清華大學 交叉信息研究院

Institute for Interdisciplinary Information Sciences, Tsinghua University

学术科研简报

2021.07-12

人工智能

量子信息

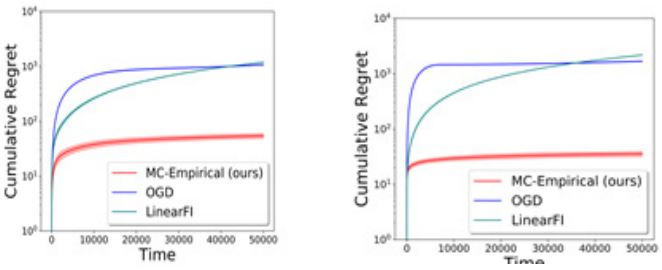
人工智能

一、机器学习

主要完成人：李建研究组、黄隆波研究组、房智轩研究组

连续决策空间的均值 - 协方差多臂老虎机问题

多臂老虎机（Multi-Armed Bandit）问题是一个经典的在线学习模型，它刻画了在线决策过程中探索与利用之间的权衡关系。近年来，现实的在线决策任务对风险控制有着越来越高的要求，如临床试验、金融投资。现有的风险感知的在线学习研究往往关注基于均值－方差的多臂老虎机（Mean-Variance Bandit）问题，这类问题的目标是在最大化奖励均值和最小化奖励方差之间寻求一个最优的平衡点。虽然基于均值－方差的多臂老虎机问题提供了一种有效的风险感知的在线决策模型，但它只适用于离散决策空间，且只考虑到每个选项自身的风险。然而，在许多现实的在线决策任务中，一个决策通常同时涉及多个相关的选项，选项之间的相关性对决策的风险控制具有重要的影响，如金融投资。因此，现有的基于均值－方差的多臂老虎机模型和研究无法直接应用于这些涉及选项相关性的在线决策任务中．为了解决上述风险感知的在线决策模型的局限性，黄隆波研究组首次提出了一种连续决策空间的均值－协方差多臂老虎机（Continuous Mean-Covariance Bandit）问题。这种新模型不但适用于连续的决策空间，而且刻画了选项之间的相关性对决策风险的影响。在这个新模型下，为适应不同的在线学习现实场景的需要，黄隆波研究组针对三种常见的反馈设定分别设计了高效的算法，并从理论分析和实验验证方面证明了算法的（近似）最优性，首次为量化协方差对在线学习过程的影响提供了严格的理论结果。



该成果研究论文：Yihan Du, Siwei Wang, Zhixuan Fang, Longbo Huang, “Continuous Mean-Covariance Bandits,” Proceedings of the Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS), 2021.

组合多臂老虎机的新算法

组合多臂老虎机问题是一类重要的在线学习问题。李建研究组研究了在有攻击者篡改某些臂的奖励时的对抗情景下的组合多臂老虎机问题。这是安全在线学习中的一个基础问题。该研究组设计了一个简单的组合算法，可以实现 O(C+d2K/Δmin) 的后悔度，其中 C 是攻击者修改的奖励的总数，d 是每轮可以玩的集合的最大个数。如果每一轮只选择一个臂，我们设计了一个简单的组合算法，改进了 [Gupta et al., COLT2019] 之前的组合算法，并且几乎匹配 [Zimmert 和 Seldin, AISTATS2019] 得到的最优后悔度。Zimmert 等人的算法需要解决一个复杂的凸优化问题，而李建研究组的算法是组合算法，非常容易实现，需要的假设更少，并且具有非常低运行复杂度。实验也验证了李建研究组的算法更快的运行时间，以及在多个数据集上接近最优算法的后悔度。

CMAB-AC	Time	C=0	C=6000	C=30000
HYBRID	45min	800	10544	44982
CBARBAR	5min	9816	17046	43278

Table 1: Regret and running time comparison between CBARBAR and HYBRID on CMAB-AC with $d = 3, K = 7, T = 10^7, \Delta = 0.1$ and different corruption amount

该成果研究论文：Simple Combinatorial Algorithms for Combinatorial Bandits: Corruptions and Approximations. Haike Xu, Jian Li. Uncertainty in Artificial Intelligence (UAI 2021).

强化学习中基于收益的对比表示算法

在深度强化学习中，如何学习高效的状态表示是解决复杂决策问题和提升样本效率的一个关键问题。最近，很多工作利用基于对比学习的辅助任务在强化学习过程中加强状态表示的学习，取得了很好的实际效果。不过已有基于对比学习的辅助任务并没有充分地考虑到强化学习问题的特性，而且大多是无 / 自监督的。而他们探究了如何利用强化学习中最为重

要的反馈信号——收益分布（Return Distribution）来构建一个新的对比学习式辅助任务。

首先，李建研究组提出了 Zˆπ-irrelevance 聚类函数。直观上来说，Zˆπ-irrelevance 聚类函数会把收益分布类似的状态动作对聚集到一起。相比于之前的抽象函数，该函数能够在不损失过多信息的同时，大幅缩小状态动作空间，从而提高学习效率。接下来，为了从采样数据中学习得到 Zˆπ-irrelevance 聚类函数，提出了基于对比损失函数的 Z 学习算法：

$$\min_{\phi\in\Phi_N, w\in\mathcal{W}_N} \mathcal{L}(\phi, w; \mathcal{D}) \coloneqq \mathbb{E}_{(x_1,x_2,y)\sim\mathcal{D}}\left[\left(w\left(\phi(x_1),\phi(x_2)\right)-y\right)^2\right]$$

尽管 Z 学习算法理论上比较完善，但在实际使用中可能会面临两个问题：1）数据集中正负样本可能不平衡，导致无法有效地训练判别器；2）需要手动地决定收益的离散分块方式。为此，他们提出了轨迹内分段的思路：在轨迹中，如果累计的奖励绝对值变动超过某个阈值，就从这里形成一个新的分段，他们认为相同分段内收益相同。在 Atari 游戏和 DMControl Suite 上，RCRL 算法不仅可以取得比其他算法更好的样本效率，并可以和一些已有的表示学习算法结合，共同提升算法的样本效率。

该该成果研究论文：Return-based Contrastive Representation Learning for Reinforcement Learning. Guoqing Liu, Chuheng Zhang, Li Zhao, Tao Qin, Jinhua Zhu, Jian Li, Nenghai Yu, Tie-Yan Liu. International Conference on Representation Learning（ICLR 2021）

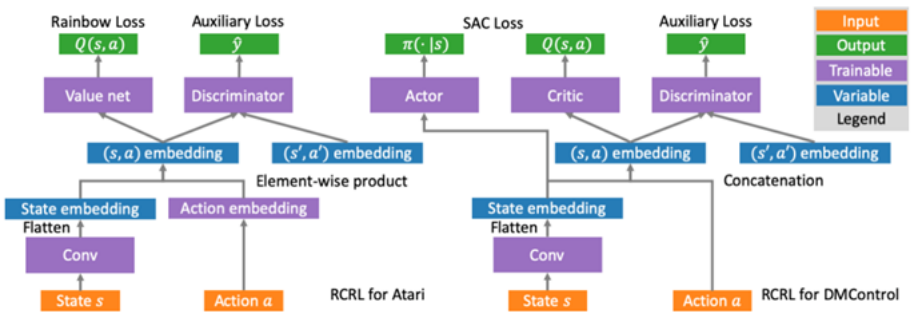


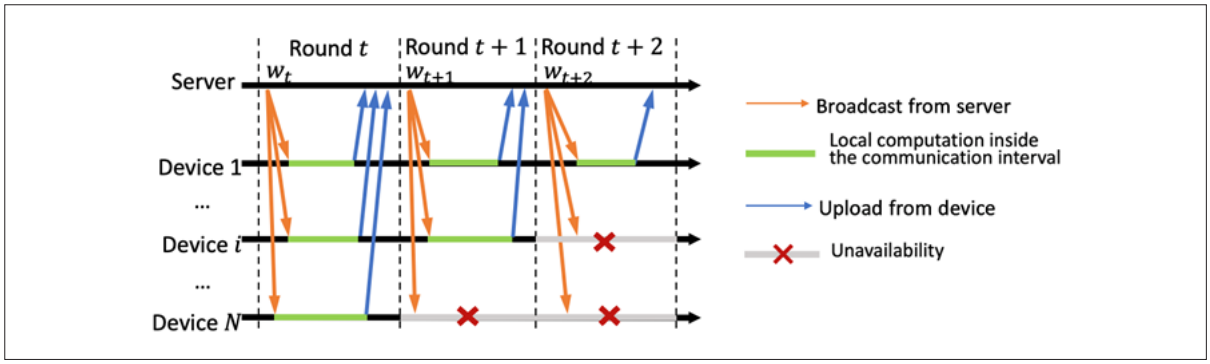
Figure 1: The network structure of RCRL for Atari (left) and DMControl Suite (right).

设备任意参与模式下的快速联邦学习

联邦学习（Federated Learning）是一种机器学习框架，在该框架下，中心服务器通过协调大量的边缘设备（例如手机、物联网设备）来训练一个共享的模型。相比传统的中心化训练，联邦学习充分利用了边缘设备产生的含量数据，并保护了用户隐私。然而，联邦学习也面临着设备参与模式差异带来的挑战。在联邦学习中，设备可能会由于低电量、网络环境差等原因随意地退出训练，因而不同的设备在训练过程中会做出不同次数的响应，可能导致算法不收敛，或大幅地延长训练用时。

针对这个问题，黄隆波研究组从优化角度给出了很好的解决方案。首先，该工作对设备的参与进行了更贴合实际的建模，该模型对设备的参与模式没有任何结构性的假设；研究组提出了 Memory-augmented Impatient Federated Averaging（MIFA）算法，该算法无需关于设备参与模式的信息就可以有效地避免由不活跃设备引起的过长的延迟，并充分利用过去梯度中关于下降方向的信息实现快速收敛；进一步，该工作分别证明了 MIFA 在光滑强凸和光滑非凸问题上的收敛速率，并给出了与收敛上界相匹配的下界，以此说明 MIFA 达到了最优收敛速率。最后，该工作通过案例研究，定量刻画了 MIFA 对 Baseline 算法的提升，并通过在真实数据集上进行数值试验来验证理论结果。该工作被国际人工智能顶级会议神经信息处理系统接收。

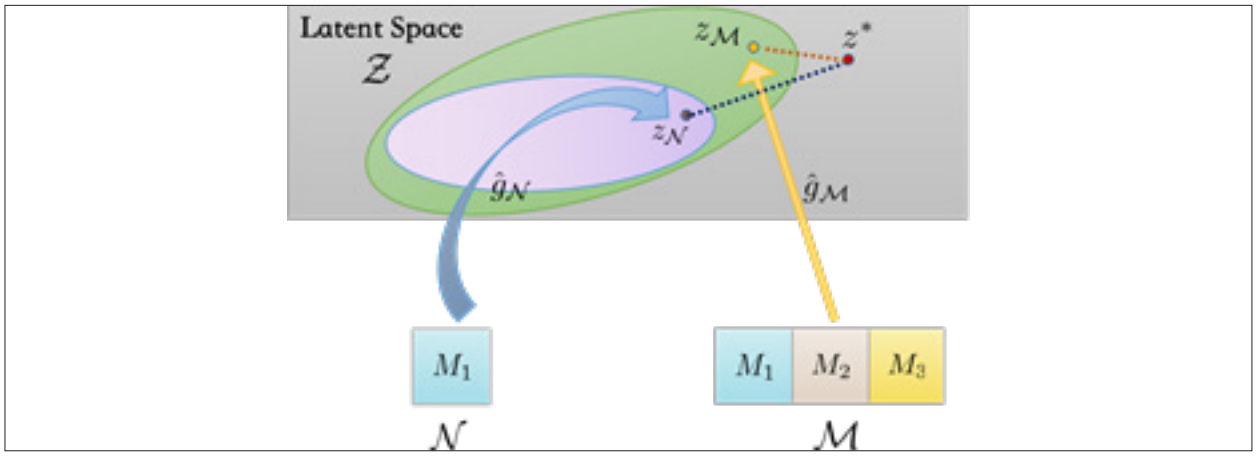
该成果研究论文：Gu, X., Huang, K., Zhang, J., & Huang, L. (2021). Fast Federated Learning in the Presence of Arbitrary Device Unavailability. Advances in Neural Information Processing Systems, 34.



多模态为什么比单模态好（可证明的）？

机器学习在图像（CV），文本 （NLP） 等单一模态领域已取得巨大成功，而融合了多种输入模态的多模态机器学习因其巨大的发展潜力成为当下研究的热点。但该领域的理论研究相当匮乏，基础的问题仍悬而未决：多模态学习能证明比单模态学习效果好么？

黄隆波研究组在理论上从两个角度回答了这一问题：1. 在何种条件下，多模态学习比单模态学习好；2. 是什么造成了其效果的提升。其分析基于一种经典的多模态 fusion 学习框架，包含潜空间学习与任务层学习。理论结果的第一部分表明随着训练数据的增加，使用多种模态训练模型的效果主要取决于它的潜空间表示的质量。进一步的分析则说明，数据量达到一定规模，模态种类越完整，其潜空间表示质量更优，而多模态模型的效果也越好。上述理论结果的正确性均在真实多模态数据集 IEMOCAP 与模拟构造数据集上得到了验证。



该成果研究论文：Yu Huang, Chenzhuang Du, Zihui Xue, Xuanyao Chen, Hang Zhao, Longbo Huang. " What Makes Multi-modal Learning Better than Single (Provably)", NeurIPS 2021.

首次对具有长期时变约束的在线凸优化问题同时实现次线性遗憾值和约束违背

具有长期且时变约束的在线凸优化问题（OCO）由于其对网络路由、在线广告展示、网络资源管理等问题的强大建模能力，成为近年来最受欢迎的在线学习框架之一。然而对在线算法的最优设计与性能保证一直是一个难点。

房智轩研究组在这一个问题上实现了突破性的进展，他们提出了一种新的基于虚拟队列的在线算法 VQB 来解决带有长期且时变约束的 OCO 问题，并针对动态遗憾值（dynamic regret）和约束违背（constraint violations）进行性能分析。算法的核心思想在于设计了一种新的对偶变量更新规则和一种将时变约束函数引入到对偶变量进行惩罚的新方法。研究组提出的算法在许多方面优于目前最先进的结果。例如，该算法不需要 Slater 条件，而已有的文献大多需要 Slater 条件；且该算法是第一个同时实现次线性动态遗憾值和约束违背的非参数算法。同时，对于一类具有很多实际应用的带有长期约束的 OCO 问题，即时变约束随时间变化足够平滑的情况，他们的算法可以实现 O(1) 程度的约束违背。此外他们还利用 doubling trick 将算法和分析扩展到时间范围 T 未知的情况，并保证了与原算法相同的性能。

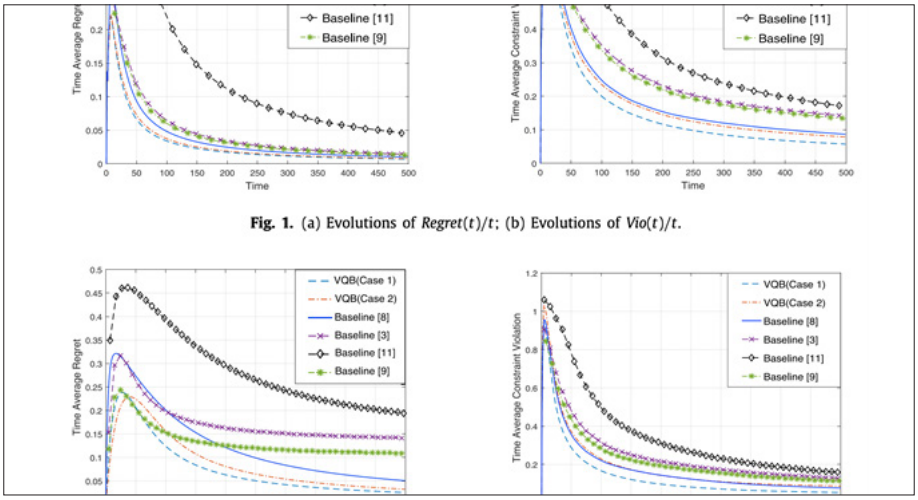


Fig. 1. (a) Evolutions of $Regret(t)/t$; (b) Evolutions of $Vio(t)/t$.

该研究成果论文：Qingsong Liu, Wenfei Wu, Longbo Huang, and Zhixuan Fang, “Simultaneously Achieving Sublinear Regret and Constraint Violations for Online Convex Optimization with Time-varying Constraints.” IFIP Performance 2021.

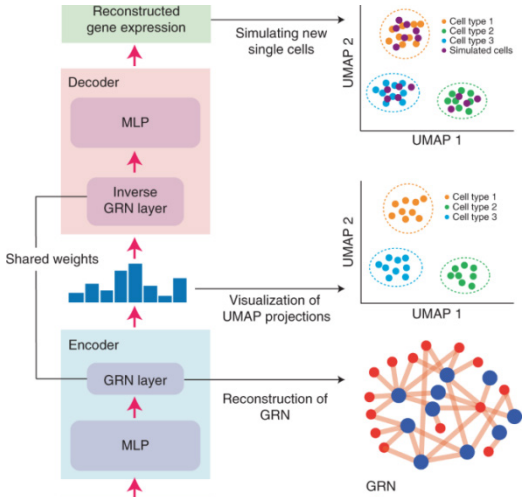
二、计算生物学

主要完成人：曾坚阳研究组、吴翼研究组

深度学习建模基因调控网络新算法

基因调控网络（Gene Regulatory Network，GRN）是研究细胞分化、细胞重编程中的关键问题。近年来，已有工作将统计学方法、自回归机器学习算法应用于这一领域，但这些算法在预测准确性上仍然存在提升空间。随着深度学习在因果推断中的应用，曾坚阳团队同合作者提出了 DeepSEM 模型，通过设计 GRN layer 和 Inverse GRN layer，首次将深度学习应用于基因调控网络的预测，提高了基因调控网络预测的准确度。

基因调控网络作为每一类细胞特有的性质，却被绝大部分 scRNA-seq 低维嵌入模型所忽视。在本项研究中，研究人员利用 GRN layer 和 Inverse GRN layer，将被预测的基因调控网络显式引入 scRNA-seq 低维嵌入中。在聚类和可视化任务中，DeepSEM 在基准数据集上优于现有方法的效果。与此同时，研究人员还进一步将 DeepSEM 模型应用于 scRNA-seq 数据模拟生成，在提出的 GRN 一致性指标上高于现有 scRNA-seq 数据模拟生成方法。该研究巧妙的将基因调控网络预测和 scRNA-seq 数据建模进行结合，实现了一个模型、多个用途，为将来 scRNA-seq 数据分析和计算方法研究提供了新的思路和切入点。

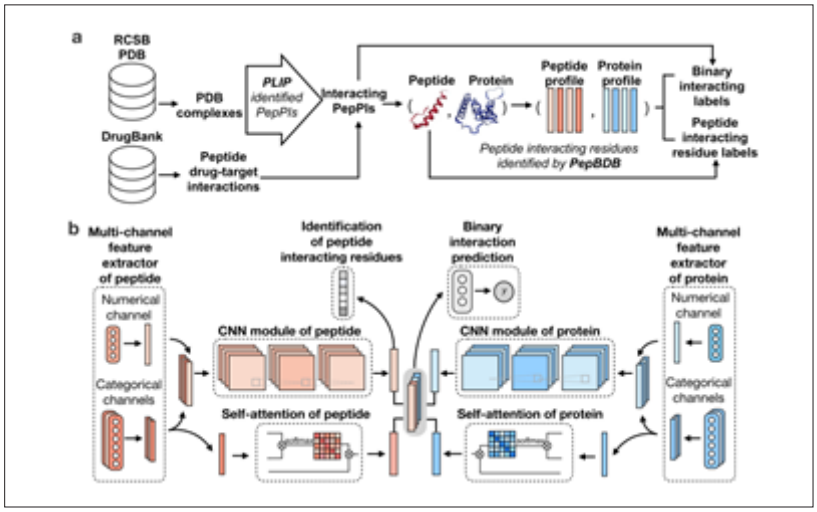


该成果研究论文：Hantao Shu, Jingtian Zhou, Qiuyu Lian, Han Li, Dan Zhao, Zeng Jianyang, Jianzhu Ma. “Modeling gene regulatory networks using neural network architectures”, Nature Computational Science, 2021.

多层次多肽 - 蛋白相互作用预测的深度学习算法

多肽涉及和参与生物体内各种细胞过程，比如信号传导、基因表达调控、细胞增殖和凋亡，在生物体内扮演着不可或缺的关键作用。识别和解析多肽与蛋白质的相互作用及其机制，有助于精准定位多肽药物的靶点，并为多肽药物的化学修饰提供关键信息，从而加速多肽药物的研发进程。目前有主流的计算框架分别基于序列和结构来识别蛋白质和多肽配体的相互作用。然而，这些方法主要集中于识别蛋白质表面与多肽结合的残基，无法直接提取多肽序列中的结合残基。此外，基于结构的方法需要用到三维结构信息，但通过传统的实验方法测定特定的蛋白质－多肽复合物的结构非常昂贵且耗时。

为了解决这个问题，曾坚阳研究组开发了一套深度学习框架，基于多肽和蛋白质序列，多尺度地预测多肽和蛋白质相互作用的深度学习模型。该研究为多肽和蛋白质相互作用的机制提供了一个高效的预测框架，可以在为多肽药物预测结合靶点的同时，识别多肽序列上的结合位点。该模型将蛋白质和多肽的氨基酸序列、二级结构、理化性质、序列灵活性得分和蛋白质的 PSSM 矩阵作为模型输入，利用卷积神经网络模块和自注意力机制来预测给定的肽－蛋白对之间是否存在相互作用，同时识别多肽序列上的结合位点。实验结果表明，模型在基准数据集上的表现均优于现有的方法，且可以准确地预测多肽序列上的结合残基，从而为进一步理解多肽与蛋白质的结合机制提供有效的帮助。曾坚阳研究组还进一步研究了模型在三个相关任务中的应用潜力，即多肽－蛋白结合域相互作用预测、结合亲和力评估和多肽的虚拟筛选。结果表明，模型在这三个相关任务上均取得出色表现。

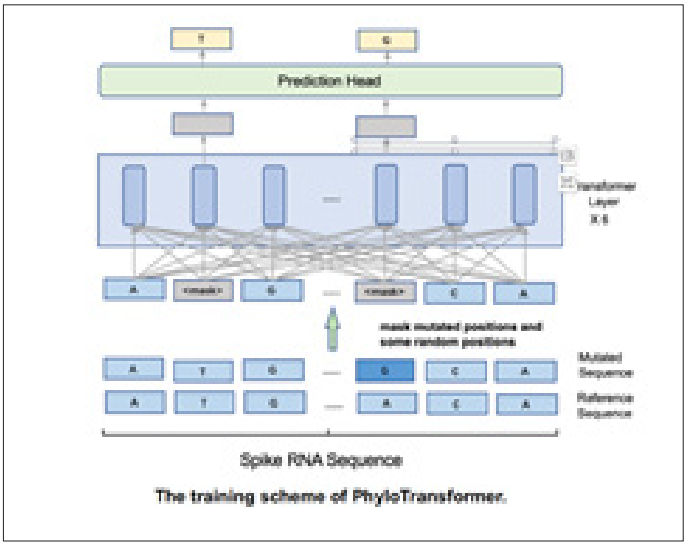


曾坚阳研究组开发了一个多层次的多肽－蛋白相互作用预测的深度学习框架，能够同时预测多肽和蛋白质之间的相互作用和以及识别多肽序列的结合残基，有助于研究者理解多肽与蛋白质结合的潜在机制。

该成果研究论文：Lei, Yipin, Shuya Li, Ziyi Liu, Fangping Wan, Tingzhong Tian, Shao Li, Dan Zhao, and Jianyang Zeng. “A deep-learning framework for multi-level peptide－protein interaction prediction”, Nature Communications, 2021.

利用 Transformer 模型进行 covid 病毒 RNA 突变的预测

自从 2019 年 SARS-CoV-2 病毒引起的疫情爆发以来，直到 2021 年 10 月，全球累计总共感染了 2.19 亿例，其中死亡率高达 3.6%。目前，SARS-CoV-2 已经严重影响了人们的日常生活，并且导致了各种各样的社会问题。在病毒的传播过程中，其 RNA 会进行各种各样的突变，并不断积累优势变异，这给疫情防控工作带来了更多的困难和挑战。注射疫苗是当前疫情防控的一个重要手段，但是病毒变异的不可控性和多样性有可能导致花费了巨大成本研究出的疫苗失效。而如果能够更好地预测未来病毒的变异方向，掌握病毒的变异规律，那么防疫工作和有针对性的疫苗开发工作也更好开展。吴翼研究组将预测病毒变异的问题建模成长序列的处理问题，引入了注意力（attention）机制极大地提升了 RNA 突变预测的准确率，接着研究组预测了未来最有可能发生的几个新的变异。



The training scheme of PhyloTransformer.

病毒的 RNA 就是一系列由 A、C、T、G 碱基组成的序列。在处理 RNA 序列时主要存在两个困难，一个是 RNA 的突变通常是非常稀疏的，从而导致模型在学习过程中非常容易只记住了突变前的 RNA 序列，另一个是 RNA 的序列通常都很长。为了解决这些问题，吴翼研究组采取了 MLM（Masked Language Model）的训练方式，并让模型在训练的时候能够更加关注于发生了突变的位置。同时，该研究组采用注意力机制让模型在预测某个位置的突变时能够关注到整个序列的信息，相比于将序列切短然后挨个预测的做法，这样对长序列进行建模的做法将单个位置突变的预测准确率从 33.1% 提升到了 65.4%。

该成果研究论文：Wu, Yingying*, Shusheng Xu*, Shing-Tung Yau, and Yi Wu. "PhyloTransformer: A Discriminative Model for Mutation Prediction Based on a Multi-head Self-attention Mechanism." arXiv preprint arXiv:2111.01969 (2021).

三、计算视觉

主要完成人：弋力研究组、马恺声研究组、高阳研究组

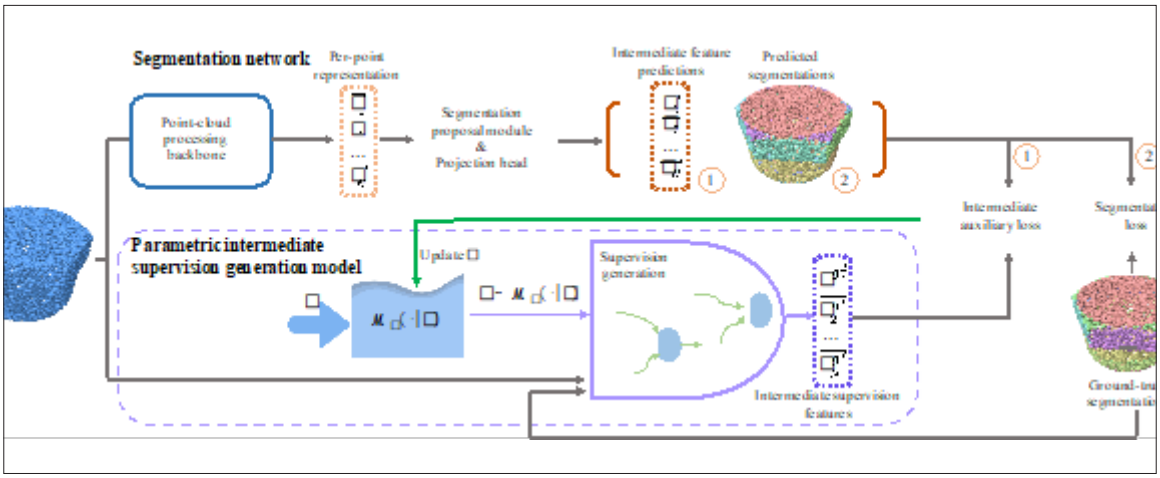
首次提出使用自动搜索中间监督的方法来训练泛化能力更强的 3D 物体部件分割网络

基于学习的部件分割算法因其有效性近年来引起了广泛的关注。但这类算法中较为普遍地存在着难以泛化的问题。为了训练出更容易泛化的 3D 物体分割网络，以往的工作倾向于采用两种思路：或针对所要解决的具体问题基于人对该问题的理解精心设计网络结构，或使用较为通用的为泛化设计的方法来训练的得到更容易泛化的网络。这两种策略虽然有一定的作用，但却仍然存在较为明显的缺点。

人类对问题的理解很可能和机器处理问题的方法并不完全一致，从而基于这种理解设计出来的网络不一定具有较好的性能。此外，针对泛化问题提出的通用的解决方法因为没有很好地考虑问题本身的特点从而很难保证最优性。

为此，弋力研究组着眼于在神经网络中引入中间监督，并设计基于自动搜索的优化算法来使得网络可以自动地找到其最适宜的中间监督以提高网络的泛化能力。在神经网络中引入合适的中间监督信息可以帮助网络更好地学习到容易泛化的信息，并依赖于这些信息完成任务。通过基于对 3D 部件分割任务的先验知识设计的监督信息空间可以使得该空间包含更多的有利于解决这类问题的监督信息。

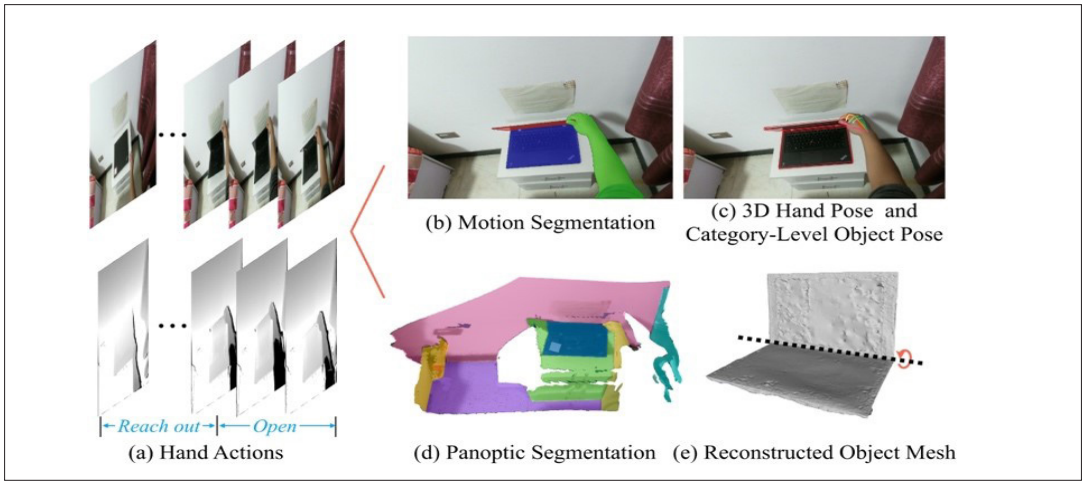
最后，设计合理的自动学习算法让每一个解决分割问题的网络找到其所适合的监督信息可以避免人为设计网络所带来的的人类认知和机器认知的差别从而找到比人类所设计的学习策略更优的结果。在训练过程的每一迭代中，一个监督信息被从所设计的监督空间中采样得到，并应用在分割网络中以获得有关其泛化性的评价指标，之后该评价信息被用来更新所设计监督信息空间（如图所示）。在空间优化结束后，一个贪心风格的方法被进一步使用来从优化后的空间中选择合适的监督信息。AutoGPart 三个 3D 部件分割任务上证明了其相较于以往方法的优越性。它将自动寻找合适的中间监督引入到自动机器学习领域中，无论对泛化网络的设计或是自动机器学习都具有更为深远的启发意义。



该成果研究论文：Xueyi Liu, Xiaomeng Xu, Anyi Rao, Chuang Gan, and Li Yi. “AutoGPart: Intermediate Supervision Search for Generalizable 3D Part Segmentation” CVPR 2021 in submission.

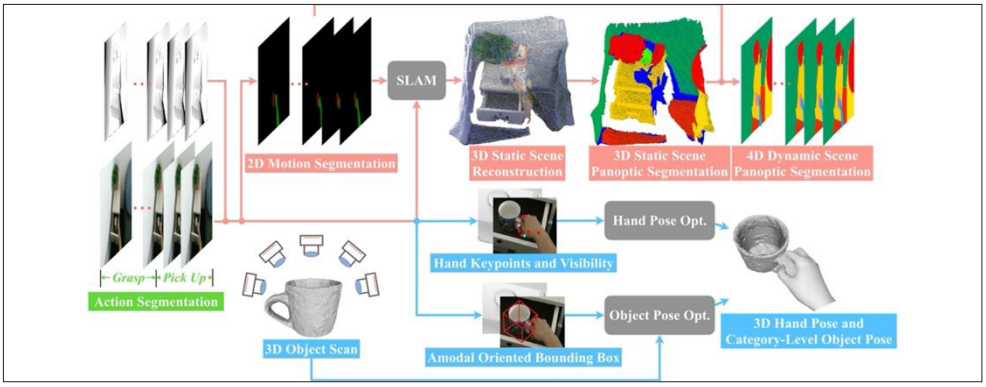
HOI4D：首个针对人和类别级物体交互的第一人称视角四维数据集

近年来的视觉感知方法已在现有数据集的帮助下取得了长足进步，然而大多数工作主要关注静态的被动感知，难以有效地应用于机器人代理和增强现实等重要应用中。近年来，越来越多的工作开始关注交互场景下的感知，已有数据集均关注几个或几十个物体的交互任务，难以支持类别级物体交互的研究。其次，现有数据集大多是第三人称视角的交互，这与人类的真实感知方式间存在显著差异。另外，忽略交互物体的功能性也是目前数据集存在的一大问题。基于此，构建一个新的人类和物体交互的数据集来支持交互感知的研究显得尤为重要。现有的人和物体交互的数据集通常是针对特定类别的特定物体收集交互数据，单一的被交互物体妨碍由此训练的模型对世间物体多样性的认知。



弋力研究组针对现有数据集的局限性，打造出首个大规模的人和类别级物体交互的第一人称视角四维数据集 HOI4D。人类与每个物体类别的拥有不同几何结构的众多物体进行交互，以此加强人工智能对物体多样性和功能性的理解；以第一人称视角观察人和物体的交互过程，这使得人工智能能以最主观的方式感知世界。HOI4D 是一个具有丰富标注的大规模 4D 第一人称的数据集，以支持类别级人和物体交互的研究。 HOI4D 由 3M 张 RGB-D 第一人称的图片组成，总计 5000 个视频序列，由 9 个参与者采集，涵盖 610 个不同的室内场景以及 20 个交互类别的 1000 个不同的物体实例。HOI4D 提供了每一帧的全景分割、运动分割、3D 手部位姿、类别级物体位姿的标注以及重建的物体网格模型和场景点云。

弋力研究组基于 HOI4D 建立了三个基准任务：4D 动态点云序列的语义分割、类别级物体位姿跟踪和具有不同交互目标的第一人称动作分割。实验分析表明，HOI4D 对现有方法提出了巨大挑战，并向人们提供了许多研究契机。除此之外，针对刚性物体的模仿学习实验证明了 HOI4D 数据可以作为学习材料支持仿真环境中灵巧手机器人与物体交互的研究。



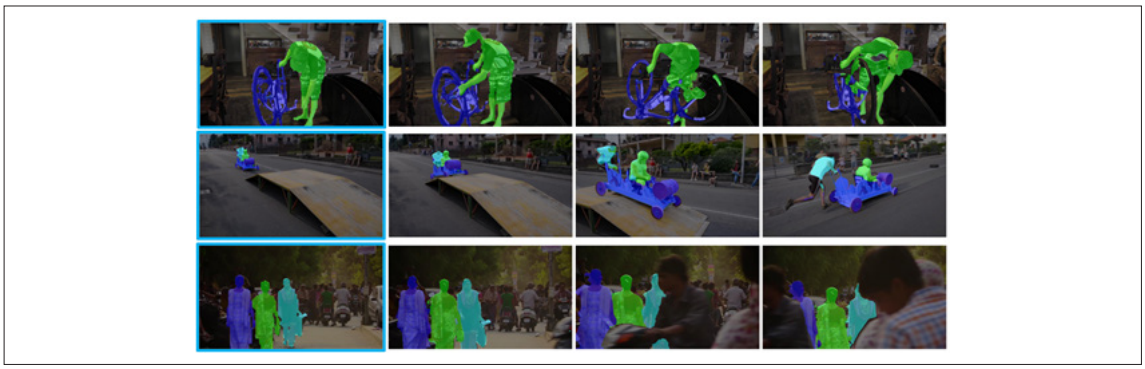
该成果研究论文：Yunze Liu, Yun Liu, Che Jiang, Zhoujie Fu, Kangbo Lyu, Weikang Wan, Hao Shen, Boqiang Liang, He Wang, Li Yi. “HOI4D: A 4D Egocentric Dataset for Category-Level Human-Object Interaction”, CVPR2022 in submission.

利用自监督学习算法学习视频图像中的对应关系

在计算机视觉中，对应关系一直以来都被认为是最基础和重要的研究问题之一。在非常多的下游任务中，比如视频物体分割，物体跟踪和光流估计，都需要能够对图像中的对应关系有很好的表示。之前大部分研究对应关系的工作都依赖在有人工标注的数据集上做监督学习，这样的方法很难泛化到大规模真实数据中。最近越来越多的工作关注到用自监督的方法来学习对应关系，这节省了昂贵的人工标注成本。

高阳研究组发现这些自监督的方法还存在关键的缺陷：没有很好利用到语义信息。这和人类的视觉系统非常不同，人类在完成追踪一个物体这样的任务时，首先是能够在语义上大致将这个物体和场景中其他物体区别开，然后再关注到一些细节的特征，比如纹理，轮廓，以此来做更精准的判断。从人类视觉系统中受到启发，他们提出学习语义可知的细粒度对应关系（Semantic-aware Fine-grained Correspondence），简称 SFC。

SFC 首次提出要同时考虑语义对应关系和细粒度对应关系。并且仅仅使用图像数据，没有任何的人工标注，SFC 能过让神经网络学到对应关系表征，这样的表征在下游任务中展示出非常好的表现，比如仅仅给出视频第一帧中物体的分割图，SFC 利用学到的图像物体间的对应关系，能自动将后面帧的物体也分割出来。SFC 为众多需要对应关系的计算机视觉任务提供了很好的对应关系表征。



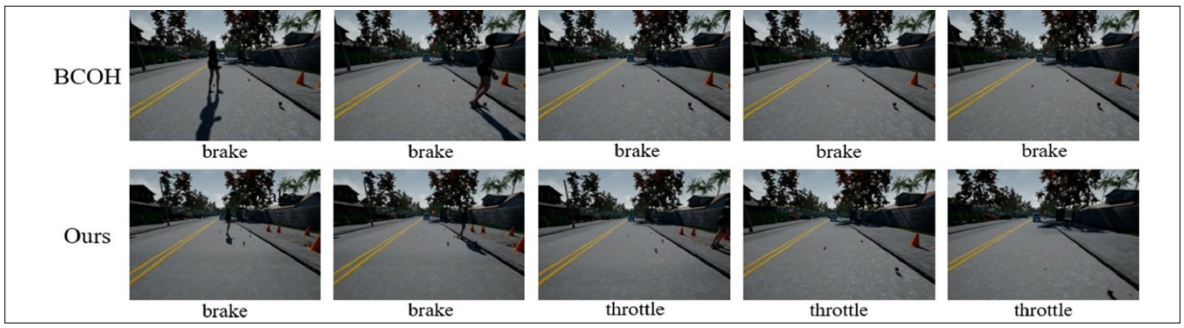
该成果研究论文: Yingdong Hu, Renhao Wang, Kaifeng Zhang, Yang Gao. “Semantic-Aware Fine-Grained Correspondence” submitted to CVPR 2022.

解决模仿学习中存在的抄袭问题

模仿学习（Imitation Learning）是训练策略模型的主流模式之一，其精神是藉由专家行为让模型学到所需知识，行为克隆（Behavioral Cloning）是一个基于监督式学习的实现方法并且在诸多任务中取得优秀成效。在实际视觉基础的任务中，历史信息常被用来增强模型对于环境的认知能力，然而当使用历史信息来训练时时常会有抄袭问题（Copycat Problem）的发生，这是由于模型从历史信息中学到了错误的因果关系导致时常会抄袭先前的动作来当成输出，并且这个现象存在于许多的任务中导致效果不佳。

高阳研究组提出了全新的训练架构分为记忆模块与策略模块，记忆模块能消除历史信息中对于当前决策中先前动作的信息并且提取特征，然后策略模块结合该特征与当前信息进行决策。新的训练架构在 CARLA 自动驾驶任务中取得了相当优秀的成绩且优于先前的方法。更进一步，他们分析记忆模块得到了其能够有效的消除由于错误因果关系带来的影响，并且

并且能够让整体模型能从专家行为中学习更正确的知识，所以此方法能有效地解决抄袭问题。



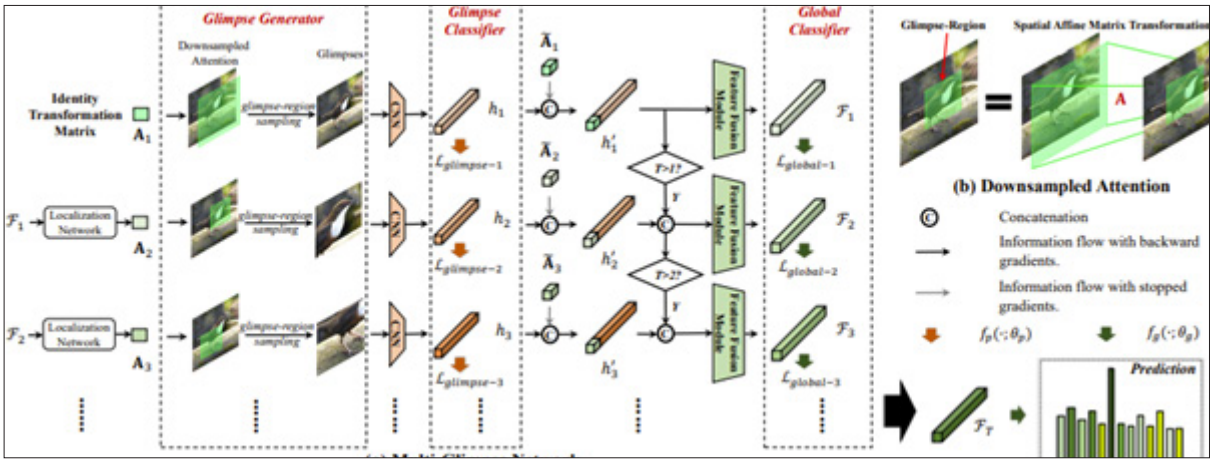
该成果研究论文: Chia-Chi Chuang, Donglin Yang, Chuan Wen, Yang Gao “Resolving Copycat Problems in Visual Imitation Learning via Residual Action Prediction” CVPR 2022, submitted.

深度学习图像分类模型结合人类视觉机制获益

深度学习在计算机视觉领域取得巨大的成功，但依然面临不少挑战，比如模型鲁棒性低以及需要大量算力。然而人类的视觉系统可以在大脑低功耗的运作下得到相对鲁棒的结果。这让马恺声研究组思考深度学习模型是否有机会借鉴人类视觉系统来得到进一步的提升。

这两者最大差别之一在于，人类视觉系统进行图像识别是一个与图像交互的过程，人们只会把大部分精力花费在图片中少数关键的区域，但大部分深度学习模型花费在每一个像素点上的算力是大约相同的，无论这个像素是在关键的主体或是无关紧要的背景上。马恺声研究组提出新的模型框架来模拟视觉系统中眼动和注意力的交互。该框架能扩展现有的深度学习模型，并借助函数可导的特性实现端到端的模型训练。他们验证了在多种常见的深度学习模型主干中，结合人类的视觉机制均可有效降低模型计算量，更重要的是，模型的一般鲁棒性以及对抗鲁棒性都得到显著的提升。

该成果研究论文：Sia Huat Tan, Runpei Dong and Kaisheng Ma. “Multi-Glimpse Network: A Robust and Efficient Classification Architecture based on Recurrent Downsampled Attention.” BMVC 2021.



四、自然语言处理

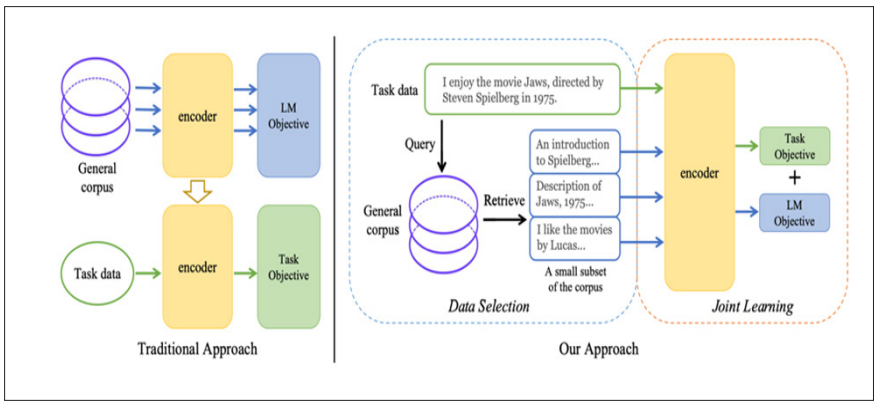
主要完成人：杨植麟研究组、李建研究组、吴翼研究组、

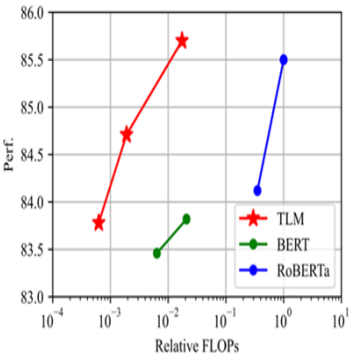
首次提出无需预训练高效 NLP 学习框架：1% 算力 +1% 语料即可比肩预训练语言模型

预训练语言模型因其强大的性能被广泛关注。基于预训练－微调 (Pretraining-Finetuning) 的范式也已经成为许多 NLP 任务的标准方法。然而，当前通用语言模型的预训练成本极其高昂，这使得只有少数资源充足的研究机构或者组织能够对其展开探索。这种“昂贵而集权”的研究模式限制了平民研究者们为 NLP 社区做出贡献的边界，甚至为该领域的长期发展带来了障碍。

杨植麟研究组首次提出一种完全不需要预训练语言模型的高效学习框架：任务驱动的语言建模 (TLM, Task-driven Language Modeling)。这一框架从通用语料中筛选出与下游任务相关的子集，并将语言建模任务与下游任务进行联合训练。相较于传统的预训练模型（例如 RoBERTa），TLM 仅需要约 1% 的训练时间与 1% 的语料，即可在众多 NLP 任务上比肩甚至超出预训练模型的性能。作为这一领域的一项突破性进展，TLM 的提出具有重要意义，它引发更多对现有预训练微调范式的思考，并进一步推动 NLP 民主化的进程。

该成果研究论文：Xingcheng Yao, Yanan Zheng, Xiaocong Yang, Zhilin Yang: From Scratch Without Large-Scale Pretraining: A Simple and Efficient Framework. CoRRabs/2111.04130 (2021).





首次提出小样本自然语言理解评价基准，揭示小样本自然语言理解研究发展现状

以 GPT-3 为代表的预训练语言模型展示出强大的小样本学习能力，仅使用少量标签数据即可快速学习新的自然语言理解（Natural Language Understanding，NLU）任务并取得良好性能。然而该领域的现状是，尚缺乏一个标准的评价准则；且小样本学习高方差不稳定的特点为其客观评价基准的提出进一步带来挑战。已有的各种相关研究工作采取各不相同的评价方式，导致它们或者面临高估真实性能的风险（超参数选择是在既往经验中观测了测试集性能的结果）；或者对于多种因素敏感导致严重性能评价偏差。

针对小样本自然语言理解领域这一严峻挑战，杨植麟研究组首次提出一个新的小样本自然语言理解评价框架 FewNLU，并且从三个关键方面（即测试集小样本学习性能、测试集和验证集相关性、以及稳定性）量化评估该评价准则的优势。基于该基准的重新评价结果表明：现有多数小样本学习方法的真实性能提升有限，小样本自然语言理解领域发展依然面临着严峻的挑战。此外 FewNLU 还揭示：已有工作未准确估计现有小样本学习方法的绝对性能和相对差距；目前尚不存在单一的能够在多数 NLU 任务均取得优势性能的方法；不同方法的增益是优势互补的，最佳组合模型的性能接近于全监督 NLU 系统等关键结论。FewNLU 研究成果的提出是对小样本自然语言理解评价的修正、改进和统一，对于促进该领域未来研究的快速发展具有重要意义。

该成果研究论文：Yanan Zheng, Jing Zhou, Yujie Qian, Ming Ding, Jian Li, Ruslan Salakhutdinov, Jie Tang, Sebastian Ruder, Zhilin Yang: FewNLU: Benchmarking State-of-the-Art Methods for Few-Shot Natural Language Understanding. CoRR abs/2109.12742 (2021).

五、强化学习

主要完成人：张崇洁研究组、陈建宇研究组、高阳研究组

基于逆向模型想象的离线强化学习

强化学习通过大量与环境交互学习，在视频游戏和仿真机器人等虚拟环境取得了巨大的进展。然而在现实应用中（比如医疗保健和自动驾驶等），在线收集大量的交互数据往往是不现实的。离线强化学习提供了一种解决这个挑战的新范式，即只利用离线数据来进行高效的学习，从而通向现实应用的强化学习场景。

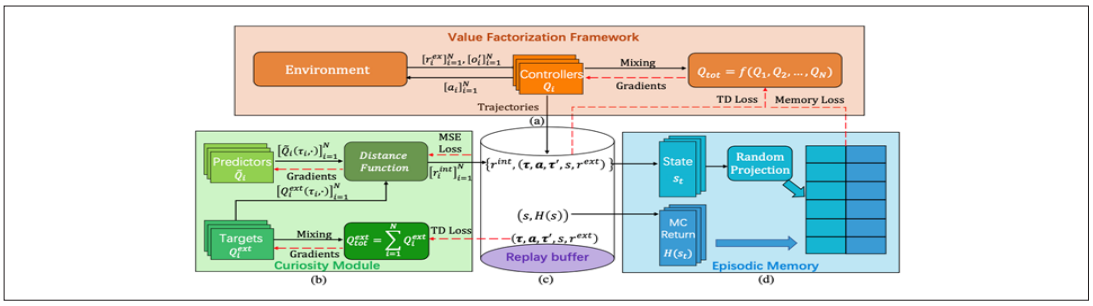
目前的在线强化学习无法直接应用到离线场景中，这是由于在训练和执行时发生了数据分布偏离的现象。张崇洁组提出许多模型无关和模型有关的离线强化学习算法来解决这个问题，但都无法有效地实现在数据集上完成安全泛化性能。张崇洁研究组提出了基于逆向模型想象的离线强化学习算法（ROMI），第一次在离线强化学习中引入逆向动力学模型的概念。该工作利用逆向模型的特性，生成指向数据集的想象轨迹，完成知情性数据增强来实现数据集内的插值泛化。逆向模型利用目标（数据集）指向型想象，从结构上保证了训练安全型，又利用模型本身的泛化性能，实现了结构性安全泛化的能力。本项目提出示例任务并进行理论性分析，并在仿真机器人环境 D4RL 离线强化学习基线的 24 个任务中的 16 个任务取得了最优或并列最优的性能。该工作首次利用逆向模型在离线强化学习中提出了结构性安全泛化的思想，为模型有关离线强化学习方向中提出全新思路和方向。

该成果研究论文：Jianhao Wang, Wenzhe Li, Haozhe Jiang, Guangxiang Zhu, Siyuan Li, Chongjie Zhang. Offline reinforcement learning with reverse model-based imagination. In Advances in Neural Information Processing Systems (NeurIPS), 2021.

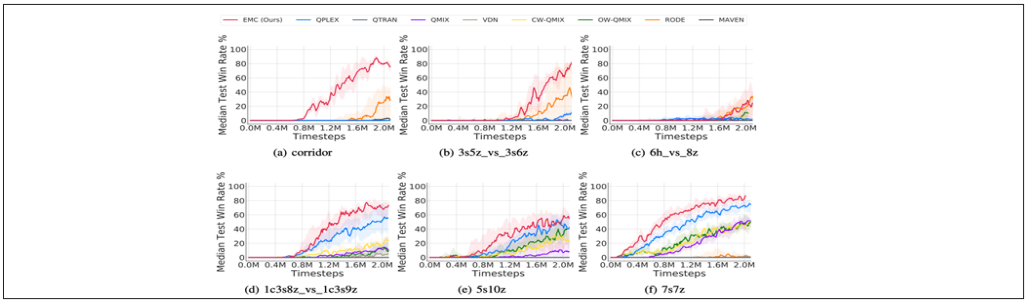
多智能体学习的好奇心探索和情景记忆

多智能体系统存在于无人机控制、游戏博弈、网络优化、工业运输优化等诸多领域，和人类的生活息息相关、密不可分。协作多智能体强化学习研究如何高效学习协作策略，从而构建有效合作的多智能体系统，来解决相应的控制问题。

和单智能体强化学习相比，多智能体强化学习具有状态空间指数级增大、部分可观测性这两大特性。 这使得在复杂任务中，协作多智能体强化学习的高效探索极具挑战性，相关工作寥寥无几。张崇洁研究组提出了一种新颖的基于好奇心驱动的多智能体探索机制（EMC），创造性地把好奇心定义在了个体状态价值函数空间上，然后设计并结合情景记忆学习方法，从而更高效地利用探索到的有价值经验以加快协作策略学习。



这种好奇心探索模式有两大好处：1. 与集中式探索相比，由于个体状态价值函数空间比原始状态空间更加紧凑，这种探索将更高效；2. 与分布式探索相比， 因为个体状态价值函数可以隐式地动态捕捉智能体之间的影响，即使信息部分可观测，也可以诱导智能体对新的或者有潜力的状态的协同探索。实验结果显示提出的方法在星际争霸基准等多种多智能体任务上取得目前最优的性能。

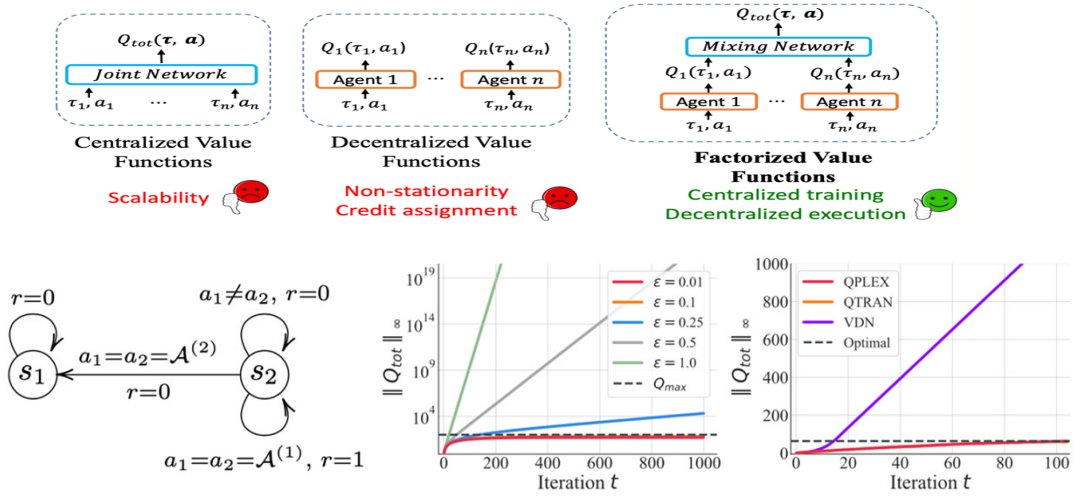


该成果研究论文： Lulu Zheng*, Jiarui Chen*, Jianhao Wang, Jiamin He, Yujing Hu, Yingfeng Chen, Changjie Fan, Yang Gao, Chongjie Zhang, In Advances in Neural Information Processing Systems (NeurIPS), 2021.

协作多智能体学习理论的初步探索

多智能体强化学习是人工智能领域未来发展的下一个重大突破点。在现实世界中，大部分复杂任务都需要多智能体协作完成，如机器人系统、无人车集群和雷达网络等。如何在多智能体环境下进行有效的学习已经成为了当前急需解决的关键问题。特别的是，在多智能体强化学习领域，目前缺乏对当前优化算法的理论研究。

张崇洁研究组着眼于这一挑战，对于协作多智能体强化学习理论展开了初步探索。此项工作提出了多智能体值分解学习方法的理论模型，给出了不同带值分解的优化方法的算法性分析。具体地，针对线性分解算法，通过利用理论模型推导出其内在的奖赏分配机制，证明了在一定条件下的局部收敛性。同时分析发现存在一些任务，线性分解的优化算法会在任意初始化条件下发散。针对这问题，该项研究进一步对具有更强表达能力的 IGM 值分解算法进行分析，从理论上证明其全局收敛性以及最优性保证。最后，通过复杂基准任务在实验上印证了理论结论。该工作在多智能体强化学习中首次利用理论模型分析了多智能体值分解学习算法性质，为整个领域提供了理论基础，也解释了已有算法表现优异的特性。



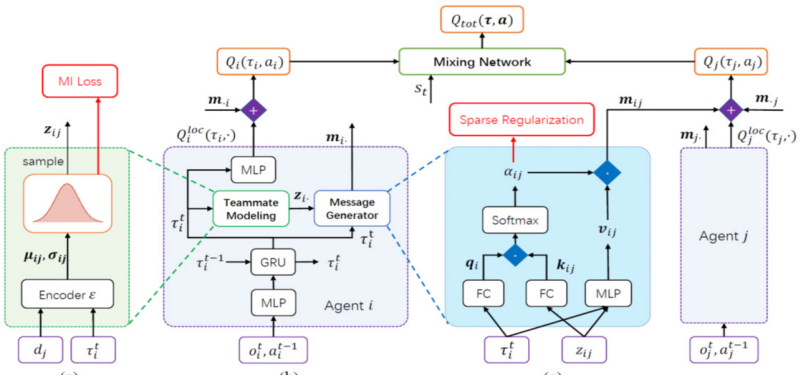
(a) Two-State MMDP **(b)** Different Choices of ϵ **(c)** Deep MARL

该成果研究论文：Jianhao Wang, Zhizhou Ren, Beining Han, Jianing Ye, and Chongjie Zhang. Towards understanding linear value decomposition in cooperative multi-agent q-learning. In Advances in Neural Information Processing Systems (NeurIPS), 2021.

多智能体激励通信的学习范式

多智能体强化学习与人类生活息息相关，许多现实生活中的应用都可以被多智能体强化学习建模，如智能家居系统、无人车合作与雷达网络等。如何在多智能体环境下进行有效的学习已经成为了当前急需解决的问题。更具体地，在多智能体强化学习领域，存在一个长期存在的重要挑战：部分可观测下的通信问题。

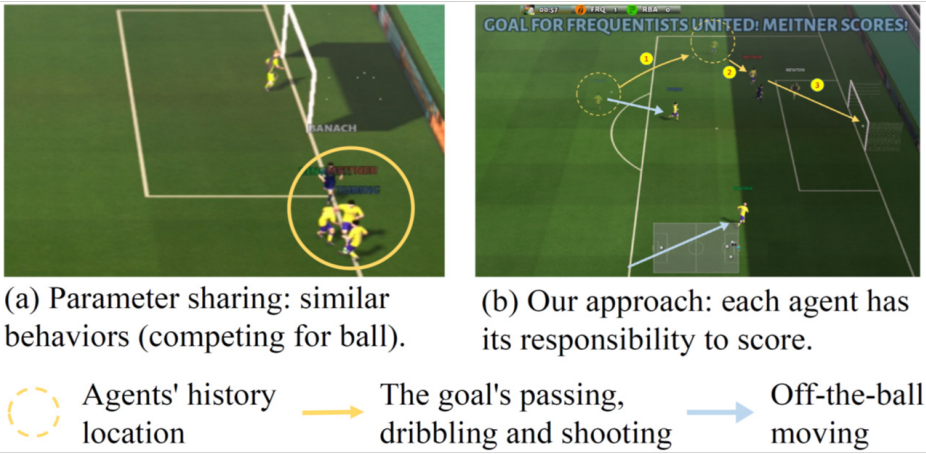
由于多智能体系统的部分可观测问题，智能体之间需要通信来进行合作。在之前的工作中，多智能体群体通信需要发送其观测历史的潜在嵌入，但智能体群体发送的潜在嵌入提供了过于多信息且使得策略空间过于大。为了解决这个问题，张崇洁研究组首次提出了激励通信的学习范式，直接影响其他智能体的价值函数来进行通信。这个方法的策略是首先对每个智能体学习构建对其他智能体的信念，并对其量身定做生成特有的激励通信。由于通讯带宽的限制，该项目还提出一个新正则化项，来获得通信稀疏性并提升通信效率。张崇洁研究组利用此基于激励通信以全新通信方式来建立多智能体系统高效且具有最小带宽的通讯策略。在实验中，该方案展现出通信稀疏性与准确性，并在复杂多智能体基准通信任务（即《星际争霸 2》）中也取得了目前最优的性能。该项目在多智能体通信领域提出了一种全新的通信范式，即激励通信，展现出了稀疏并高效的通信能力，推动了该领域发展。



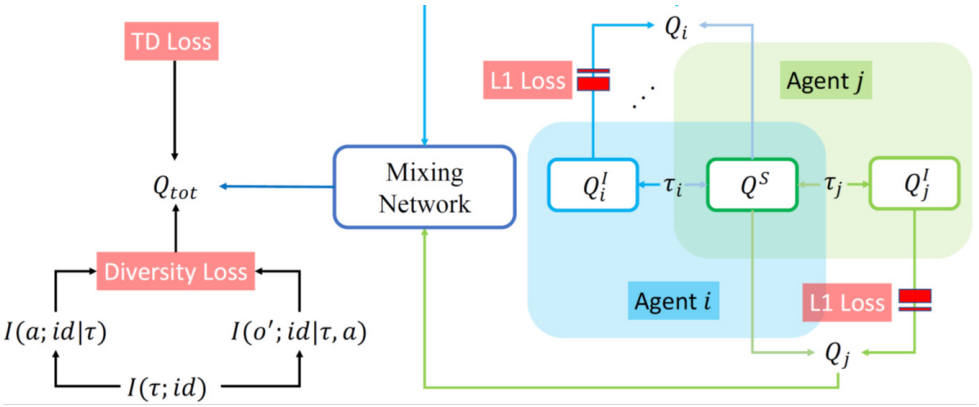
该成果研究论文：Lei Yuan, Jianhao Wang, Fuxiang Zhang, Chenghe Wang, Zongzhang Zhang, Yang Yu, Chongjie Zhang. Multi-Agent Incentive Communication via Decentralized Teammate Modeling. In Thirty-Sixth AAAI Conference on Artificial Intelligence (AAAI), 2022.

基于平衡共享经验和多样性的多智能体强化学习

共享学习是人类高效学习的重要基础，即通过分享经验提高学习效率。近年来在人工智能领域，多智能体强化学习算法利用共享学习机制，通过在智能体之间共享值网络或者策略网络，极大地增加了训练效率。然而目前研究工作大忽略了参数共享导致的模型表达能力不足的问题，同时诱使所有智能体作出相似的行为。许多现实任务往往需要差异性行为的智能体合作来完成，已有的共享学习方法无法有效解决这些任务。如何在利用共享带来的训练效率优势的同时保有必要的多样性是一个亟待解决的问题。



张崇洁研究组对于这一领域进行了开创性研究，首次建立了多样性以及共享之间的平衡框架，拓宽了多智能体强化学习能够解决问题的范围。在该框架网络架构中，保有一个共享的状态－动作值评估网络来在智能体之间共享经验，同时每一个智能体保有一个独立的状态－动作值评估网络来为可能需要的多样性提供足够的表达能力。在优化算法上，通过极大化智能体身份标示和采样轨迹之间的互信息来同时优化动作和观测的多样性，以促使每一个智能体采集到更有价值的经验来在群体里共享。该框架在每个智能体独立的状态－动作值评估网络上增设的一范数正则项可以引导智能体仅在关键的状态－动作组合上考虑多样性，和环境的外部奖励一起约束多样性使其匹配要解决的任务。在极具挑战性的星际争霸 2 微管理测试集和谷歌足球（Google Research Football）上的实验结果表明，该方法形成的多样性以及共享之间的平衡可以这些极具挑战性的基准任务上取得了最领先的学习效率和结果。



该成果研究论文：Chenghao Li, Tonghan Wang, Chengjie Wu, Qianchuan Zhao, Jun Yang, Chongjie Zhang. “Celebrating Diversity in Shared Multi-Agent Reinforcement Learning”, In Advances in Neural Information Processing Systems (NeurIPS), 2021.

Double Q-Learning 的估计偏差

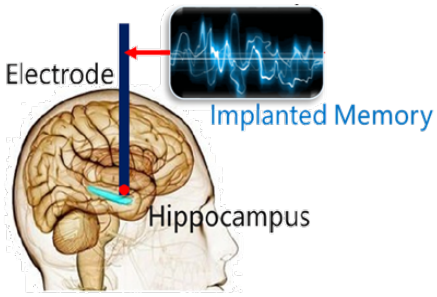
由近似误差累积造成的估计偏差问题是强化学习领域一个悬而未决的问题。在 Q-Learning 算法框架中，正值的误差会在 Bellman 算子的作用下传播和累积，从而产生过估计偏差（Overestimation Bias）。Double Q-Learning 是一个用于缓解过估计现象的经典算法，但它作用下的欠估计偏差（Underestimation Bias）也会对其有效性产生不利影响。

为了对估计偏差问题建立更好的理解，张崇洁研究组在一个已经沿用了近三十年的误差噪声模型下，发现了 Double Q-Learning 算法的一个新性质，即多个非最优稳态解的存在（图 1）。由于这些多余的稳态解在优化空间中的几何特性类似于鞍点，他们提出了一个简洁且有效的启发式解决方案。该算法利用离散化后的 MDP 经验模型建立一个非参数化的额外 Q 值估计，进而为 Q 值网络的输出提供一个近似的下界，从而促使算法跳出非最优稳态解。该成果为 Q-Learning 估计偏差问题提供了新的理论理解和新的工程指导方向，有助于设计更稳定的强化学习算法。

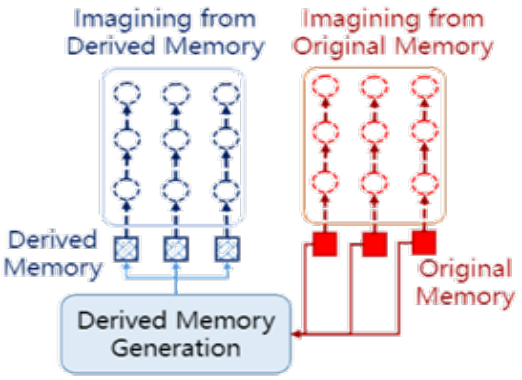
该成果研究论文： Zhizhou Ren, Guangxiang Zhu, Hao Hu, Beining Han, Jianglun Chen, Chongjie Zhang. “On the Estimation Bias in Double Q-Learning.” In Advances in Neural Information Processing Systems (NeurIPS), 2021.

基于派生记忆联想模型的强化学习

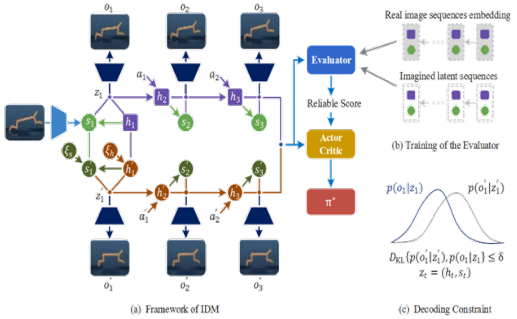
基于模型的强化学习旨在通过对环境建立动力学模型来提高策略学习的样本效率。最近，隐式动力学模型备受关注，它可以在隐空间中进行快速高效的规划，避免了在高维的图像空间进行复杂的计算。它模仿了人类的想象功能，人可以对世界进行抽象和凝练，决策时在脑海中的抽象空间里进行快速预测和规划。然而，在想象的过程中只考虑历史存储的真实经历拉低了想象能力的天花板。受神经科学家在记忆假体方面研究的启发，张崇洁研究组提出了一种新的基于模型的强化学习框架，称为衍生记忆想象模型（Imagining with Derived Memory; IDM）。它可以基于历史存储的真实经历进行联想和扩充，“幻想”出过去没有真实发生过的情况，使智能体能从更丰富多样的想象空间中学习策略。另外，该组评估了“幻想”经历的可靠性，基于可靠性高低设置了使用权重，从而在提高样本效率的同时，还保证了策略鲁棒性，张崇洁研究组在理论上的证明也验证了这一点。在谷歌 Deepmind 的机器人控制任务上，IDM 取得了世界最领先的策略鲁棒性和样本效率。



Memory prosthesis proposed by neuroscientists



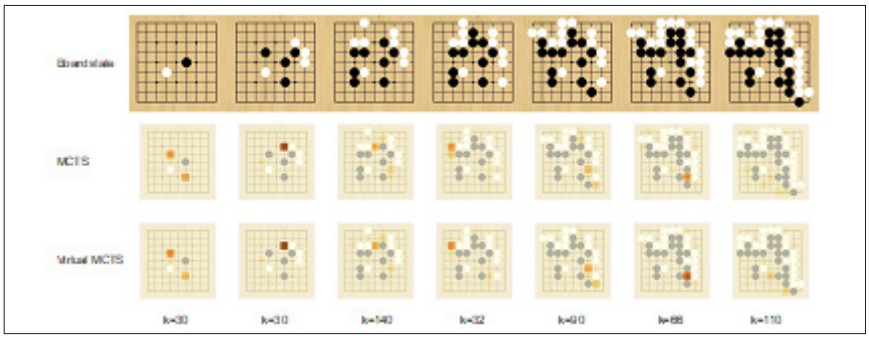
作为这一领域的一项突破性进展，华为诺亚方舟实验室和张崇洁研究组联合首次提出了在隐空间中基于模型扩充训练轨迹的方法，NeurIPS 的程序主席认为该成果会在强化学习领域产生广泛影响。他认为这种扩充轨迹的方法非常具有普适性，可以帮助很多课题组来提升他们现有方法的鲁棒性。



该成果研究论文：Yao Mu, Yuzheng Zhuang, Bin Wang, Guangxiang Zhu, Wulong Liu, Jianyu Chen, Ping Luo, Shengbo Eben Li, Chongjie Zhang, Jianye Hao. “Model-Based Reinforcement Learning via Imagination with Derived Memory.” In Advances in Neural Information Processing Systems (NeurIPS), 2021.

提出 MCTS 算法改进方案，在效果不变情况下效率提升 100%

蒙特卡洛树搜索（MCTS）算法通常用来解决复杂搜索场景问题，但是由于大量的搜索次数会使得算法本身开销极大，而简单减少搜索次数则会导致性能显著降低，受限理性（bounded-rationality）较差，在一些实时场景中难以应用。而围棋（Go）是一种状态数随搜索步长呈指数级增长的游戏，在 RL 算法应用中是一个非常大的挑战，绝大部分 RL 算法无法在围棋上取得较优性能，因此围棋被当作验证强化学习算法在复杂搜索场景下性能的指标。高阳研究组提出一种 MCTS 算法的变种，Virtual MCTS（V-MCTS）：通过一种虚拟展结点（Virtual Expansion）的技巧获取预选策略，并同时针对性地提出一种终止规则，从而使得算法能够根据当前状态的复杂性动态地选择搜索次数，最终提升算法效率。该算法基于此前高阳研究组提出的 EfficientZero 算法，在 9x9 围棋上从零开始训练，通过自我对弈提升算法，并选择与主流围棋引擎 GNU GO（level 10）对弈进行算法性能测试。结果表明，所提出的 V-MCTS 算法能够与原算法保持相当的水平，但是可以减少平均 50% 的搜索次数，最终能够在胜率和速度上均超越该围棋引擎。

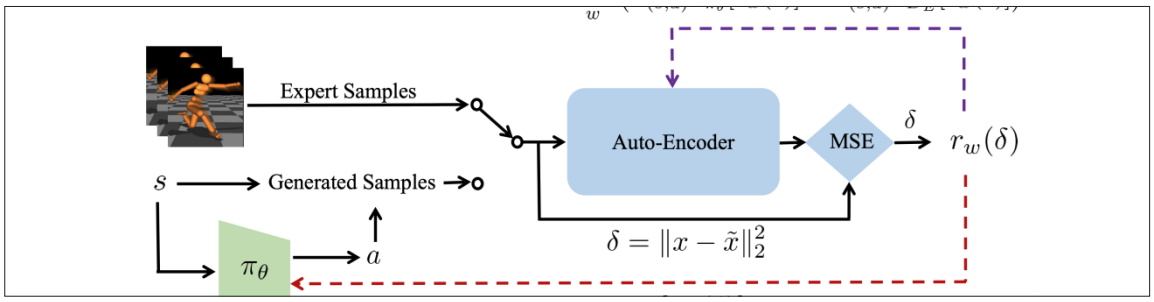


通过算法分析，V-MCTS 的终止过程是针对性的，在简单状态会大幅提前终止搜索，而在复杂状态下会大幅延后终止搜索，从而智能地动态调整搜索开销。因此，该成果为基于 MCTS 的强化学习算法在实时复杂场景下的应用提供了可能。

该成果研究论文：Weirui Ye, Pieter Abbeel, Yang Gao. “Spending Thinking Time Wisely: Accelerating MCTS with Virtual Expansions” ICLR 2022, in submission.

基于 encoding-decoding 过程提升对抗模仿学习模仿能力及鲁棒性

模仿学习是通过大量专家决策轨迹学习专家决策策略的一类方法。对抗模仿学习通过 bi-level optimization 的模型实现两步优化得到专家策略，(1) 通过专家决策轨迹学习环境反馈信号函数；(2) 基于当前所学的环境反馈信号优化策略；进行迭代得到专家策略。目前基于对抗模仿学习的方法主要有两类，一类是基于 Discriminator 的反馈信号函数模型；另一类算法利用离线的手工设计的反馈奖赏函数，通过拉近专家样本和策略采样样本之间的分布获得近似专家的策略。然而基于 Discriminator 的反馈信号函数非常容易过度关注专家样本和当前策略采样样本之间微小的特征上的差异，从而容易过拟合并给予一个较为稀疏的反馈信号函数。此外当专家样本采集过程中存在噪声时，目前的方法将会受到非常严重的影响。



高阳研究组首次提出通过 encoding-decoding 的方式来定义新的反馈信号函数模型(auto-encoder based)。新方法中的反馈奖赏函数能够很好地关注到专家数据和当前策略采样数据之间的全局的差异，给予了一个较为稠密的奖赏用于优化策略。不仅如此，新的方法能够通过 encoding-decoding 的过程对专家数据进行去噪。实验结果表面新的方法在无噪声/含噪声专家数据上都超越了当前最优基准算法 FAILR/PWIL。

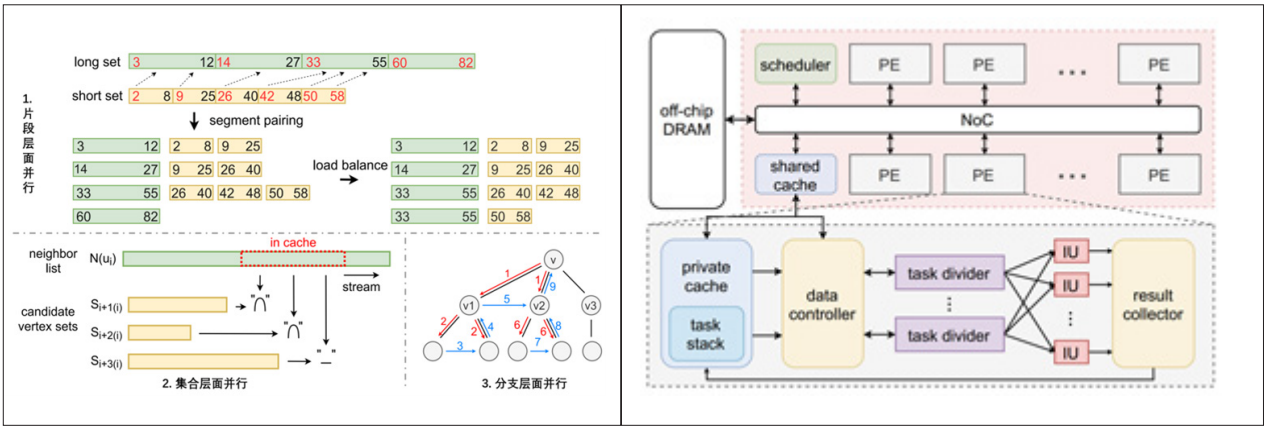
该成果研究论文：Kaifeng Zhang, Rui Zhao, Ziming Zhang and Yang Gao. “Auto-Encoding Inverse Reinforcement Learning.” ICLR 2022, in submission.

六、计算机系统结构

主要完成人：高鸣宇研究组

图挖掘硬件加速芯片架构中的细粒度并行优化

图挖掘（Graph Mining）算法在大规模（数百万节点及数千万边）图中匹配与指定图案同构的子图。其在社会科学、生物信息学等领域有着广泛的应用。目前最先进的普适图挖掘算法将每个图节点的邻接节点表示为集合，将每个起始图节点的匹配任务建模为一个搜索树，在每一级中依次考虑图案中各个节点，通过集合相交和相减进行匹配。由于其巨大的计算复杂度，近年来图挖掘算法开始利用专用软件系统和硬件加速芯片以优化性能和效率。



为了进一步提升图挖掘算法的性能和效率，高鸣宇研究组提出了名为 FINGERS 的硬件加速芯片架构。在利用不同搜索树之间的粗粒度并行执行的同时，该架构提出并实现了三种不同的细粒度并行方式。（1）将两个输入集合分别拆分成片段并配对执行集合相交和相减，利于并行处理和平衡工作量；（2）集合层面并行：在提前计算未来结果的同时复用输入数据；

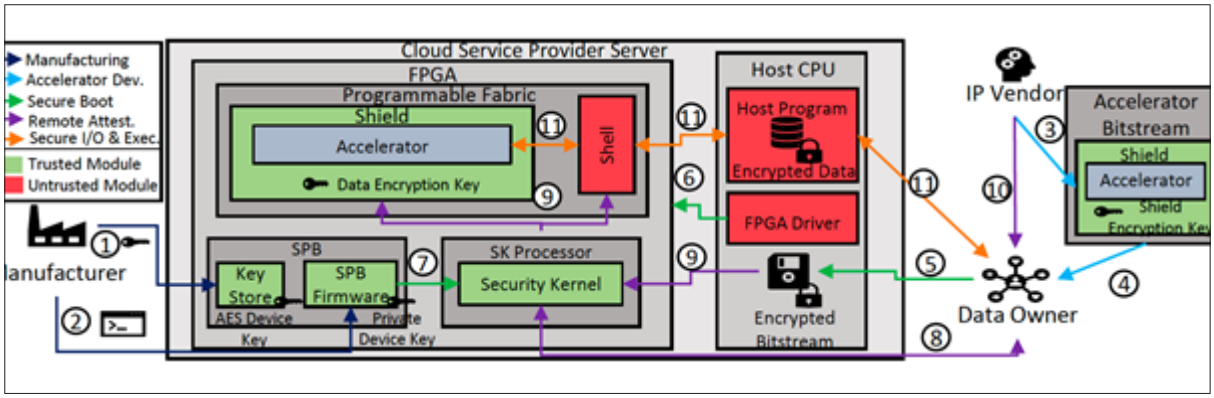
（3）分支层面并行：提供更多可灵活选择的子任务，最大化利用计算资源，并掩盖访存时间。在硬件设计上，该研究组采用资源分配、负载均衡、轻量计算单元间通信等创新设计，高效实现上述并行方式，在同等面积下相比当前最佳架构达到平均 3 倍、最高 8.9 倍的性能提升。

该研究成果论文：Qihang Chen, Boyu Tian, Mingyu Gao, “FINGERS: Exploiting Fine-Grained Parallelism in Graph Mining Accelerators,” ASPLOS 2022.

云计算数据中心 FPGA 硬件上的可信执行环境

现代云计算数据中心中，在传统通用 CPU 之外，FPGA、GPU 等加速硬件正越来越多被应用于高速、高效的数据处理。如微软、亚马逊 AWS、百度等云计算服务商，均开始在其数据中心中广泛部署 FPGA。然而，不同于 CPU 上 Intel SGX、ARM TrustZone 等硬件可信执行环境，当前 FPGA 上尚缺少硬件层面的安全机制保护。这对于金融、医疗等隐私数据在云平台上的处理带来了数据泄漏的风险。

高鸣宇研究组提出了一种针对云计算平台上 FPGA 设备的硬件可信执行环境架构，ShEF。即使在假设攻击者完全控制所有运行软件和接口、并可直接物理接触 FPGA 设备的情况下，该架构仍可提供可靠的安全保护。ShEF 包括一套安全启动和远程认证协议，以及一个名为 Shield 的运行时硬件单元提供数据访问时的加密和验证。其不依赖于主机 CPU 的可信执行环境，不需硬件改动即可在现有 FPGA 上实现，并基于 FPGA 的可重构特性，可高度自由地配置和扩展其组件，允许用户选择最合适的安全机制，以达到保护强度和性能开销间的最佳折中。



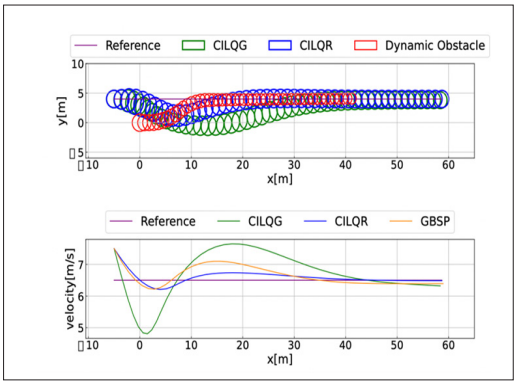
该成果研究论文：Mark Zhao, Mingyu Gao, Christos Kozyrakis, “ShEF: Shielded Enclaves for Cloud FPGAs,” ASPLOS 2022.

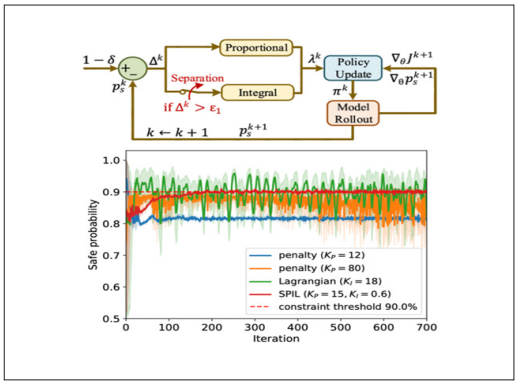
七、机器人

主要完成人：陈建宇研究组、高阳研究组、吴翼研究组

带安全约束的强化学习与控制及其应用

强化学习是实现下一代更智能的机器人系统的关键技术之一。然而目前的强化学习算法主要将最大化奖励作为唯一目标，缺乏安全性的保障。为此，陈建宇研究组需要定义合适的安全约束，并且设计合理的机制来使得强化学习算法能够满足该约束。围绕带安全约束的强化学习与控制，陈建宇研究组从确定和随机系统下的安全状态约束出发，提出了基于优化、能量函数等技术的安全强化学习与控制方法，并更好地利用模型信息，使得算法能够得到更安全、更高效的控制策略。具体研究成果为：（1）提出了一种基于控制阀函数的安全强化学习算法，对于确定系统能够更高效地提高安全性；（2）提出了一种概率约束强化学习方法，对于随机系统能够得到既安全又不保守的控制策略；（3）提出了概率约束下的迭代线性二次型高斯控制器，能高效规划出满足安全约束的机器人运动轨迹。上述成果还在无人驾驶等应用场景中进行了验证，为强化学习在真实世界的机器人应用提供了安全性的支撑。





研究成果论文为：J. Chen, Y. Shimizu, L. Sun, M. Tomizuka, W. Zhan, "Constrained Iterative LQG for Real-Time Chance-Constrained Gaussian Belief Space Planning", IROS 2021.

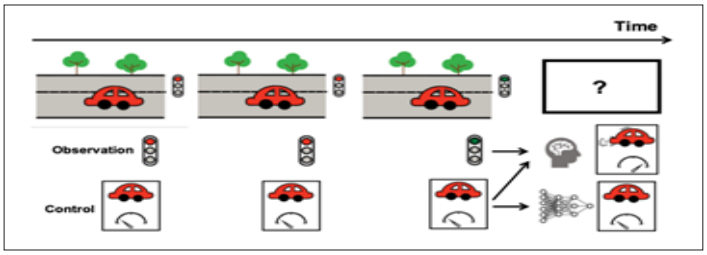
H. Ma, J. Chen, SE. Li, Z. Lin, Y. Guan, Y. Ren, S. Zheng, "Model-based Constrained Reinforcement Learning using Generalized Control Barrier Function", IROS 2021.

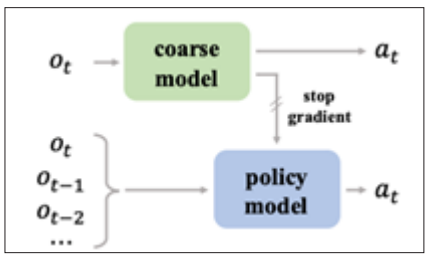
B. Peng, Y. Mu, J. Duan, Y. Guan, SE. Li, J. Chen, "Separated Proportional-Integral Lagrangian for Chance Constrained Reinforcement Learning", IV 2021.

使用正确的捷径来解决模仿学习中因果错误的捷径

模仿学习是策略学习领域中的一项重要的技术，它通过学习从观察值到专家动作的映射来进行决策。在部分可观察马尔科夫决策过程中，模仿学习常常将历史信息作为观察的额外补充来弥补信息的缺失。但是实践表明，将包含历史信息的观察值序列堆叠后作为模仿学习模型的输入往往会导致 copycat 问题，即模型往往会学到从历史信息中恢复并复制前一时刻的动作这一错误的捷径，而不是从观察值中学到决策的真正原因。这一问题在自动驾驶和机器人控制中广泛存在，并且常常会导致严重的错误，有时甚至比单独使用瞬时观察值的效果更差。

针对模仿学习中的 copycat 问题，高阳研究组认为使用瞬时观察和历史观察的模仿学习策略各有其优点和缺点，将这两者模型进行优化组合可以实现两全其美的作用。受此启发，他们提出了一个简单的模型组合方法，其借鉴了人类决策过程：首先根据瞬时观察计算出一个粗略的动作，然后利用历史信息将其细化为一个最终的动作。其中根据瞬时观察进行模仿得到的粗略动作由于没有 copycat 问题因此对于历史信息网络是一个正确但是不够精细的捷径，但因为历史信息网络采用该捷径远远比采用 copycat 捷径更容易，因此避免了 copycat 问题。他们通过在自动驾驶和机器人控制领域的实验中验证该方法在模仿学习任务中成功缓解了 copycat 问题，并显著提升了性能。





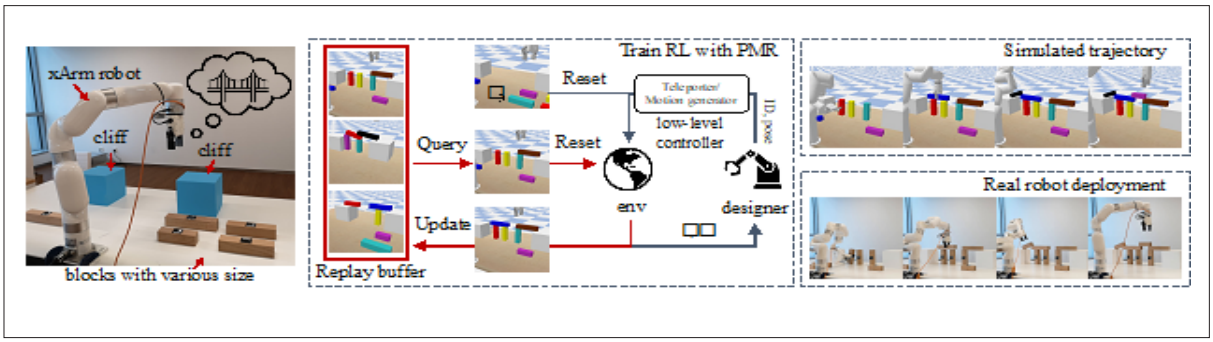
该成果研究论文：Chuan Wen, Jianing Qian, Jierui Lin, Dinesh Jayaraman, Yang Gao. “Fight Fire with Fire: Countering Bad Shortcuts in Imitation Learning with Good Shortcuts.” Submitted to ICLR 2022

使用可变尺寸物料搭桥的智能机械臂

在没有图纸的情况下，机器人可以利用长短不一的物料自己搭建稳定的桥梁吗？吴翼研究组开发了一套基于强化学习的系统解决了这个有趣的问题。在该问题中，物料尺寸不一，所以所有可能的组合方式极多；另一方面，机器人只在一座桥完整地搭出来时才会奖励信号，因此怎样高效地探索到有意义的结构是最大的难点。

吴翼研究组首先把机器人设计建造任务建模成高层和底层两个层次的问题。高层负责设计，由强化学习智能体一次改变一个物体的位置，直到搭出能连通两岸的结构；底层负责执行移动单个物体的子任务，由基于规划的机器人控制算法实现。

为了提高强化学习智能体的探索能力，该研究组设计了一种自动重启机制，让智能体自己挑选曾经访问过的关键状态并从这里继续探索，该机制会比每次从零开始探索更容易获得奖励信号。该研究组还提出了一个自监督的预测任务来促进表征学习，即使在没有外部奖励信号时，智能体也能通过预测逆转移关系获得有效的训练。最后，为了确保智能体的决策是适用于真实机械臂的可行动作，他们结合真机的低层控制策略微调了高层强化学习智能体。



吴翼研究组将系统部署在 xArm 机械臂上，实现了在给定不同组合的物料情况下，自主设计建造不同形态的桥梁，在给定 7 块不等长物料的任务中成功率可达 71.8%。

该成果研究论文： Yunfei Li, Tao Kong, Lei Li and Yi Wu, Learning Design and Construction with Varying-Sized Materials via Prioritized Memory Resets. In submission to ICRA 2022.

八、自动驾驶

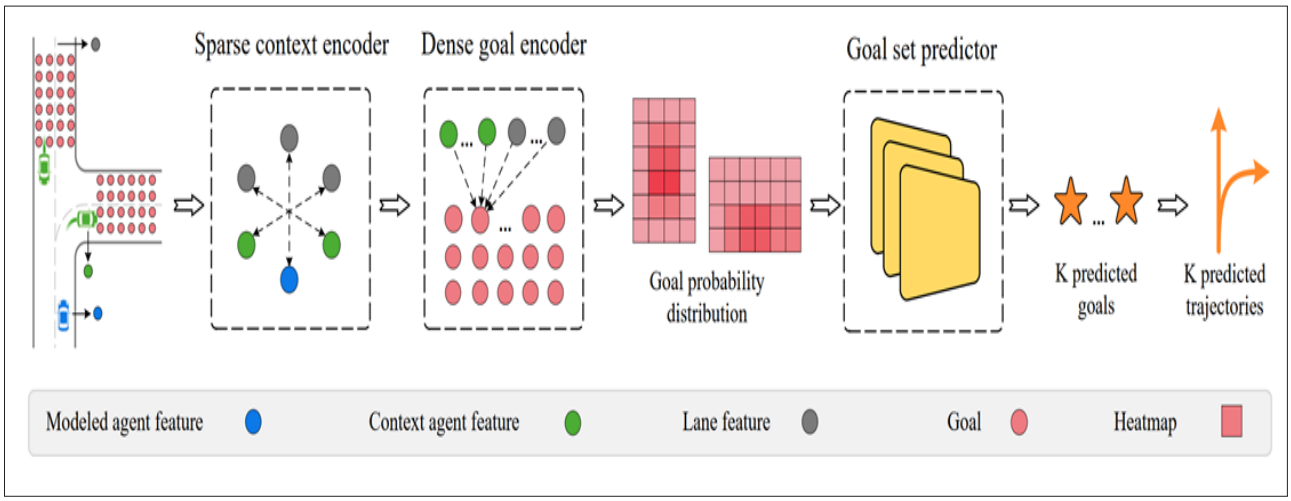
主要完成人：赵行研究组

基于密集终点的端到端轨迹预测

赵行研究组提出了自动驾驶的轨迹预测模型 DenseTNT。在自动驾驶中，为了让自动驾驶车辆更加安全地行驶，避免未来可能的碰撞，需要预测其他车辆的轨迹。因此需要尽可能考虑到车辆不同的情况，即不同的模态，如前行或左转，并预测出对应的概率。

该研究团队提出的 DenseTNT 旨在解决自动驾驶中的轨迹预测的不确定性。作为一个基于终点的轨迹预测模型，即首先预测轨迹的终点，然后基于这个终点补充完整条轨迹，DenseTNT 能够做到 Anchor-free 和 End-to-End。DenseTNT 首先通过直接预测终点的概率分布，从而做到 Anchor-free；再通过设计一个 Set predictor，使得其在推理时能够 End-to-End 输出最优的轨迹方案。

实验表明，DenseTNT 拥有最好的性能表现，在 Argoverse 运动预测基准上排名第一，并在 2021 年 Waymo Open Dataset 运动预测挑战赛中获得第一名。



该成果研究论文：Junru Gu, Chen Sun, Hang Zhao. "End-to-End Trajectory Prediction From Dense Goal Sets", ICCV 2021

九、密码学

主要完成人：陈一镭研究组

Fiat-Shamir 转换中需不需要复杂的哈希函数？

Fiat-Shamir 转换是一种将多轮交互协议转换为无交互协议或数字签名的方法，由 Amos Fiat 和 Adi Shamir 在 1986 提出。目前使用的高效的数字签名，比如 Schnorr 和 Lyubashevsky 的方案，就是基于 Fiat-Shamir 转换来构造的。在过去的三十多年中，人们一直认为 Fiat-Shamir 转换需要用一个很复杂的哈希函数才能安全实现，并且安全性证明都使用的了一种很强的安全模型（称为 random oracle model）。

陈一镭研究组探讨了使用一个复杂的哈希函数是不是必要的。研究组的主要发现是，在诸如 Schnorr 和 Lyubashevsky 等的数字签名方案中，并不需要使用复杂的哈希函数，而只需要很简单的哈希函数，就能在特定的假设下证明安全。目前这两个简单版的数字签名还需要更多的安全分析来判断其实用价值。但这项技术至少有望会启发用简单的哈希函数来实现 Fiat-Shamir 转换，从而得到更高效的数字签名。

该成果研究论文： Yilei Chen, Alex Lombardi, Fermi Ma, Willy Quach. “Does Fiat-Shamir Require a Cryptographic Hash Function?” Advances in Cryptology - CRYPTO 2021

十、理论计算机科学

主要完成人：段然研究组、李建研究组

更快的有限差 (min,+)- 矩阵乘法算法

min,+)- 矩阵乘法是算法领域的经典难题，因其与每两点间最短路问题等价。目前仍然没有真正的快于立方时间 O(n3) 的算法。这里 (min,+)- 矩阵乘法是指结果中 Cij=mink{Aik+Bkj}。在 FOCS 2016 的论文里 [Bringmann et al. 2016]，段然研究组研究了有限差矩阵（即相邻元素的差为常数）的 (min,+)- 矩阵乘法，给出了快于立方时间的算法，其中随机算法的复杂度能达到 O(nˆ{2.824})。

如果矩阵乘法能在 n^{2+o(1)} 的时间内完成，那么复杂度为 O(n^{2.755})。段然研究组还证明了像上下文无关文法的编辑距离、RNA 折叠等一系列问题都能规约到有限差 (min,+)- 矩阵乘法，这就表明了这个问题的重要性。因为之前此类利用快速矩阵乘法的问题很多做到了 O(n^{(ω+3)/2})=O(n^{2.686})，所以对于这个问题应该也有很大改进空间。

段然研究组改进了有限差 (min,+)- 矩阵乘法的算法。与之前的方法类似，他们的方法也将矩阵分成小块先找到结果可能存在的小块。如果这样的小块过多他们利用 [Fredman 1976] 的思路随机选择一些列然后相减，这样就只需比较相反的区间。在利用矩阵乘法时，他们采用随机打乱区间然后多个不同区间重叠计算，最后排出错误块的方法，这样错误块分布较均匀不会过多。新的时间复杂度做到了 O(n^{2.779})，如果矩阵乘法能在 n^{2+o(1)} 的时间内完成，那么复杂度为 n^{8/3+o(1)}。

该成果研究论文：Faster Algorithms for Bounded-Difference Min-Plus Product, Shucheng Chi, Ran Duan, Tianle Xie, to appear in Proceedings of the 33rd ACM-SIAM Symposium on Discrete Algorithms (SODA 2022).

最优的后缀数组构造算法

后缀数组是字符串搜索和数据压缩算法中广泛应用的一个基础数据结构。李建研究组设计了第一个线性时间，常数工作空间（或称为 in-place）的后缀数组构造算法。这些算法在（只读）整数字母表的时间和空间上都是最佳的。具体来讲，李建研究组做出了以下贡献：

1. 对于整数字母表，李建研究组获得了第一个仅使用常数工作空间的线性时间后缀排序算法。输入字符串可以在算法执行期间修改，但应在算法终止时恢复输入。

2. 通过加强第一个算法，李建研究组设计了针对只读整数字母表（即输入字符串不能修改）的第一个线性时间，常数工作空间的算法。该算法解决了 Franceschini 和 Muthukrishnan 在 ICALP 2007 中提出的未解问题。

3. 此外，对于只读的通用字母表（算法只允许比较两个字母），该研究组提出了一种最优时间最优空间的后缀排序算法。

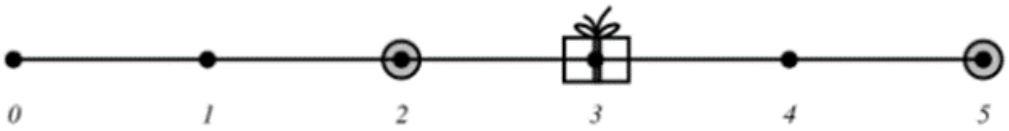
Index	0	1	2	3	4	5	6	7	8	9	10	11	12
Type	L	S	S	L	L	S	S	L	L	S	L	L	S
SA	(12)	(11	1	5	9	2	6)	(E	E)	(E	E	E	E)
SA	(12)	(11	1	5	9	2	6)	(10	E)	(E	E	E	E)
SA	(12)	(11	1	5	9	2	6)	(10	0)	(E	E	E	E)
SA	(12)	(11	1	5	9	2	6)	(10	0)	(4	E	E	E)
SA	(12)	(11	1	5	9	2	6)	(10	0)	(4	8	E	E)
SA	(12)	(11	1	5	9	2	6)	(10	0)	(4	8	3	E)
SA	(12)	(11	1	5	9	2	6)	(10	0)	(4	8	3	7)

该成果论文：Optimal in-place suffix sorting. Zhize Li, Jian Li, and Hongwei Huo. Information and Computation (2021).

有转换成本的马尔可夫博弈问题

李建研究组研究了一个具有转换成本的一般马尔可夫博弈：在每一轮中，玩家自适应地选择几个马尔可夫链中的一个来推进，目标是最小化至少 k 条链达到其目标状态的预期成本。如果玩家决定玩不同的链，则会产生额外的转换成本。 Dumitriu, Tetali, Winkler 利用著名的 Gittins 指数高雅的解决了没有转换成本的马尔科夫博弈问题。然而，对于有转换成本的马尔可夫博弈问题，即使切换成本是一个常数，Banks 和 Sundaram 的 94 年的经典论文表明没有任何类似 Gittins 指数的策略是最优的。

Suppose you are invited to play the following game. Tokens begin on vertices 2 and 5 of a path connecting vertices $0, \ldots, 5$ (see Figure 1). A valuable gift awaits you if either token reaches vertex 3. At any time you may pay \$1 and point to a token; that token will then make a random move (with equal probability to its left or right neighboring vertex if it has two neighbors, otherwise to its only neighbor). Which token should you move first?



李建研究组首次从近似算法的角度研究了该经典问题，并设计了一个简单的索引策略，可以达到常数的近似度。该算法是第一个具有转换成本的一般马尔可夫博弈问题的常数近似度的算法。对于一般度量空间，李建研究组提出了一种更复杂的常数近似度算法。该研究组的技术包括一个对随机 k-TSP 问题的非平凡归约，以及一个用单一随机变量近似一个马尔科夫链的奖励的算法。同时，他们提出了一个新的分析方法，其中广泛使用了 Gittins 指数的各种有趣属性。

该成果论文：Multi-Token Markov Game with Switching Costs. Jian Li, Daogao Liu. ACM-SIAM Symposium on Discrete Algorithms (SODA22).

量子信息

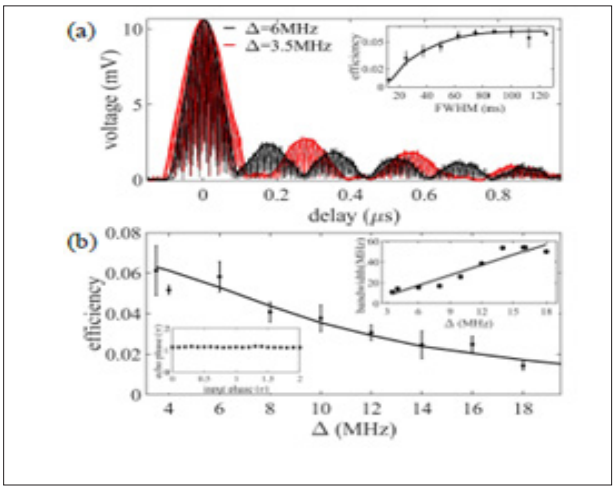
一、量子存储器

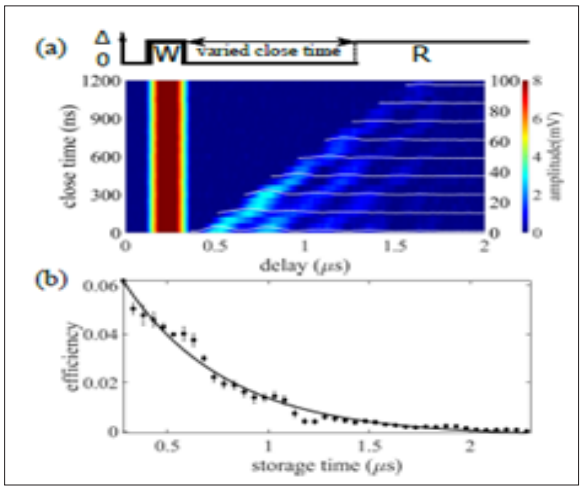
主要完成人：段路明研究组

首次实现微波频段的按需存取量子存储器

量子存储器是构建长程量子通信和大规模量子计算系统的重要组成部分。微波量子存储器主要工作在超导量子比特的特征频段，全功能性的微波存储器是冯诺依曼架构下超导量子计算系统的必备模块，如何借助微波量子存储器实现“飞行量子比特”（flying qubit）与静态量子比特的高效转换是该领域的一个研究热点。

作为这一领域的一项突破性进展，段路明研究组在量子存储领域取得重要进展，首次在实验中借助对多谐振器系统的动态调控实现了对单光子水平微波脉冲的保相存储和读取，并利用此方法展示了对时分编码量子比特（time-bin qubit）的按需存取。此器件结构类似一个原子频率梳，将谐振腔作为人工原子提供了更多的设计和调控自由度，并且能够兼容超导量子计算芯片工艺和结构，易于集成，对于发展含存储模块的超导量子计算系统有重要价值。





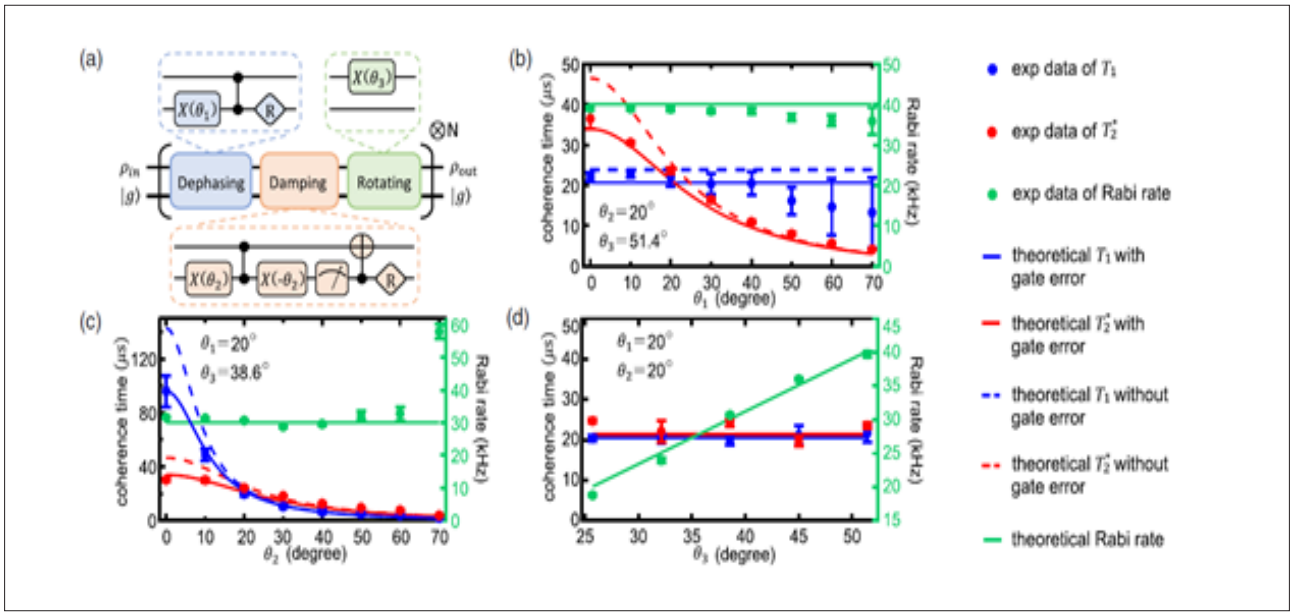
该成果研究论文：Zenghui Bao, Zhiling Wang, Yukai Wu, Yan Li, Cheng Ma, Yipu Song, Hongyi Zhang, and Luming Duan. “On-Demand Storage and Retrieval of Microwave Photons Using a Superconducting Multiresonator Quantum Memory.” Phys. Rev. Lett. 127, 010503, July 2021.

二、超导量子计算

主要完成人：孙麓岩研究组

基于 Trotterization 的开放量子系统动力学实验模拟

数字量子模拟器为解决具有复杂哈密顿量的量子系统的演化提供了多样化的工具，具有广泛的应用潜力。目前封闭量子系统的幺正演化受到了极大的关注，但耗散和噪声对于理解实际量子系统的动力学必不可少。因为真实的量子系统是开放的，其不可避免地会与环境相互作用，从而引入耗散和噪声。孙麓岩研究组演示了在可控马尔可夫环境中进行开放量子系统的数字模拟。



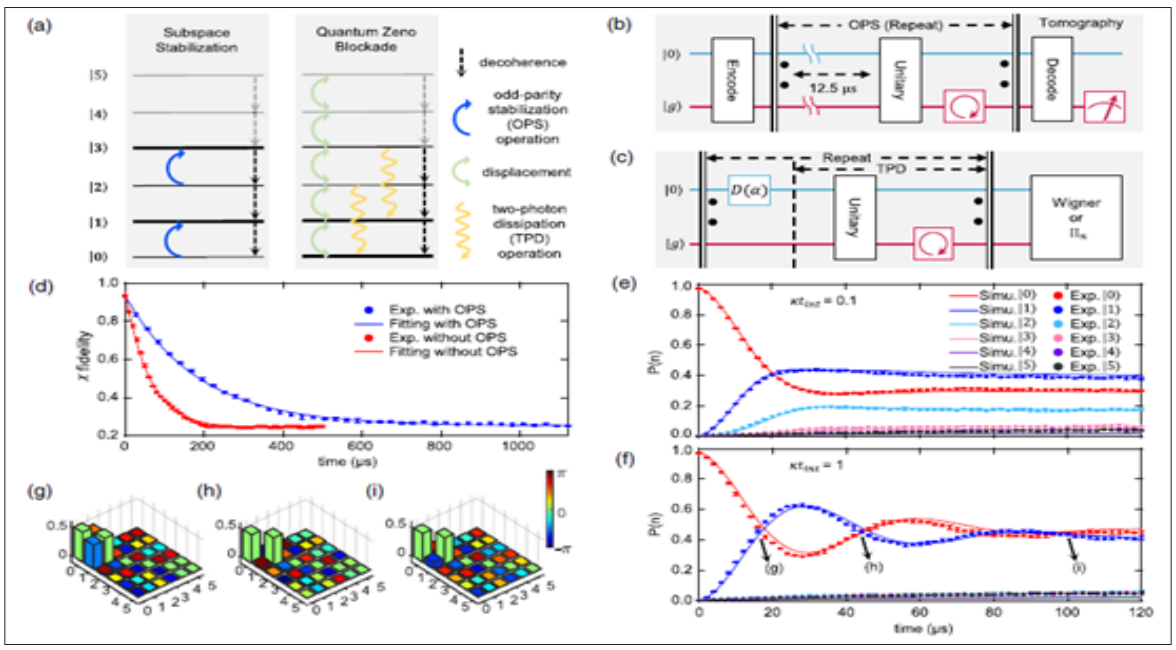
通过将量子刘维尔量进行 Trotterization 形式的分解，有效地实现了开放量子系统的连续演化。并且，实验团队通过调整模拟噪声强度证明了数字模拟在误差缓解中的可行性。该团队更进一步地研究了高阶的 Trotterization，实验结果显示高阶的 Trotterization 用于实现开放量子系统的连续演化具有更高的准确性。该实验工作的结果代表了朝着“硬件高效的开放量子系统模拟”和“嘈杂的中等规模量子算法中的误差缓解”迈出的重要一步。

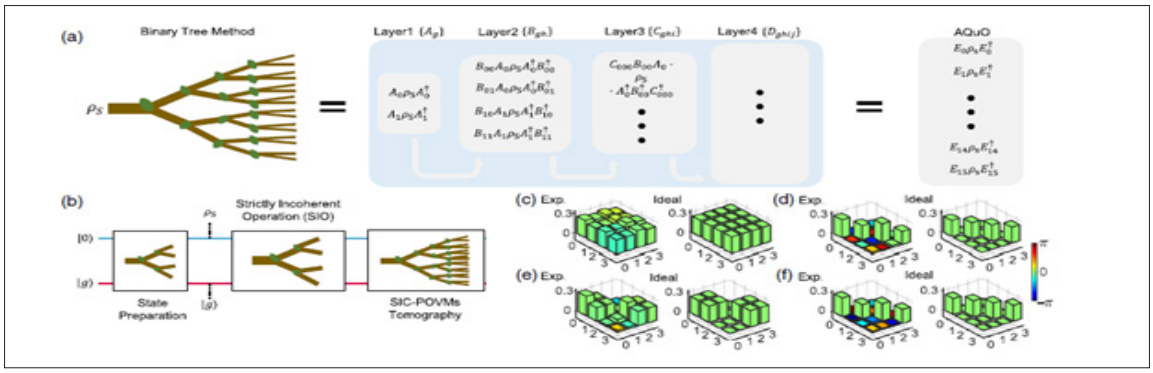
该成果研究论文：J. Han, W. Cai, L. Hu, X. Mu, Y. Ma, Y. Xu, W. Wang, H. Wang, Y.P. Song, C.-L. Zou, and L. Sun, “Experimental Simulation of Open Quantum System Dynamics via Trotterization”,Phys. Rev. Lett. 127, 020504 (2021).

高维量子系统上的高效任意量子操作

量子技术蓬勃发展的背后是成熟的量子控制技术，它允许人们以前所未有的精度操纵物理系统的量子态。量子控制技术的发展从任意量子态制备到任意量子门操作，但都主要关注于封闭的量子系统。然而孤立的量子系统在现实中并不存在。一方面量子系统被环境包围，不可避免地暴露在噪音中。另一方面量子系统的控制和读出是量子任务所必需的，需要与外部系统进行通信。因此，实际的量子系统是开放的。

操纵量子系统的能力是量子技术发展的核心。量子控制的最终目标是为所有可能的开放量子系统动力学实现任意量子操作。然而传统的方法在实现任意量子操作上需要巨大的物理资源，这给相关的实验研究带来了巨大的障碍。孙麓岩研究组演示了在高维量子系统上实现任意量子操作的新通用方法，该方法仅要求一个二能级量子比特的辅助（最小物理资源）和 log_2d 规模的电路深度（d 为高维量子系统的维度）。 运用该方法，该实验团队将任意量子操作应用于量子轨迹模拟，以实现量子子空间稳定和量子芝诺动力学。接着，实验团队进一步演示了，通过任意量子操作实现高维量子系统上的非相干操作和广义测量。该工作所展示的用于完全量子控制的任意量子操作将在量子信息科学中发挥不可或缺的作用。





该研究成果论文：W. Cai, J. Han, L. Hu, Y. Ma, X. Mu, W. Wang, Y. Xu, Z. Hua, H. Wang, Y.P. Song, J.-N. Zhang, C.-L. Zou, and L. Sun, Phys. Rev. Lett. 127, 090504 （2021）。

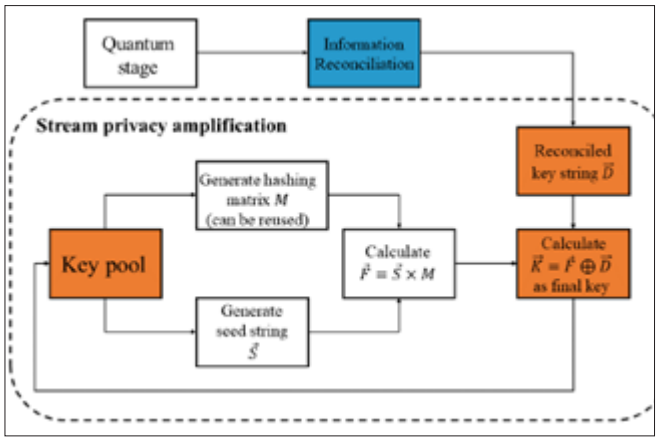
三、量子密码学

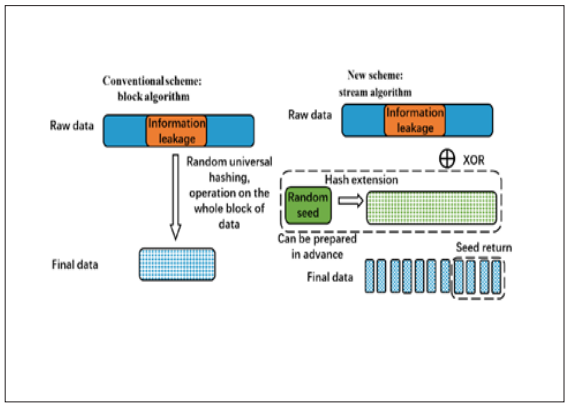
主要完成人：马雄峰研究组

提出量子密码的流式隐私放大方案

隐私放大（privacy amplification）是保证量子通信安全的关键步骤。现有的隐私放大方案基于矩阵乘法，需要积累足够的原始密钥才能进行计算并的到最终密钥。这类似于经典密码学中的分组密码。出于效率的考虑，一次隐私放大所处理的数据块通常非常巨大，导致数据累积的时间会比较长，进而会延迟最终密钥的生成。这个问题在通信链路流量较小或是链路不稳定时，例如卫星－地面链路中，尤为明显。此外，信息协商中遗留的任何错误都会导致整个分组的最终密钥无法使用。

通过重构基于量子纠错的安全性证明过程，马雄峰研究组提出了一种类似于经典流密码的流式隐私放大方案，以解决最终密钥生成延迟和错误扩散的问题。此外，流式隐私放大方案与传统方案不同，可以在信息协商步骤进行，从而有助于提高量子通信后处理的灵活性。应用方面，流式隐私放大方案还可以与量子网络中的延迟隐私放大方案相结合，有助于降低对量子网络中中继节点的安全性要求。相关理论文章见 arXiv:2111.14108（2021）。



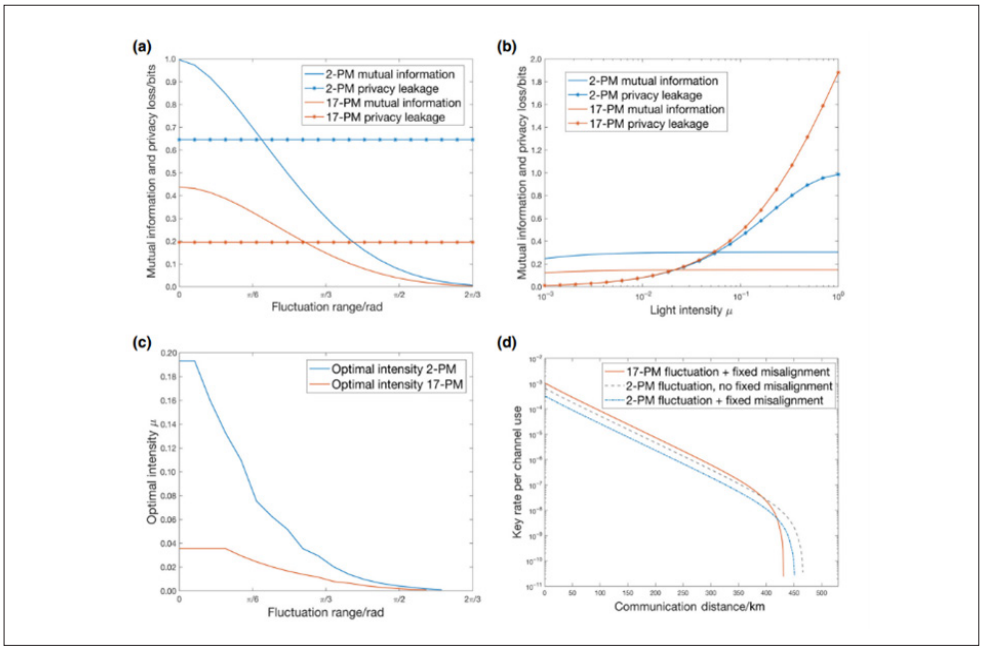


相关成果参加第十七届“挑战杯”竞赛“揭榜挂帅”专项赛“信息安全中的‘矛’‘盾’交锋——抵抗量子计算冲击的量子安全应用”比赛获特等奖。

参考系统无关的相位匹配量子密钥分发设计

最近提出的相位匹配量子密钥分发（PMQKD）协议为突破线性密钥率的限制提供了可行的道路。在该协议中由于密钥信息被编码到相干态的相位上，两个远程参考相位之间的错位会导致错误，大大降低最终密钥生成率。

马雄峰研究组通过引入高维密钥编码空间，提出了一种不依赖于参考的相位匹配量子密钥分发的设计方案。通过令编码相位张成单位圆，可以单独处理并修正任意固定参考相位差的误差，进而计算出偏差角。很自然地，可以通过将二维编码的对称性和互补性拓展到高纬度高维度，该工作还提供了对于高维相位匹配量子密钥分发的安全性证明，并基于该证明演示了在 17 维编码的条件下，相位匹配协议不受任何程度的固定偏差影响，并且对于缓慢的相位波动具有鲁棒性。

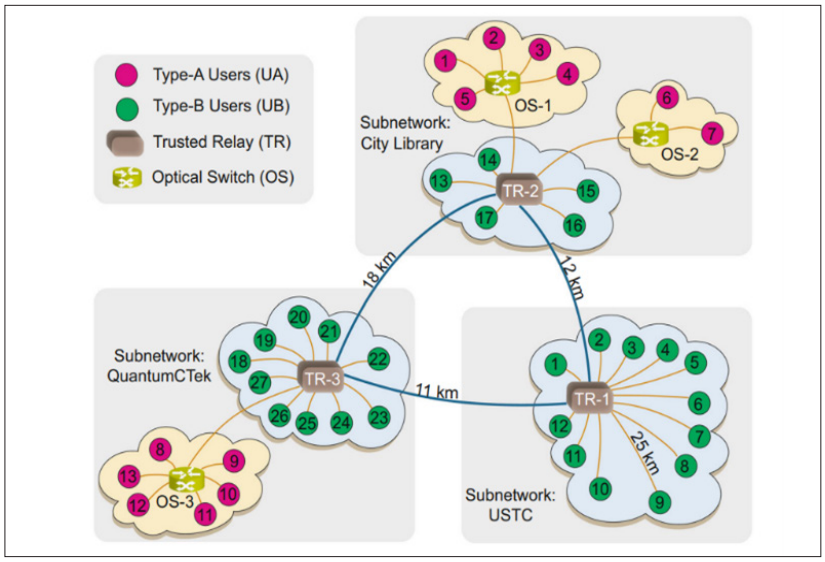


该成果研究论文：Anran Jin, Pei Zeng, Richard V. Penty, and Xiongfeng Ma, Phys. Rev. Applied16,034017(2021).

46 节点量子城域网的实现

量子密钥分发（QKD）实现了两个远程用户之间的安全密钥交换。量子安全通信的最终目标是建立全球量子网络。而现有的实场测试表明，构建量子网络是完全可行的。但是为了构建一个实用的量子网络，还需要克服几个挑战，包括实现大规模的通用拓扑、进行简单的网络维护、实现配置的可扩展性和保证对节点故障的鲁棒性。

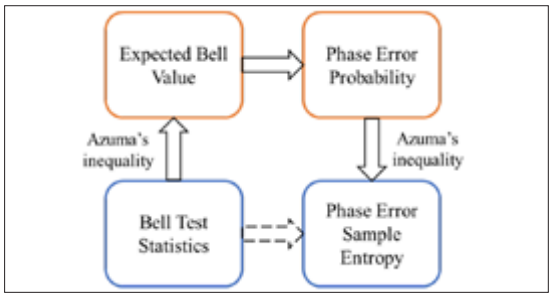
为此，马雄峰研究组与中科大潘建伟团队合作，提出了一个拥有 46 个节点的量子城域网络的操作，并表明所有上述挑战都可以通过现有的尖端量子技术来克服。具体来说，团队实现了不同的拓扑结构，并通过使用具有可扩展配置的标准设备进行网络维护，使网络连续运行 31 个月。团队通过一个先进的密钥控制中心实现了 QKD 配对和密钥管理。在该网络中，最终密钥已被用于实时语音电话、短信和文件传输等安全通信，并采用一次性加密，可支持 11 对用户同时进行音频通话。该工作可以与城际量子骨干网络和地面卫星链路相结合，为全球量子网络铺平了道路。



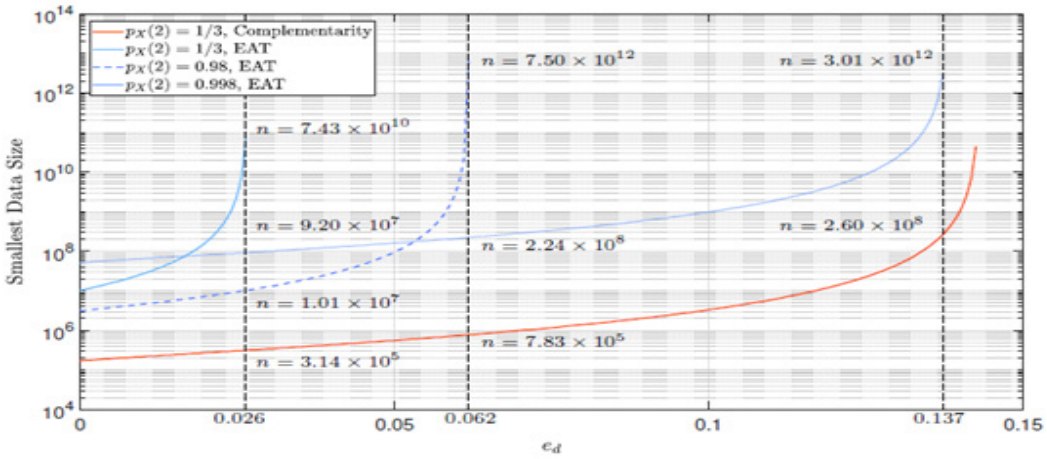
该成果研究论文：Chen, TY., Jiang, X., Tang, SB. et al. npj Quantum Inf 7, 134 (2021)。

基于测量互补性的设备无关安全性分析

测量互补性是量子力学的本质特征之一，是量子不确定关系的来源。对于实验中制备的物理观测量本征态，如果对其进行互补观测量的测量，那么测量结果将是完全随机的。这一原理在量子信息中得到了重要的应用，在量子密码中，基于测量互补性的共轭编码也是量子密钥分发等任务的基本编码方案。另一方面，测量互补性与量子力学的另一核心概念——量子非定域性紧密相关。贝尔测试是验证量子非定域性最为根本而有效的方案。实验者可以在完全不相信实验设备的情况下，单纯通过贝尔不等式违背，以一种设备无关的方式验证纠缠这一非定域关联的存在性。而为了使非定域性显现出来，互补的量子测量必不可少。基于量子非定域性，人们提出了需要最少安全性假设、代表最高安全性的设备无关量子密码协议，例如设备无关量子密钥分发等。然而，在缺少物理设备刻画的情形下，在先前研究中，测量互补性对于量子安全性的具体作用尚未得到充分的认识。



针对这一问题，马雄峰研究组与多伦多大学 / 香港大学 Hoi-Kwong Lo 教授等人合作，揭示了测量互补性在设备无关量子密码任务中的具体角色。通过以基于 Clauser-Horne-Shimony-Holt 不等式的设备无关量子密钥分发协议为例，研究小组定义了设备无关情景下的量子相位错误纠错虚拟协议，定量地将贝尔不等式违背与量子相位错误概率联系起来。此外，通过推广经典信息论中的采样熵参数估计方法，并借助鞅论等随机过程分析手段，建立起了适用于最一般相干攻击的包含有限码分析的完整安全性分析。相比于先前基于量子熵的设备无关安全性分析，新的结果显著降低了实现设备无关量子密钥分发协议的最小实验测量轮数，使基于 NV 色心、冷原子等实验平台的实验实现能够在合理的实验时长完成。此外，这一安全性分析框架可以与 B 步骤、噪声预处理、利用损耗标签等优势密钥提取方案结合，进一步降低实验探测效率、量子态制备保真度等要求。



该成果研究论文：Quantum Complementarity Approach to Device-Independent Security Xingjian Zhang, Pei Zeng, Tian Ye, Hoi-Kwong Lo, Xiongfeng Ma; arXiv:2111.13855.

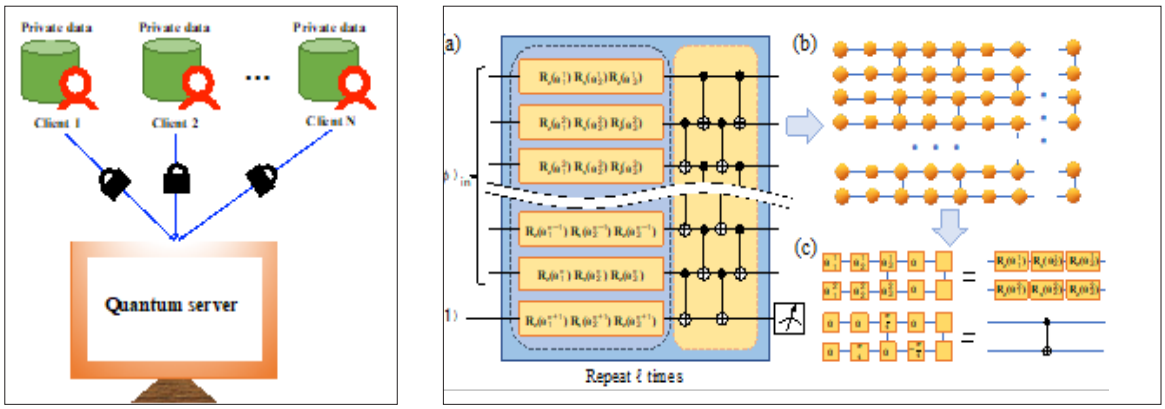
四、量子人工智能

主要完成人：邓东灵研究组

基于盲量子计算的量子联邦学习

量子机器学习是一个新兴的交叉前沿领域，近年来受到了广泛的关注。在未来量子计算机可能以服务器的形式存在于云端供用户使用的情景下，如何在用户远程提交计算任务和数据的过程中保护用户的隐私成为一个重要的问题。这对基于量子计算机的远程代理机器学习、联邦学习任务等具有重要意义。

在此背景下，结合盲量子计算方案与微分隐私，邓东灵研究组首次提出了量子联邦学习。研究组所提供的框架可以支持多个用户合作的联邦学习，并在学习过程中保护每一个用户的隐私数据不被泄露。同时，研究组为量子联邦学习提供了一个完整的方案框架，包括了量子线路构造、模型优化方案、模型对噪声的鲁棒性、微分隐私方案对数据的保护效果、以及不同嵌入数据的方式对训练效果的影响等，将为未来相关领域的工作提供有益的指导。研究组首次提出了一个完整的量子联邦学习框架，并结合了盲量子计算和微分隐私来保证多方学习过程中的用户数据安全。与经典方案相比，研究组提出的方案结合了量子计算能力优势和经典联邦学习的隐私保护优势，在未来量子服务器商业化的阶段，具有潜在应用价值。



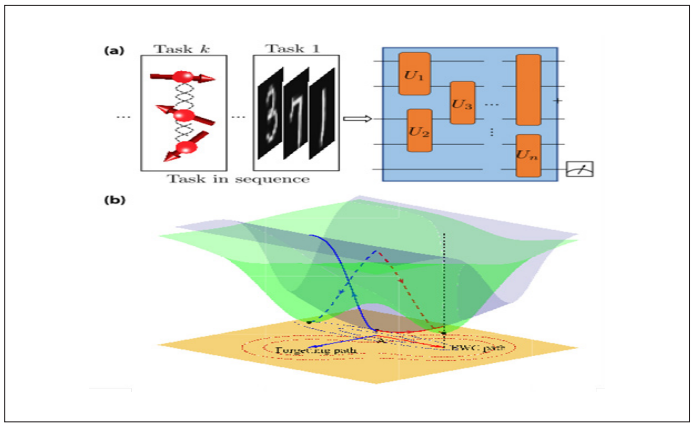
本研究在机器学习与量子计算之间构建了新的联系，将对今后相关方向的理论和实验研究提供一定的指导。

该成果研究论文：Weikang Li, Sirui Lu, Dong-Ling Deng, Quantum federated learning through blind quantum computing, Sci. China-Phys. Mech. Astron. 64, 100312 (2021).

量子持续学习

量子机器学习主要研究量子物理和机器学习的交叉领域，近年来受到广泛的关注。这一交叉领域发展非常迅速，有许多新奇的量子机器学习算法被提出，其中部分算法被认为是相较于经典机器学习算法具有指数级的加速。研究组首次将持续学习的概念引入量子机器学习领域，探究了量子机器学习算法的灾难遗忘现象，并结合经典机器学习中的 elastic weight consolidation 方法，在量子学习模型上实现了持续学习。

本研究主要关注基于变分量子线路的分类器的量子持续学习。研究组首先探究了在持续学习的场景下，量子分类器在学习后续任务的过程中，会遗忘先前任务的信息，即是灾难遗忘现象。本文研究表明，即使是持续学习两个相似的任务，量子分类器也会显著遗忘已经学习到的信息。进一步的，本研究将基于学习模型损失函数局部几何信息的 elastic weight consolidation 方法应用于量子分类器，实现了量子分类器对多个任务的持续学习。 结果表明，在使用 elastic weight consolidation 方法后，量子分类器可以显著保留已经学习到的信息，从而使得在新任务的学习过程完成后，量子分类器依然可以有效区分先前任务的数据。与经典持续学习相比，本研究采用的量子分类器相较于经典分类器具有更强的表示能力，这对实现通用学习具有重要意义。本研究将持续学习引入量子机器学习领域，构建了机器学习与量子计算之间新的联系，将对今后相关方向的理论和实验研究提供指导。



该成果研究论文：Wenjie Jiang, Zhide Lu, Dong-Ling Deng, Quantum continual learning overcoming catastrophic forgetting, arXiv:2108.02786.

Edited by Qin XIE

Reviewed by: Luming Duan, Jian Li, Xiamin Lv