走进实验室——王源老师专访

新一期《走进实验室》栏目又和同学们见面了！

《走进实验室》栏目是由信科学生会学术部推出的一档栏目，定期（不定期）采访信科各个实验室中不同专业方向的老师，让未分流的大一同学们对各个实验室的研究方向与大致内容有所了解，也让考虑本研与进组的同学们了解各个导师的研究领域，同时介绍各个领域的学术大咖，扩展同学们的视野。

这次我们有幸采访到王源老师，他为我们深入浅出地介绍了他的课题组的一些基本信息，我们非常感谢王老师对我们的采访给予的极大支持与帮助！

|  |  |
| --- | --- |
| 王源老师 | **研究领域**：集成电路设计与新计算架构  **电子邮件**：wangyuan@pku.edu.cn  **办公电话：**010-62767910  **个人主页**：  <http://scholar.pku.edu.cn/wangyuan>  **个人简介：**  北京大学集成电路学院党委书记，教授，博士生导师。 |

1. 能具体介绍一下您的实验室目前的研究方向吗？

我们的实验室在大方向上属于集成电路设计，具体而言，我们的实验室着眼于**新型计算范式**的集成电路研究，主要包括**类脑计算、存内计算、随机计算**这三大方向。有别于其它课题组的是，我们实验室主要研究**基于传统硅基集成电路**的新型计算范式，注重成果能够与现有大规模集成电路标准工艺结合；其它课题组可能会更多地关注基于新型器件的新型计算研究。基于以上研究方向，我们团队的研究范围覆盖从底层电路到顶层应用的方方面面，从芯片设计到算法研究，再到供应链，各个方向都有涉及。

1. 可以科普一下这几个方面具体是做什么的吗？在未来现实生活中的前景如何呢？

这三个方向其实都在一定程度上源于**生物脑启发**与集成电路设计的结合。我们知道传统的计算结构基本都是基于冯诺依曼体系，在冯诺依曼体系下存储和运算是严格分离的，在计算时就需要在存储器和运算器之间进行数据搬运。然而对于现在的许多人工智能应用场景，它的数据量都很大，如果采用冯诺依曼体系，运算过程中就需要反复地搬运，这就造成了很大的冗余功耗。对于端侧而言，譬如我们所用的手机，这样的高功耗往往是不能接受的。而我们所研究的新型计算所追求的一大特点就是**低功耗高性能**，希望能够比传统的GPU更好地运用在端侧。

除了降低运算功耗，新型计算的另一大特点就是**生物脑启发**。我们说，人脑是强人工智能的唯一参照物。人脑的体积不大，但其中的连接很复杂，通过10个11次方个神经元，10的15次方个突触，它们之间的相互连接，就可以以很低的功耗，完成多种多样的工作——一天三顿饭，就能做到图像识别、目标跟踪、目标识别，路径规划，你想象的事情，机器能做的人脑都能做，而且速度更快，功耗更低；此外人脑还能实现情感和自主意识。由此可见，尽管迄今为止人类人工智能领域已经取得了诸多成果，但是目前所做出来的我们还是把它叫做弱人工智能，它和人脑相比还是相形见绌的。我们通过类脑计算等由生物脑启发得来的计算体系的研究，或许可以探索出一条**走向强人工智能**的道路，同时也可能可以**促进生物脑的研究**。譬如我们知道艾隆·马斯克以及有些国内的研究团队现在在做脑机接口，做生物脑智能体的研究，我们做计算机或者说芯片智能体的研究，这两边有相互融合、相互促进的可能。

1. 目前的研究中有哪些瓶颈呢？

对于我们的研究方向，目前的一大瓶颈在于生物学家还做不到一个人的全脑的解析。一方面，前面说到人脑中有10的11次方个神经元，10的15次方个突触，它们相互连接的拓扑就会非常复杂，这种**网络拓扑的复杂**程度可能比我们想象中的计算系统都要复杂地多得多。另一方面，人脑的**生物化学机制非常复杂**，其中包括各种神经递质的作用机理，它们对神经元会产生抑制、反抑制、长时间增强、短时间增强等多种作用。就单个作用机制而言，生物学家目前研究地还是比较清楚的，但是对于全脑而言，这些作用机制之间产生复杂的相互作用，为什么就能形成影像、情感、意识，目前还没有研究清楚，所以理论基础并不是那么完整，但是我们也不可能等到生物学家解决了所有问题再去做。从某种程度上讲，我们也在摸着石头过河，希冀可以和生物脑研究形成双轨互相驱动发展的这种模式。

1. 类脑芯片的优势和特点是什么？

第一点优势是**低功耗**，第二点优势是**存算一体**，人脑是个存算一体的结构，神经元和突触也是存算一体，脉冲神经网络和类脑计算它也是个存算性的结构。第三点，人工智能芯片大部分在云端训练，端侧推理，这种方式就是最高效的，因为在一个大模型训练的时候需要有一个高精度参数去训练，端侧去做的时候人工化低一点去做推断，但有一个问题就是它没办法**在线学习**，大模型训练的时候基于大样本，训练完之后得到结果，如果我见到一个新样本，一个完全没见过的东西的时候，就没法推理出来。现在的神经网络解决方法基于样本迁移学习。但是生物上其实是在线学习的过程，人脑随时都在学习，他有**自主学习**和在线学习的过程，所以它有一些机制去**在线调整**一些参数，类脑计算我们也是在做这方面的尝试，所以它有**在线学习能力**或者**自主学习能力**这一优势，最后一个优势不容易被关注，现在的人工智能网络的**抗攻击能力**是很弱的，比如在原本样本注入一些噪声，基本就无法被辨认，识别精度不够，但是人脑是一个**抗噪能力强**的东西，它虽然可能不太精确，但是类脑计算是基于**脉冲神经网络（SNN)**，基于Spike的，不是基于一个01代码，所以它有很好的抗攻击能力或者在强干扰情况下的抗攻能力。

综上，**低功耗、存算一体、在线学习、抗攻击抗干扰能力**，这是类脑计算芯片特点与优势。

**小萌新友好部分**

1. 那现在您可以说说实验室是如何发挥它作为科研的沃土的作用呢？您和同学们平时在实验室中是怎样交流学习的呢？

我觉得实验室作为一个科研的平台，它能给大家带来的支持有几个方面。首先是“软”的方面，也就是**实验室的氛围**。在我的团队里，不会说以一个偏工程的方式每个人大家都做同一件事情，也不会说大家各自做的差异性特别大，之间相互交流的机会都没有。可以说都是向着一个宏大的目标，但是大家分工各有不同，譬如可能互相之间是上下游关系，或者说是横向的几条不同的技术路线。我觉得这就是从软的角度来讲的话，提供给大家一个**充分自由交流**的机会。

然后是“硬”的条件。对我自己的博士生而言，如果是做硬件的话，那么可能在整个博士阶段能够有**3~4次流片机会**。虽然流片的成本很高，但是我觉得作为一个集成电路学院的学生，未来肯定希望大家在流片能得到充分训练。在我的实验室，只要你有好的想法，同时你有好的方案，那么我是积极鼓励大家、支撑大家去流片的。

至于如何指导学生，我们会有组例会的形式。我们课题组的模式是：每周由一位博士生做时长**1~2小时的主题报告**，在一个学期中每位博士生轮流报告，介绍他的工作。虽然可能这个报告不是和组内所有学生的研究方向都关系十分密切，但是我会要求每位学生必须去认真听，然后**在报告后进行提问**，起到学术交流的效果，拓展每个学生的知识面。在主题报告之后，我们还会有**常规的工作汇报**，由每个同学讲讲他最近的工作，或者给流片定下进度。

除了组会以外，我还定期会找博士生**一对一地去聊一聊他的研究方向**，了解他最近的研究进展。当然，我觉得组会和一对一谈话形成的外在激励其实是次要的，我觉得你要真的要做研究，最重要的还是研究兴趣和自我驱动能力。我们从事的研究肯定都是新知识，对这些新知识而言，学生会深入地更深一点，导师则会更多地从宏观的方向了解。我们只能教给你工具怎么用，教给你实验手段是什么，但是至于这个东西到底是要怎么做，这就是你需要去研究去探索的东西，而不是课程教授的东西。

1. 您对本科生科研有什么个人的见解吗？您的实验室对想进行本研的学生在GPA/综测/科研/实习经历方面的要求标准如何呢？

我其实**不太看重GPA**，这只是体现你们学习能力的一方面。学习知识考试能力和**把书本上的知识转化成你们的个人能力**，其实中间的差别还挺大的，有些人的GPA很高，但不见得说他适合做电路。**创新**这件事情更深奥，怎样从别人研究工作当中来迸发出一个点子，创新也有各种各样的类型，比如说有的创新是一种集成创新，在别人文献看到一个a方案，看到一个b方案，我把a方案和b方案**结合**后，我就成功创新了；有的时候可能看到a方案，突然觉得可以**再往前走一走**，这又是另一种创新。

关于我对学生的要求，首先我建议先来**听听组会**，知道我到底在做什么，我的实验组做的东西是不是跟你自己的**研究兴趣**是相契合的，这是第一点。我的组会在**微纳525每周五的上午9点**，大家可以随时来听，在组会开始会有**主题报告**，这个主题报告会进行比较整体地宣传。第二点就是想进入我的实验组对于**GPA是没有什么要求的**。而且就集成电路学院来讲，我觉得大家可以在所有导师的组会上听一段时间，你想换到别的组，这都是**自由**的。

但是对于我的学生做研究，我会要求几个基本能力，第一个是**文献阅读**的能力，要能看懂基本内容并且有独到见解，第二个我个人会更倾向于**自我驱动能力比较强**的学生，他很有冲劲，在自己的驱动力牵引下，他会从事科研更容易。如果你自己都很迷茫，只是说要找一个实验室，为了有一份研究的经历，我不建议进我实验组，也不建议进入任何研究组，这样没有意义。当你有自我驱动能力，**有motivation**，**追问能力又很强**，才能往前走很多。

1. 对类脑计算这方面感兴趣的同学，您有什么推荐的**课程或书籍**吗？

坦白地说，类脑计算是蛮新的学科，讲脉冲神经网络的书没有太多。有些综述性的文章以及比较经典的工作，我觉得可以看一下，例如**IBM的TrueNorth**（真北芯片），以及**脉冲神经网络**相关的一些书籍。关于推荐的课程，如果你想做类脑芯片的话，还是希望大家对**人工智能包括深度学习**这方面的知识有所涉猎，因为它只是人工智能的一个分支而已，并没有那么神奇，不是说生物启发就是一个完全偏生物的东西，它仍然基于现有的人工智能这个框架往下走。具体来说，它和人工智能的区别是信息处理的方式不一样。人工智能在处理图像识别等任务时主要采用01代码进行处理，无论是深度神经网络（DNN）还是卷积神经网络（CNN），都在图像上进行反复卷积和特征提取。经过特征提取后，常伴随着池化操作和非线性函数的应用，最终进行分类。这一流程一直是人工智能处理图像任务的典型方法。

在类脑计算中，虽然仍然沿用了之前的处理流程，但对中间的处理环节进行了改变。首先，我们的处理单元不再是传统的处理单元，而是变成了**神经元**，这些神经元又具有多种类型。由于类脑计算更加生物化，神经元在经过非线性函数时不再使用激活函数，而是采用**阈值函数**。这是因为它模拟了生物膜电位的过程，外界不断对膜电位进行刺激，这相当于权重不断增强的过程。当达到一个阈值后，神经元就会发出一个脉冲，传递信息。因此，这一过程被称为阈值函数。

所以，大家如果想做类脑计算方面的研究，要对人工智能相关的基本概念要了解，例如人工智能中**卷积、池化、非线性激活函数，全连接**等这些基本定义，我推荐大家多看一些**人工智能、神经网络的书**。

除此之外，我觉得还需要学习**基本的电路设计**，比如要学**集成电路原理与设计**课，至少电路的**EDA工具**需要学会，以及电路的基本设计语言，毕竟我们要做硬件。

采访&&文案: