

彪哥带你学强化学习

13. 深入理解A3C算法

DEEPLY UNDERSTAND REINFORCEMENT LEARNING

讲师：韩路彪

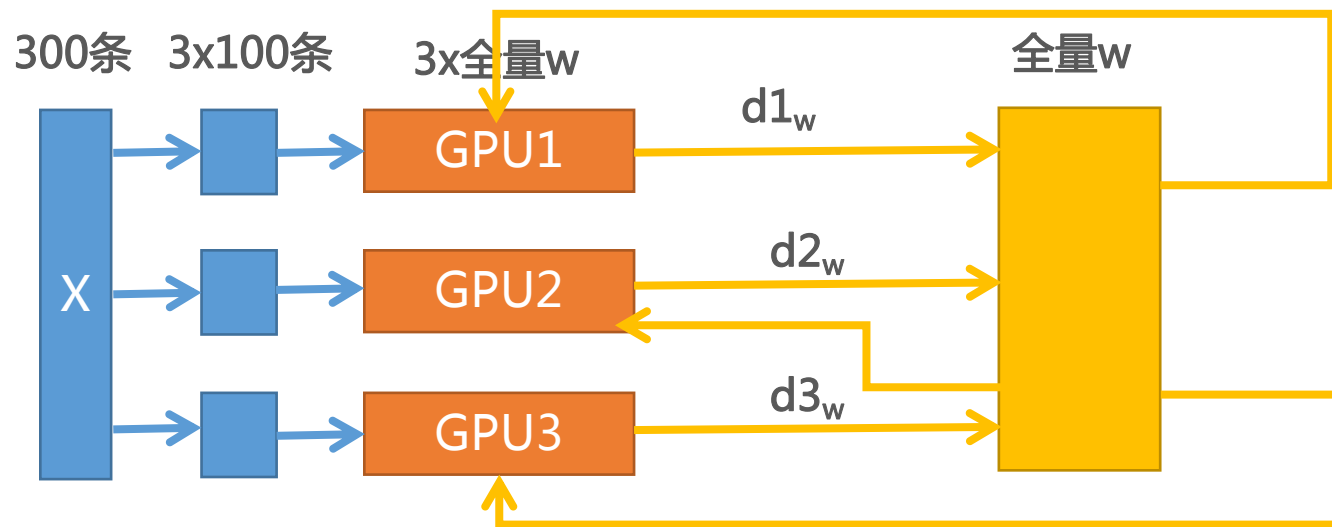


A3C 算法

A3C含义：Asynchronous Advantage Actor-Critic，异步优势演员评论家

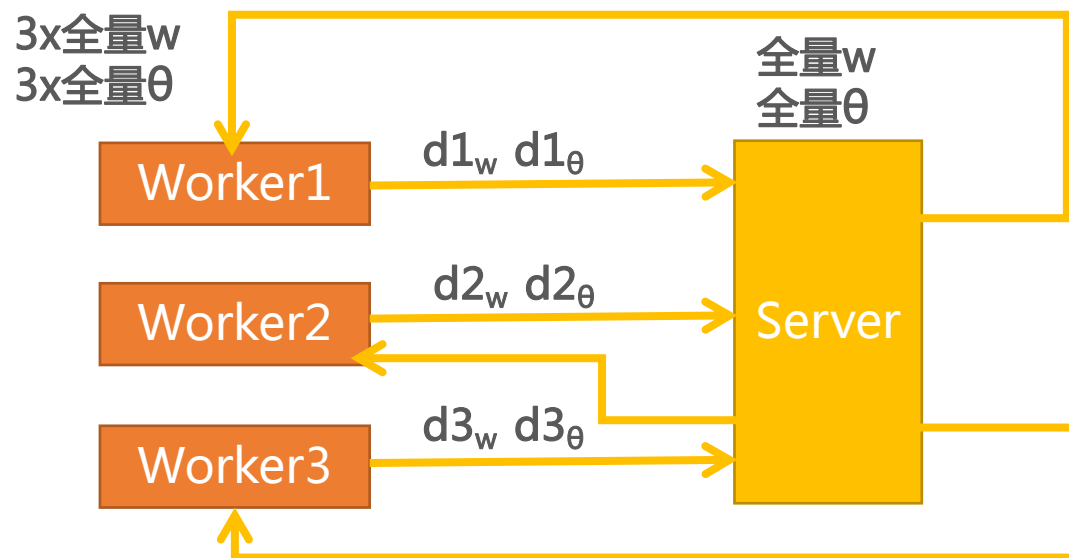
解决的问题：加快训练速度

思路：类似数据并行DP



不同点：训练样本来源

A3C 算法



每个worker单独采样，并按A2C算法计算出梯度，统一把梯度给server更新

worker向server同步什么内容？

worker什么时候向server同步？训练完1步还是训练完1个轨迹？

不同worker训练完1条轨迹后需要等待其他worker吗？