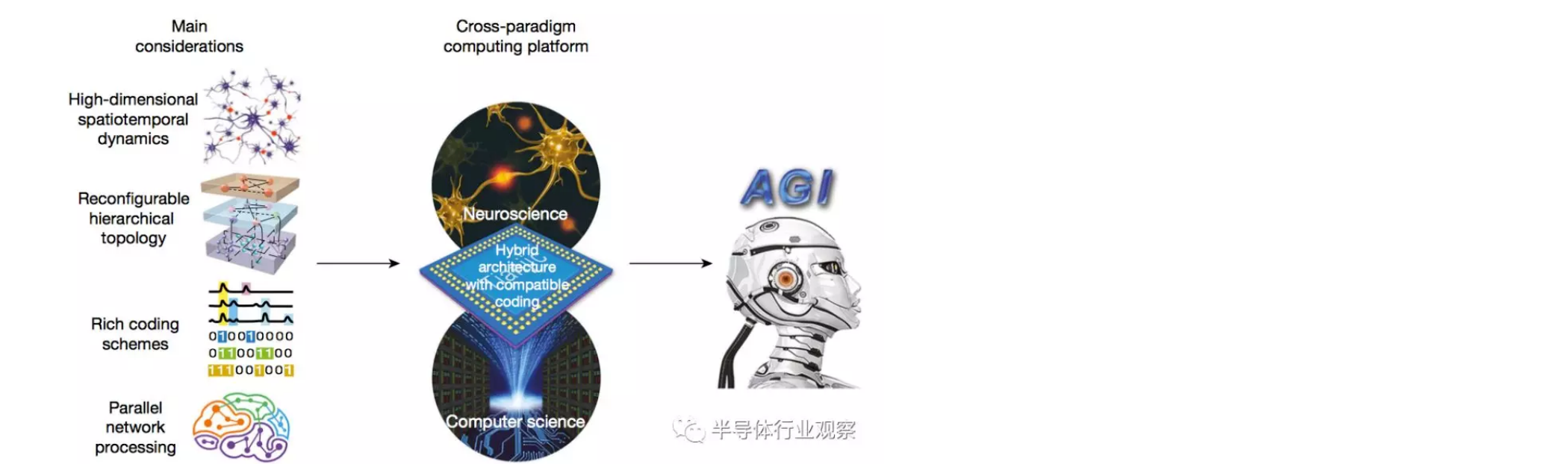
原创：李飞 半导体行业观察 8月15日



近日，清华大学一篇关于将神经模态计算和深度学习计算融合在同一块芯片（“天机芯”）的研究发表在了全球顶级自然科学期刊《自然》上，并且荣登杂志封面。该项研究获得全球顶尖杂志发表，证明了我国在神经模态和人工智能芯片研究走在了世界前列。那么，天机芯到底隐藏了什么天机？本文将为大家仔细分析。

半导体行业观察
Semilnsights.com
七年磨一剑的天机芯

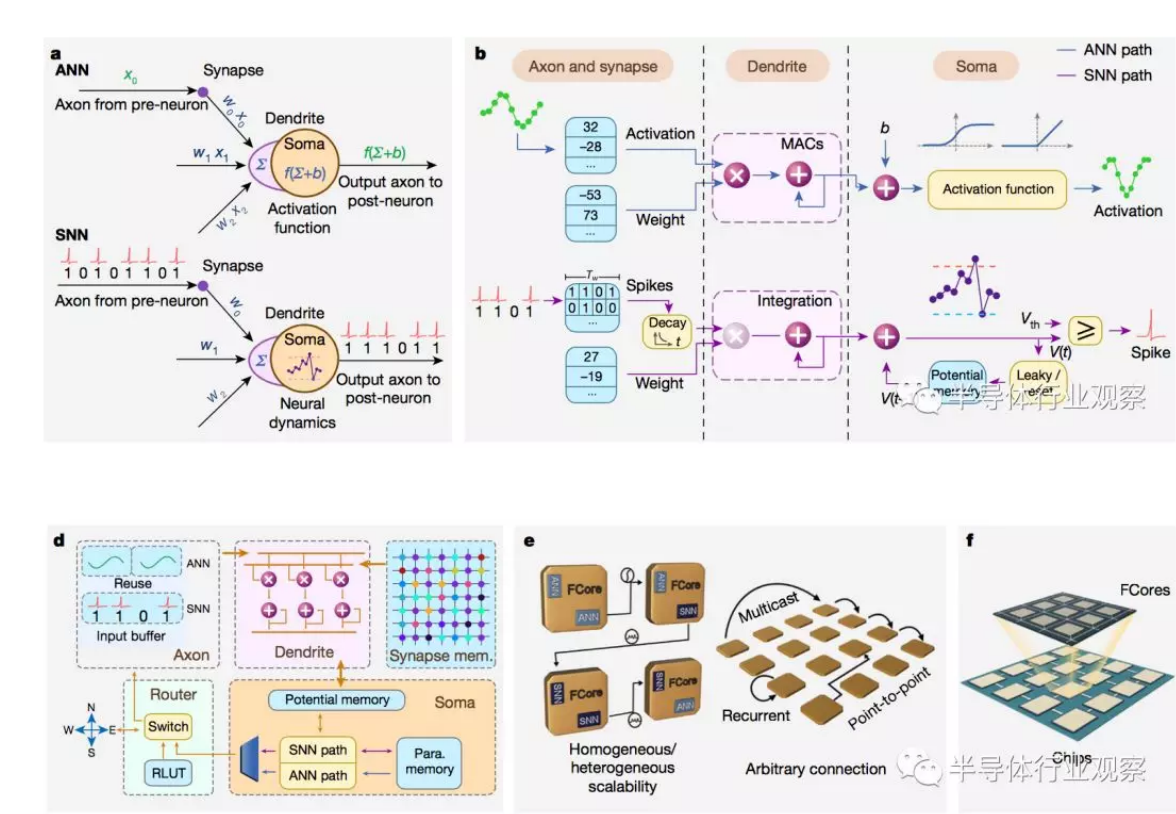
天机芯是清华施路平团队历经七年打磨的芯片，使用28nm工艺流片。该芯片的最大特点是兼容包括神经模态脉冲神经网络、卷积神经网络和循环神经网络在内的多种神经网络同时运行。



目前来说，以卷积神经网络（主要针对图像任务）和循环神经网络（主要针对语音和翻译任务）已经得到学术界和工业界的广泛应用，也已经有不少芯片（例如寒武纪等）问世。另一方面，对于读者来说神经模态脉冲神经网络比较陌生。从原理上说，脉冲神经网络和卷积神经网络／循环神经网络都是在模仿生物神经元之间连接和通信的过程，区别在于卷积神经网络／循环神经网络是使用宏观统计方法来描述生物神经元和突触，而脉冲神经网络则试图真正模拟生物神经网络。在生物学中，一个神经元接收脉冲电荷并改变神经元的内部的电势能，当神经元电势能超过一定阈值后将会发射出脉冲，该脉冲一方面清空了发射脉冲神经元内部的电势，另一方面该脉冲会沿着神经突触进入其他神经元，并在其他神经元中积累电荷改变电势能，如此反复。大量神经元和突触之间形成的网络就是神经网络，而脉冲可以认为是神经元之间传递信息的方法。比较脉冲神经网络和卷积／循环神经网络，两种神经网络中都有神经元以及神经元激活的概念；卷积神经网络／循环神经网络在前馈过程中每个神经元的输出值即该过程中该神经元被激活的次数，而神经元之间的连接强度则用网络权重来表征，每个神经元的输出乘以网络权重再传播给下一层的神经元，因此可以认为卷积神经网络／循环神经网络中的神经网络是使用数学计算的方法在模仿生物神经系统，神经网络并没有真正在发射脉冲而只是在做数学运算。与之相对，神经模态计算的前馈过程中，每个神经元电路模块每被激活一次就会真的发射出一个电脉冲（而不是直接输出整个前馈过程中神经元被激活的次数）并传播给相连的其他神经元，因此可以认为神经模态计算是在真正重现生物神经系统的物理过程。

在天机芯之前，脉冲神经网络和卷积／循环神经网络芯片是截然不同的两种芯片，设计方法完全不同，例如IBM的True North是脉冲神经网络芯片，而寒武纪则是支持卷积／循环神经网络芯片，因此并没有任何一款芯片能同时兼容这两大类神经网络。而天机芯的天机，也恰恰在

于提出了一种新的芯片架构能同时支持两种网络。



具体来说，天机芯架构采用了众核架构，每个核都可以自由配置成脉冲神经网络单元或卷积／循环神经网络单元，总共可以实现40000个神经元。因此，通过将不同的核配置成不同的神经网络单元，天机芯可以实现同时运行脉冲神经网络和卷积／循环神经网络。更重要的是，一个核还可以配置为兼容模式，即接受脉冲神经网络的输入（即电脉冲），并在计算后转化为卷积／循环神经网络的输出（即多比特的数字信号），或者反之将卷积／循环神经网络的输入转化为脉冲神经网络输出。这样一来，天机芯就打通了脉冲神经网络和卷积／循环神经网络的界限，从而允许同一块芯片把两种神经网络融合在一起运行。在应用场景中，天机芯可以实现10倍以上的能效比。

半导体行业观察
Semlinsights.com

天机芯为什么能登上《自然》封面？

天机芯能登上《自然》封面，是因为天机芯有可能成为人类通往强人工智能的重要一步。通用人工智能（AGI）历来是《自然》和《科学》等顶尖杂志的关注要点，之前IBM的True North和Google的AlphaGo也登上了《自然》封面。



天机芯与通用人工智能的最大联系在于它打通了脉冲神经网络和传统深度学习（卷积／循环神经网络）。目前，传统深度学习神经网络在图像识别等领域的表现甚至已经超过了人脑，但是在抽象推理等能力上仍然难以突破瓶颈。另一方面，脉冲神经网络由于和生物机理更接近，

因此可望在神经元数量提升到一定数量后，能实现接近生物神经网络的性能，从而有可能实现抽象推理等高级脑部功能（但是具体是否能实现还需要看学界的下一步研究）。例如，脉冲神经网络非常擅长自适应在线学习，因此可以弥补深度学习对于数据量的需求，而天机芯可以为人工智能学界提供一个极好的研究脉冲神经网络做在线学习的研究平台。因此，通过结合深度学习和脉冲神经网络，天机芯可望能实现进一步的机器智能。

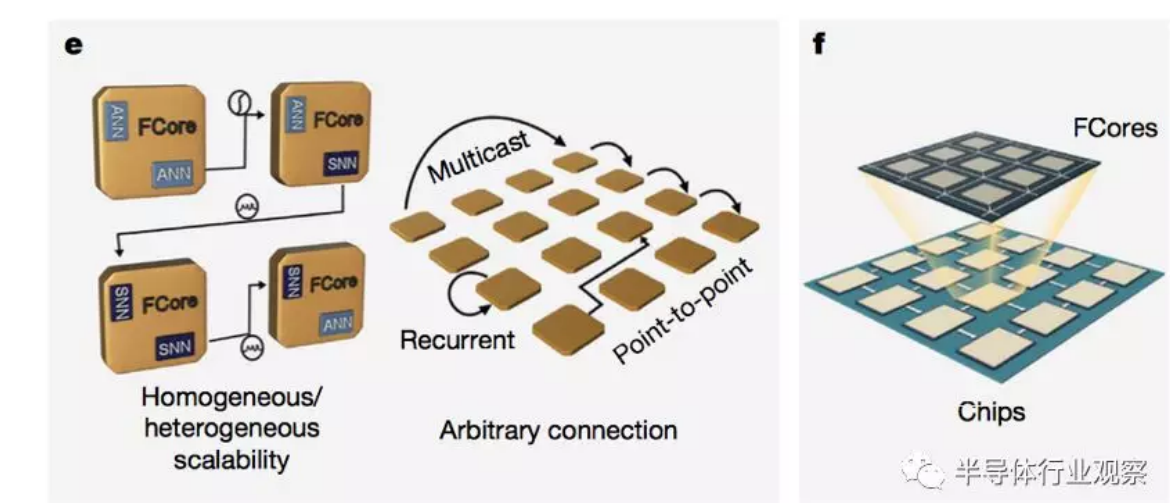
除了实现人工智能之外，天机芯的另一个用途是使用脉冲神经网络来做脑科学相关的研究。随着生物学和医学的发展，脑科学目前已经成为人类亟待攻克的下一个重要课题，而使用脉冲神经网络来模拟和研究脑部工作原理将成为脑科学研究中的重要组成部分。

综上，天机芯打通了脉冲神经网络和深度学习神经网络，因此为人工智能和脑科学研究带来了一种新平台，该平台可望在科学研究中成为下一代人工智能和脑科学的重要基础设施，因此得到学术界认可而登上了《自然》杂志封面。

天机芯的价值所在

天机芯除了在学术领域的重要意义之外，在芯片设计领域也有很高的价值，尤其是对于我国的半导体行业。这里的价值并非是狭义的完成了性能领先的人工智能芯片，而更是证明我国半导体业在下一代处理器架构领域能够跟上甚至引领潮流。

天机芯在中国处理器架构设计方面的第一个重要意义在于我国在众核处理器领域的进步。如前所述，天机芯通过众多核心互联以及互相通信可以灵活实现各种神经网络。这其实是众核架构的典型应用，即芯片上的许多核心可以由编译器和软件来做调度，实现工作量的最优分割。这不仅仅需要芯片设计，还需要编译器技术能跟上，只有在软件和硬件都做好之后才能把众核处理器的性能充分发挥出来。因此，天机芯在众核处理器方面的探索可以说是我国半导体行业的一个重要标志性事件。



除了众核之外，天机芯对于异构计算和可重构计算的探索也走在了全球前列。随着摩尔定律遇到瓶颈，单纯通过半导体工艺节点已经很难进一步推进处理器的性能，目前半导体业界的共识是使用异构计算的方法来进一步提升处理器性能，即在芯片上集成多个针对特定领域做优化的专用处理器，在遇到特定任务时调动专用的核来做处理，这样一来“特事特办”可以实现较高的性能和效率。然而，异构计算也存在一个问题，即在使用专用核来处理相应任务的时候，芯片上的其他部分即处于闲置状态，因此造成了芯片处理能力的浪费，即“暗硅问题”（dark silicon）。为了解决“暗硅”问题，一个可行的办法就是使用可重配置计算，即在不同的应用场景把同一块芯片配置成不同的模式，从而提升芯片针对不同应用场景的处理效率。然而，一旦加入可配置性，往往就意味着性能会相对专用化的设计打折扣。天机芯的设计哲学则很好地体现了在异构计算和可重配置之间的平衡。首先，在众核架构中，每一个核都可以配置成脉冲神经网络核或者卷积／循环神经网络核，这体现了可重配置计算的思想；而在不同的核之间，可以让一些核专门负责脉冲神经网络而另一部分负责卷积／循环神经网络，这又是体现了异构计算的思想。因此，可以说天机芯在下一代处理器架构需要解决的问题上提出了自己独特的思路，在全球走在了前列。

*免责声明：本文由作者原创。文章内容系作者个人观点，半导体行业观察转载仅为了传达一种不同的观点，不代表半导体行业观察对该观点赞同或支持，如果有任何异议，欢迎联系半导体行业观察。

今天是《半导体行业观察》为您分享的第2037期内容，欢迎关注。

推荐阅读

- ★[危机四伏的Lattice](#)
- ★[12nm制程的机会来了](#)
- ★[3D闪存将走向何方？](#)

半导体行业观察



『半导体第一垂直媒体』
实时 专业 原创 深度

识别二维码，回复下方关键词，阅读更多
射频 | 台积电 | 江北新区 | AMD | FPGA | EDA | 集成电路 | MRAM

回复 **投稿**，看《如何成为“半导体行业观察”的一员》
回复 **搜索**，还能轻松找到其他你感兴趣的文章！

**芯片设计服务**

Turnkey · NRE · 咨询/驻场

**供应链服务**

流片 · 封装 · 测试 · IP云

**企业服务**

IT/CAD · 孵化 · 投融资

**人才服务**

招聘 · 培训 · 媒体 · 社交

摩尔精英, 让中国没有难做的芯片

上海 | 硅谷 | 新竹 | 北京 | 深圳 | 合肥 | 重庆 | 苏州 | 广州 | 成都 | 西安 | 南京 | 厦门



长按关注
摩尔快车