

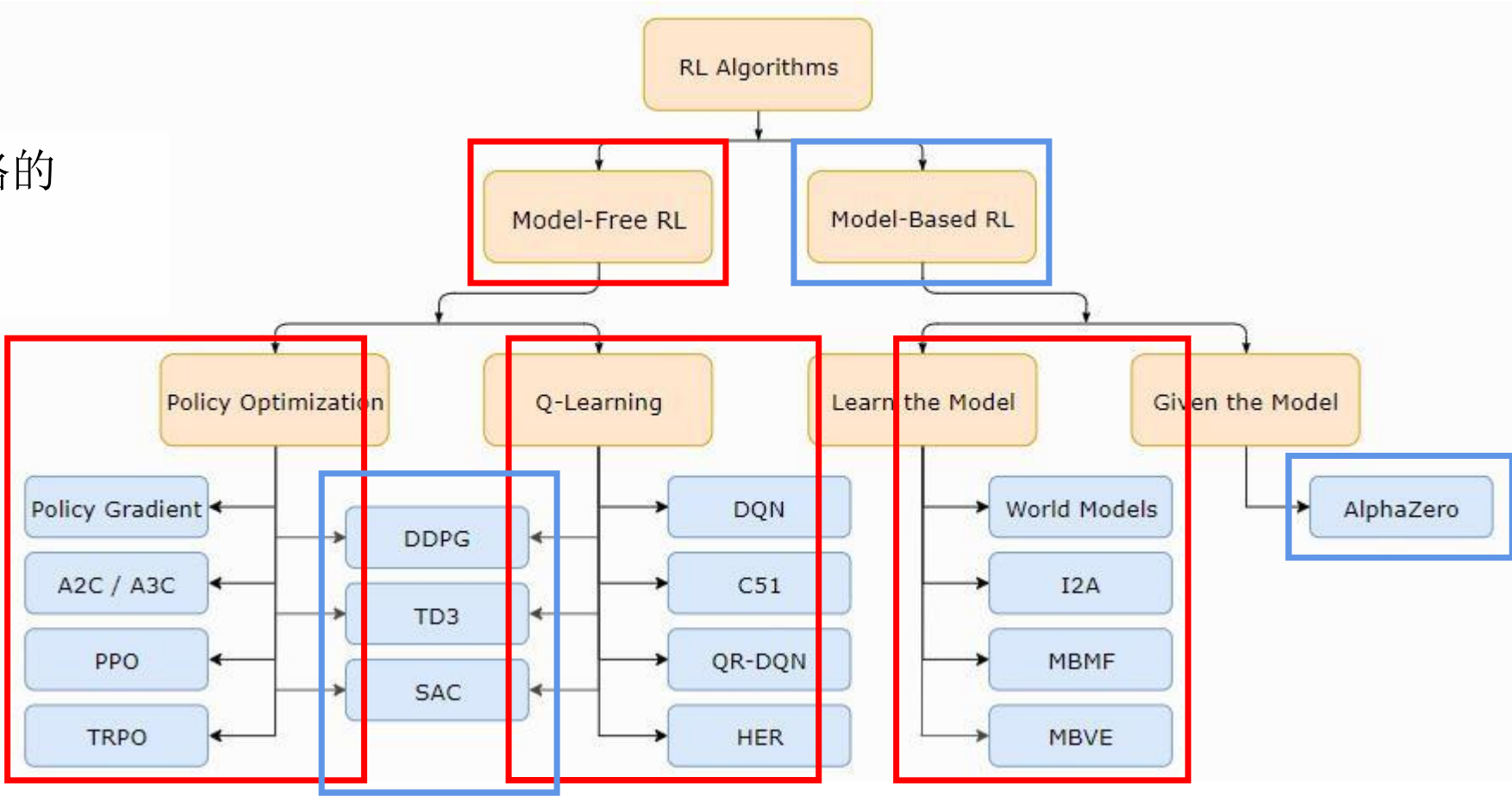
通过知识共享的方法进行的调查 在强化学习中使用动态网络

指数加速RL是否可能？

RL社区的进展

基于价值的与基于策略的

MBRL vs. MFRL



基于政策的问题
家

演员-评论

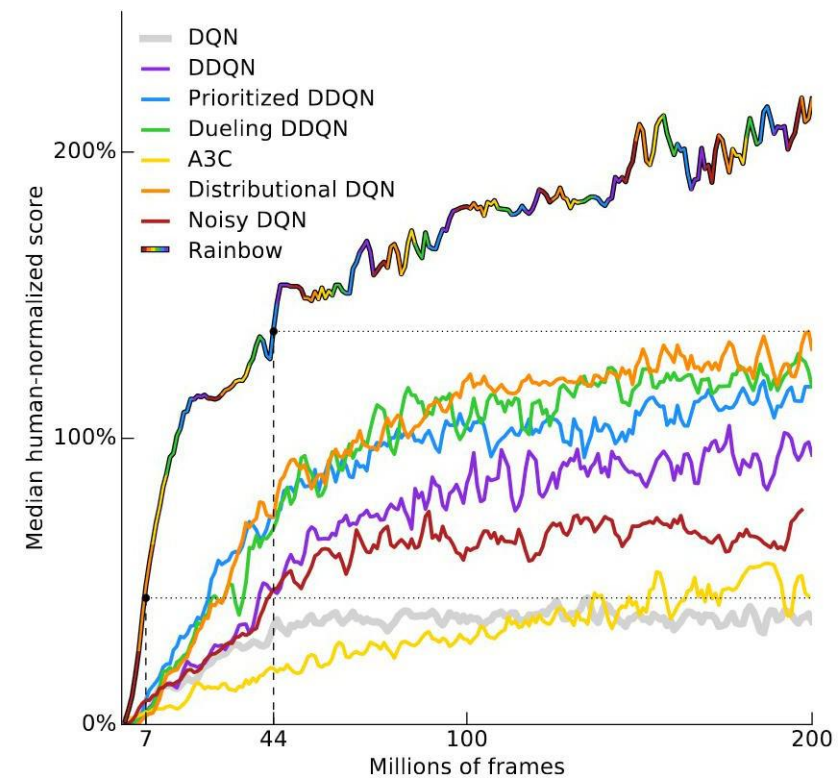
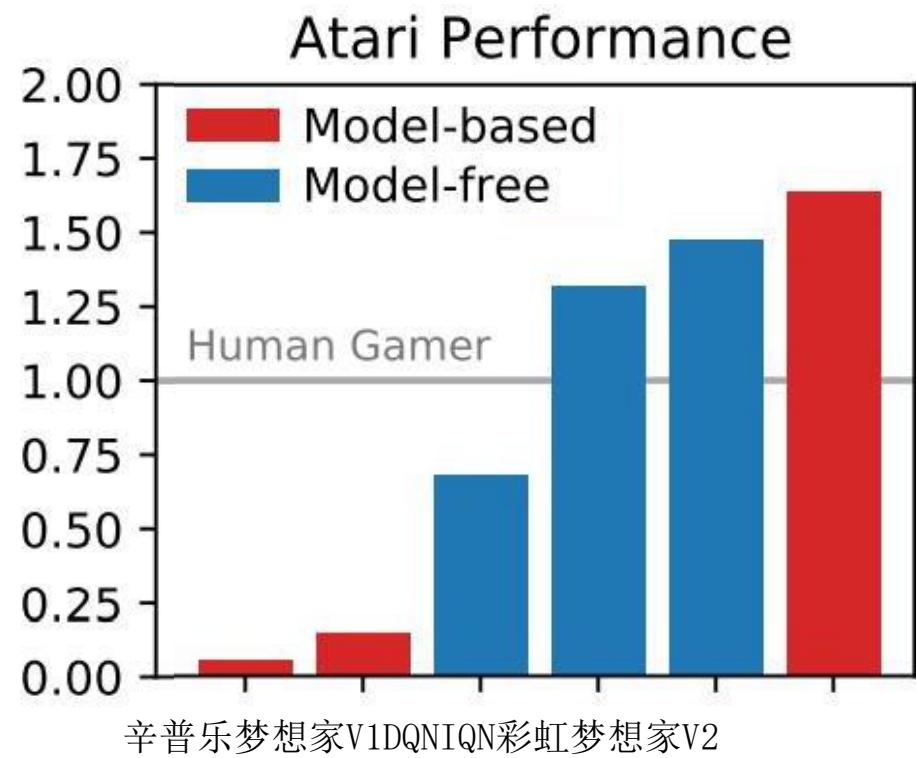
基于价值的
信息

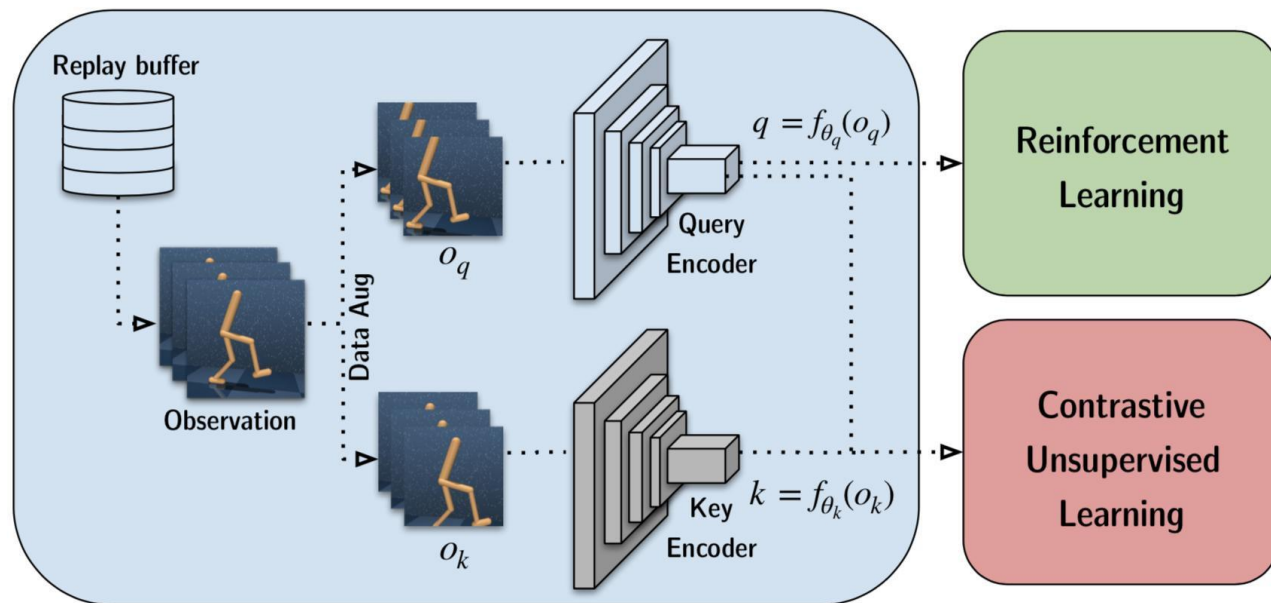
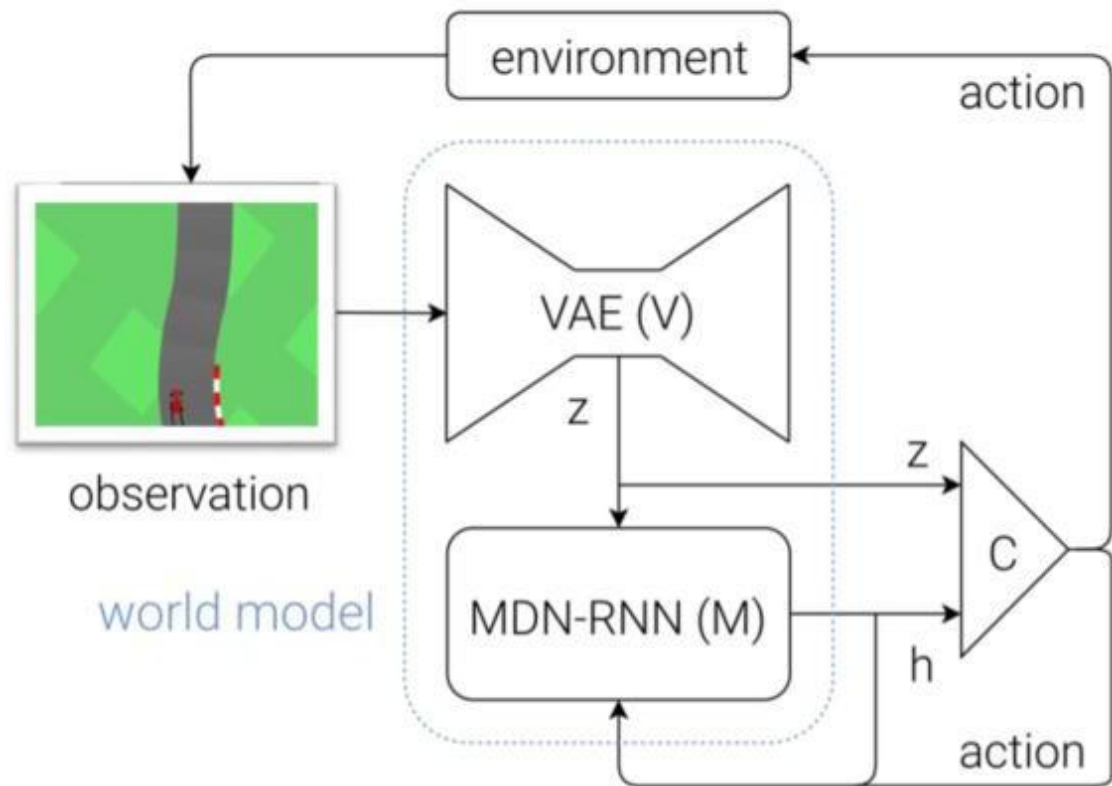
+梦想者V1V2

信用证: openAI主轴

（这里没有提到离线内容）

数据有效性限制了应用程序
而改进已经饱和了





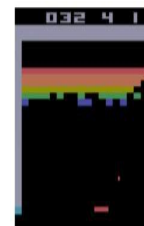
缺点的解决方案：

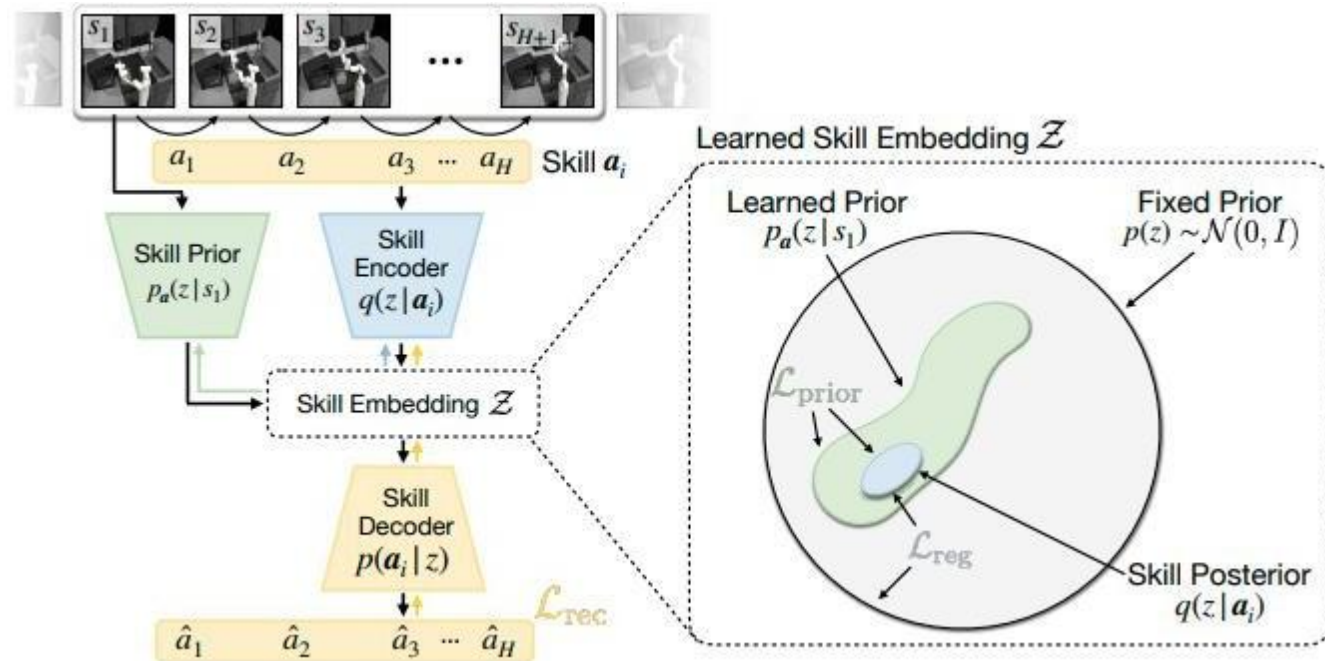
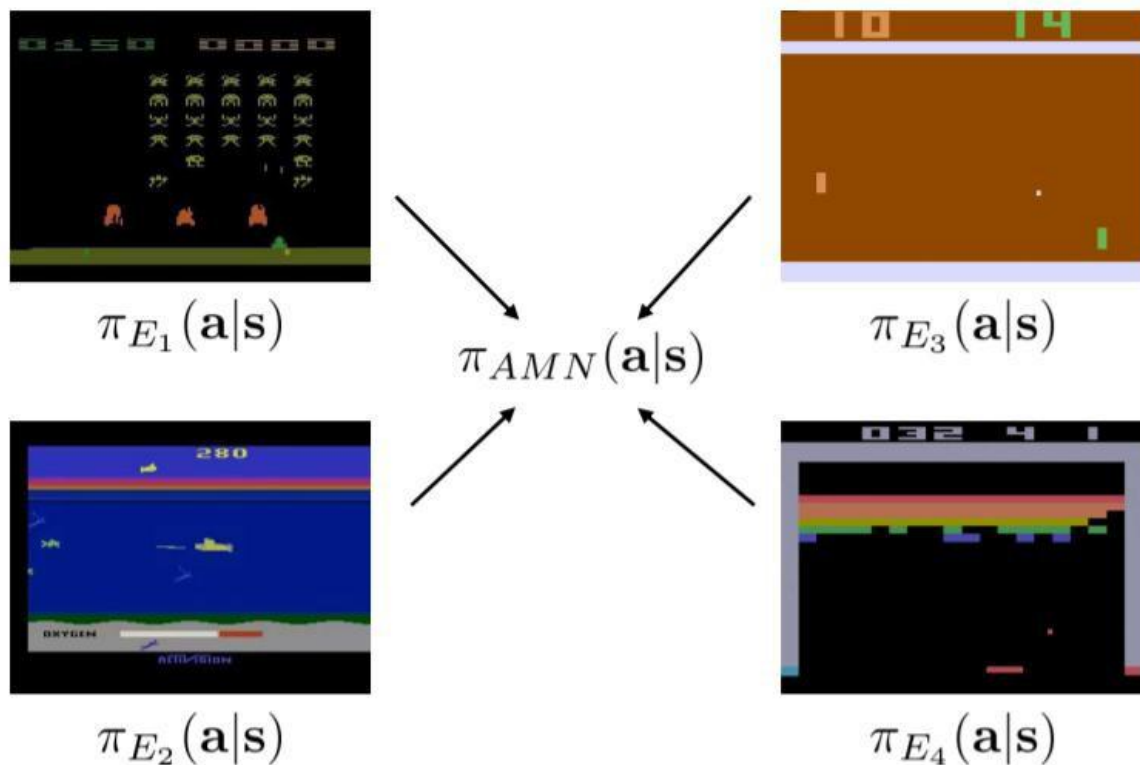
自我建模：世界模型

数据增强：视觉表示的RL表示对比学习

x30+的数据效率

单个代理可以玩单人游戏





除此之外，我们将做什么（下一代RL，NGRL

多任务的技能转移

代理应该共享跨场景的技能和理解

从蒙特祖马的复仇到连续的行动空间稀疏奖励机器人，与WM&CL

指数NxGRL（下一个吗？）

算法

$O(1)$ 、 $O(n)$ 、 $O(n\log n)$ 、 $O(n^2)$ 。

神经网络

训练
镜头
零镜头学习
...

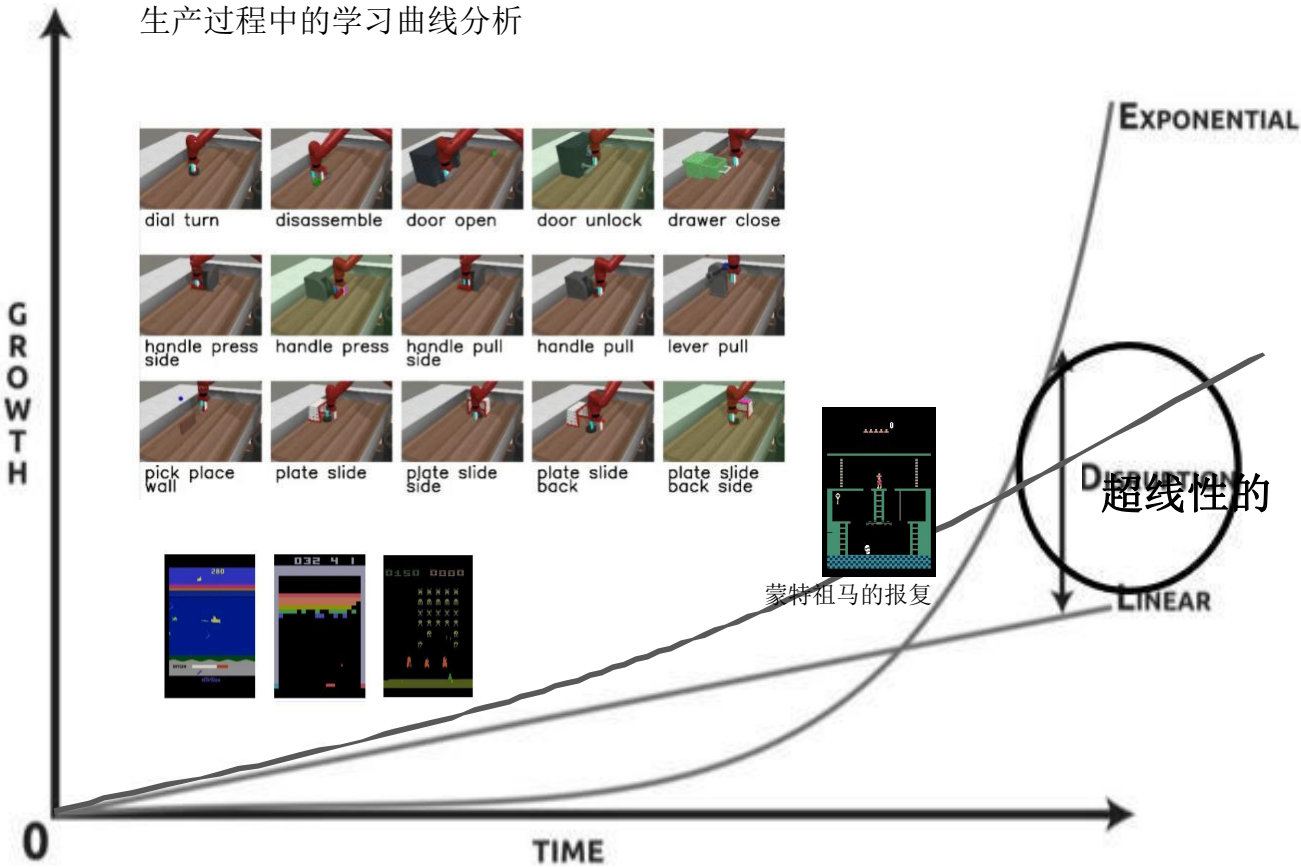
$$(e^x)' = e^x$$

从根本上说，经验/技能应该会加速整个任务/环境中的学习过程

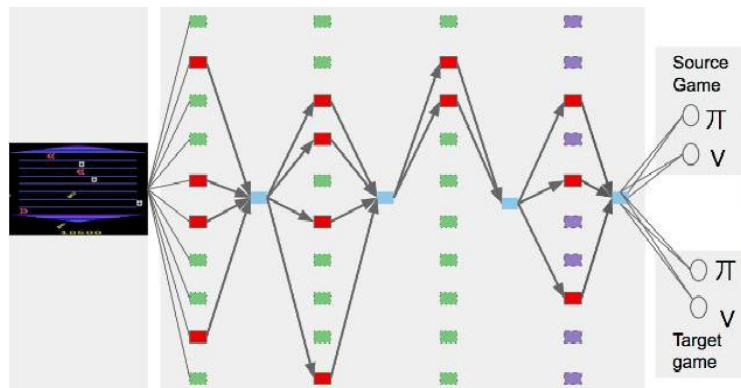
有了更多的经验，代理应该更快地学习任务。

如果它与经验成正比，

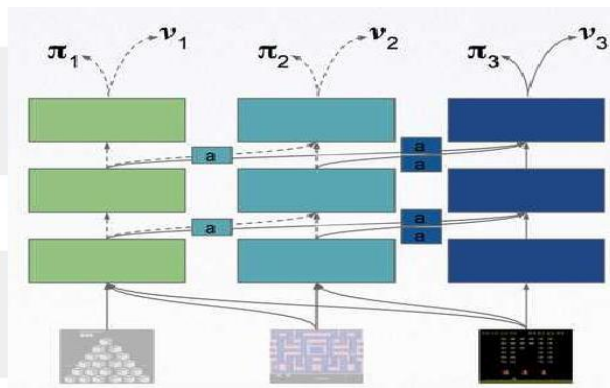
根据定义，它应该是一个指数学习曲线。



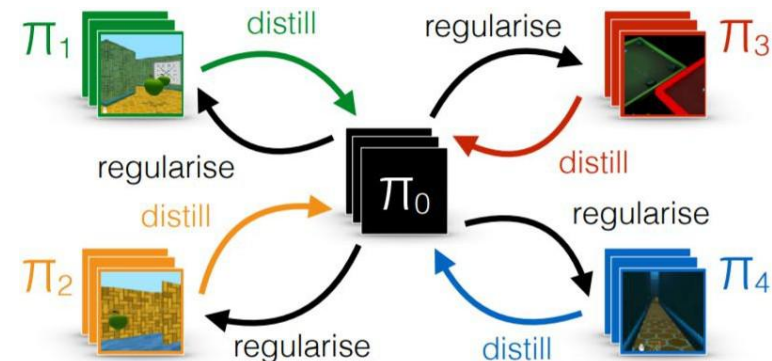
多任务RL现在已存在



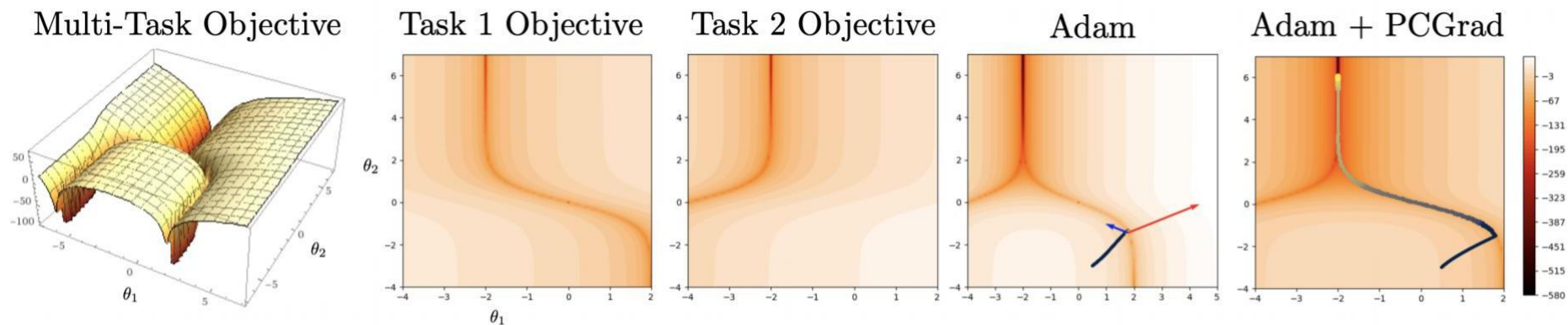
路径网



渐进式神经网络

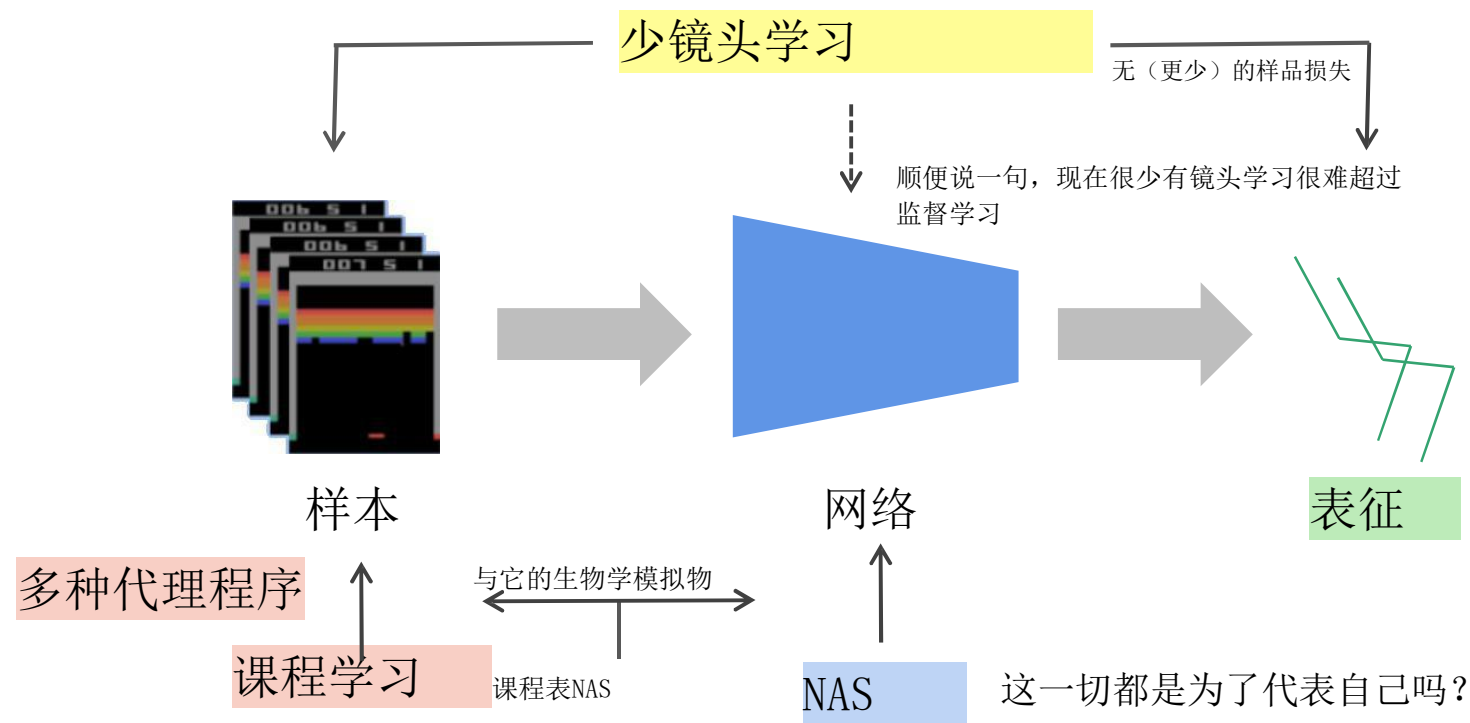


分歧：提炼和转移学习



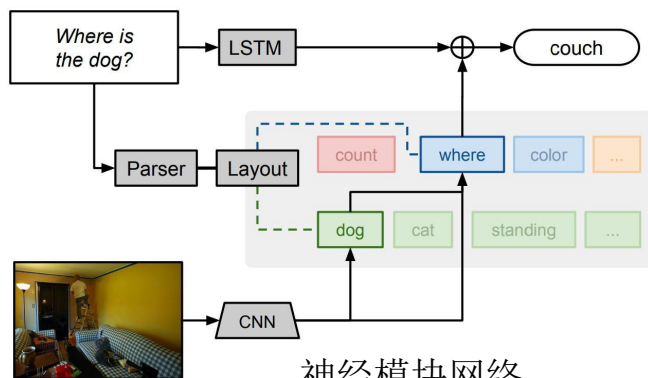
项目冲突梯度PopArt规范化

让我们再回到神经网络上来吧

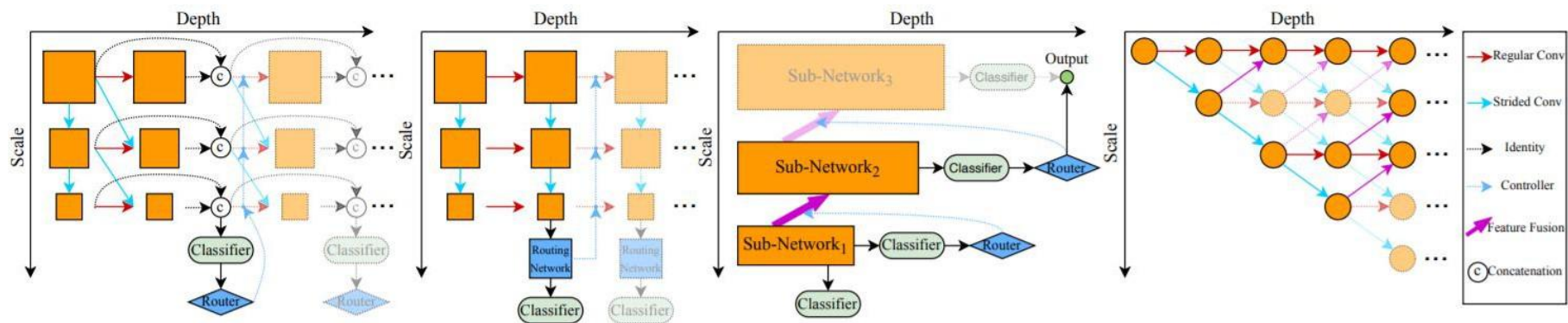
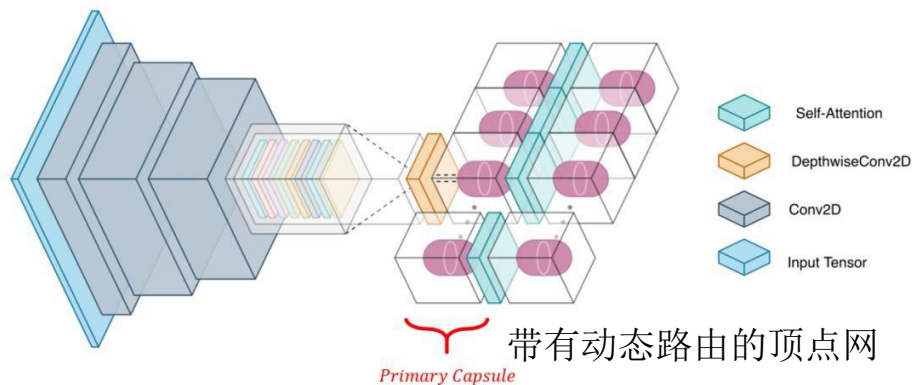


《打破太空爆炸的诅咒》，2020年，腾讯，arXiv: 2007.07197

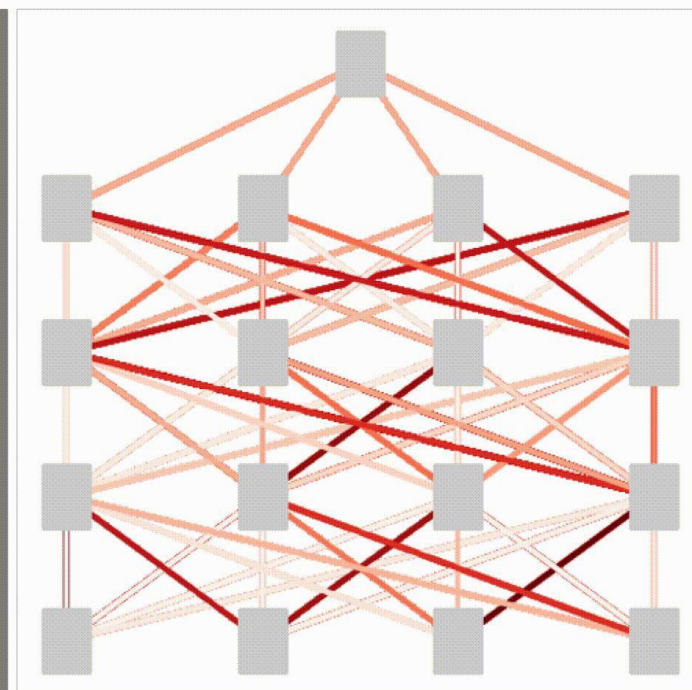
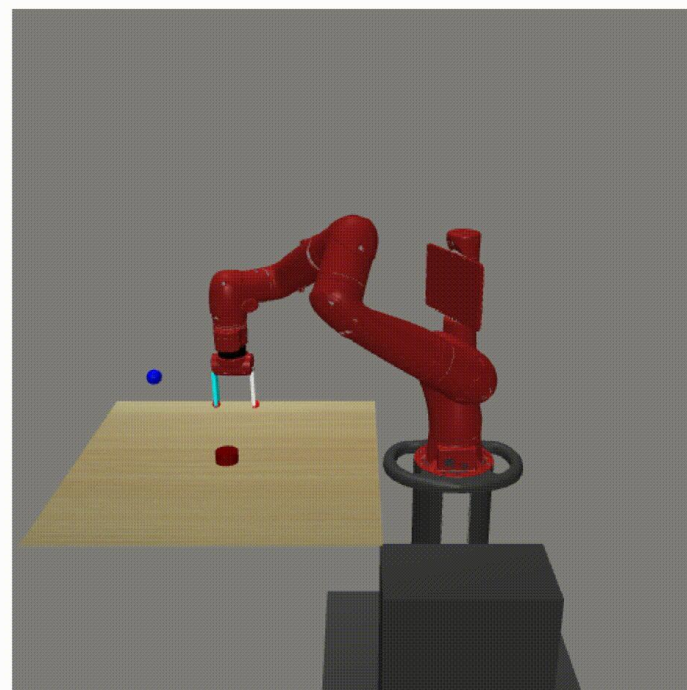
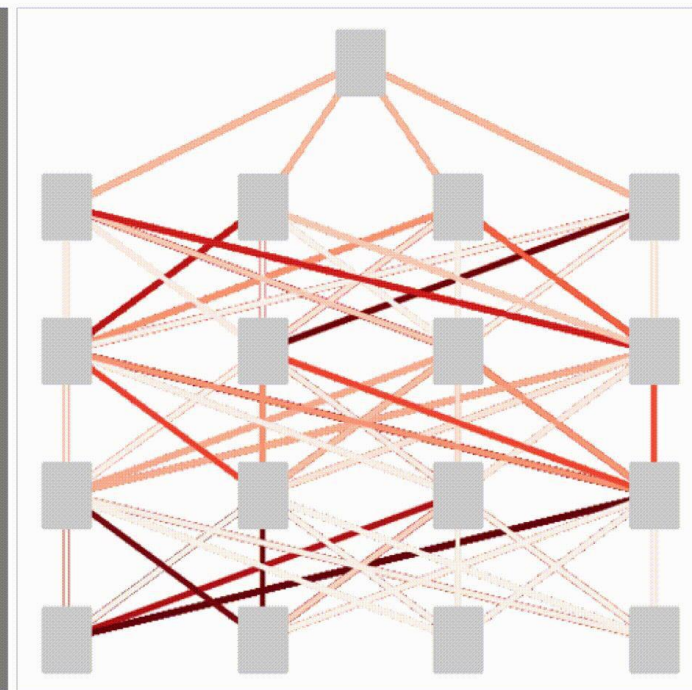
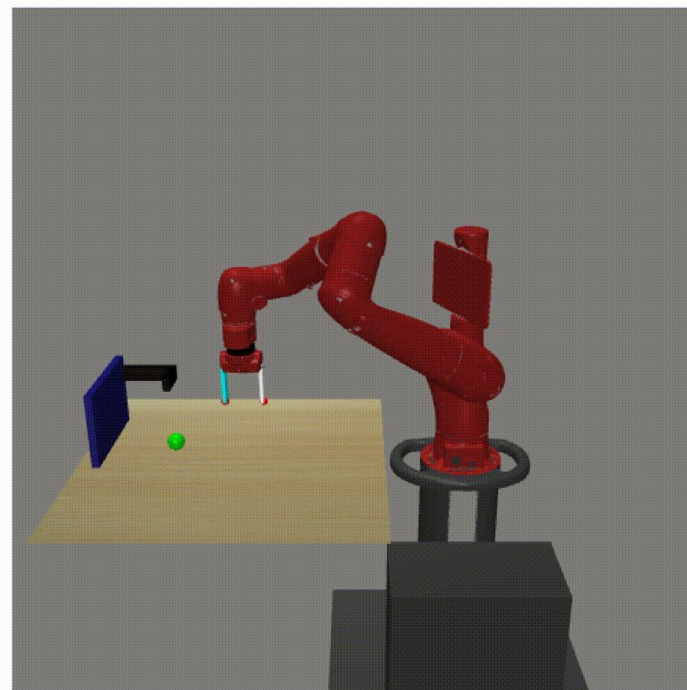
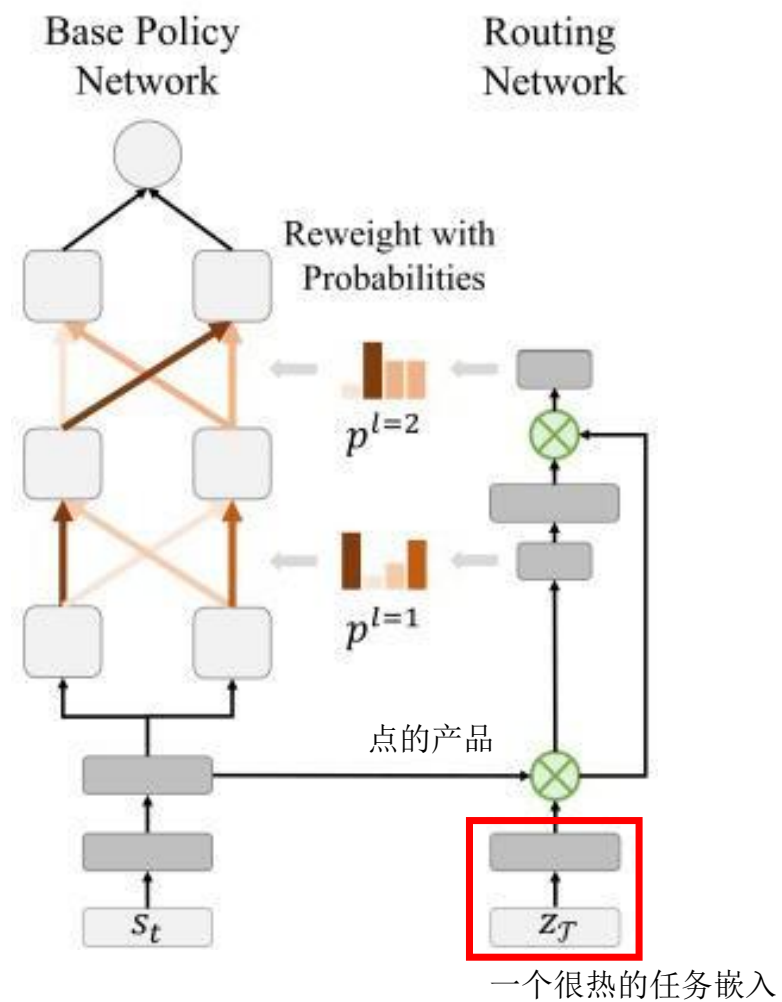
动态网络现在可用



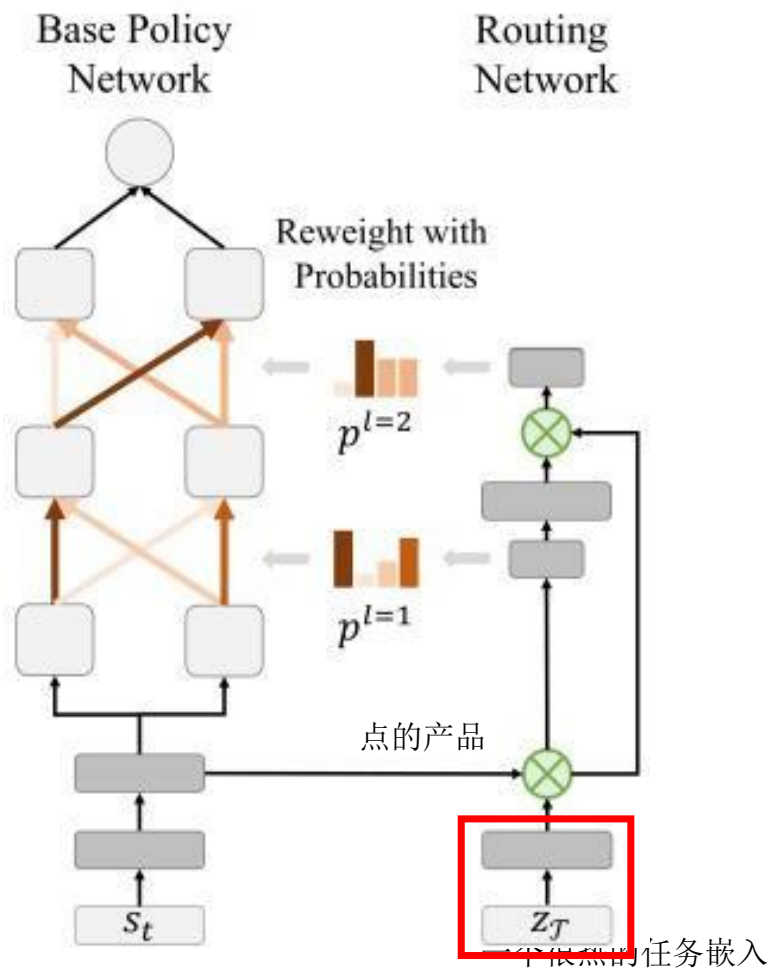
神经模块网络



主要是针对视觉任务而设计的



具有软模块化功能的多任务强化学习，



嵌入式嵌入

结合时间序列吗？

更好的任务嵌入式工作吗？包括层次结构的任务结构吗？任务的过程？这些选项之间的关系吗？

门关

网络架构的注意机制？动态的NAS系统吗？

稀疏性门关（以节省计算能力）？

模块化的功能

软模块代表什么？

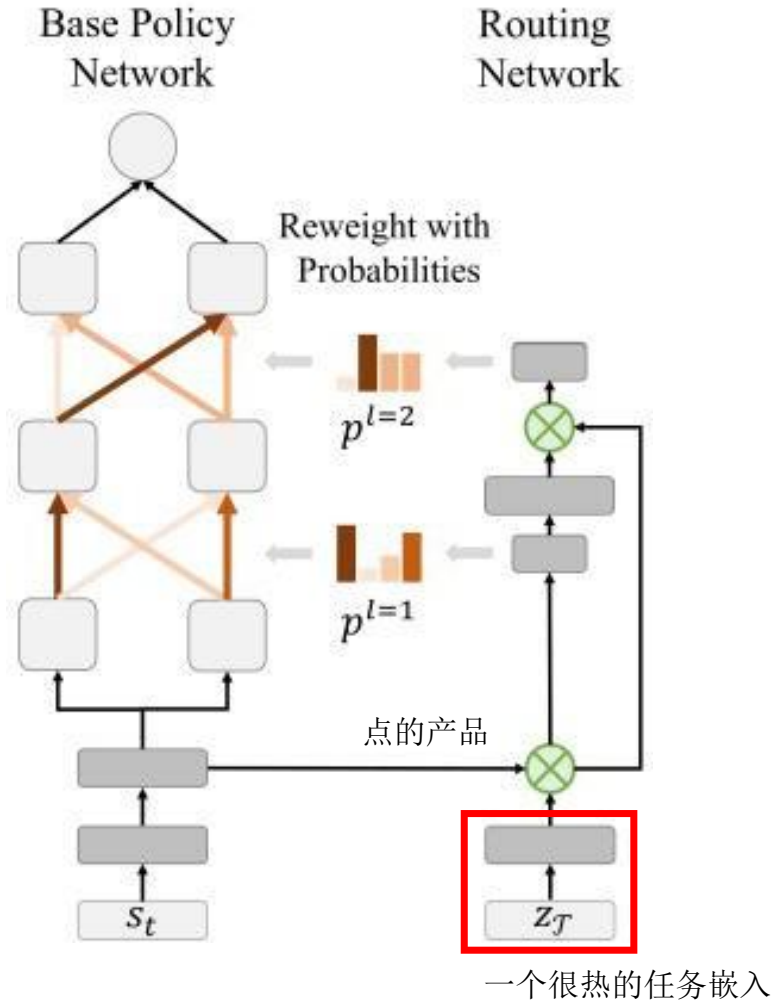
它们如何在跨任务中获得功能的

如何将此机制推广到代理的其他部分？

比如价值网络，世界模型嵌入等

一个继续/终身学习的范例？

关于ICRA2022年的主题



嵌入式嵌入

结合世界模型的时间序列吗？

更好的任务嵌入式工作吗？包括层次结构的任务结构吗？任务的过程？这些选项之间的关系吗？

门关

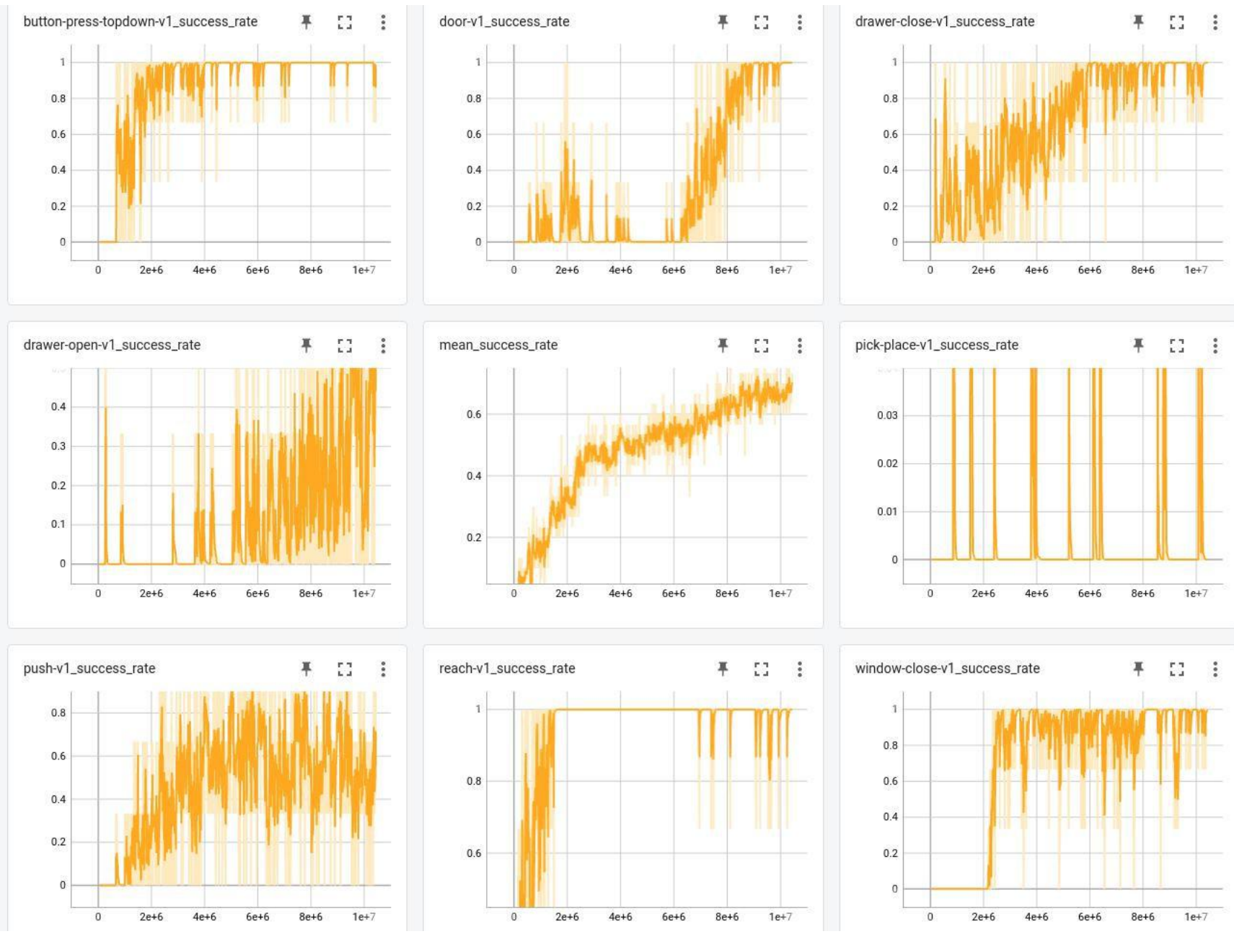
网络架构的注意机制？动态的NAS系统吗？
稀疏性门关（以节省计算能力）？

模块化的功能

软模块代表什么？
它们如何在跨任务中获得功能的
如何将此机制推广到代理的其他部分？
比如价值网络，世界模型嵌入等

一个继续/终身学习的范例？

主要的结果 关于软模块化技术





谢谢你！

