

2023.12.5周报

本周任务完成情况：

学习A2po算法的代码，阅读了论文：《ORDER MATTERS: AGENT-BY-AGENT POLICY OPTIMIZATION》，本论文做出的主要贡献如下：

1、我们利用提出的非策略修正方法PreOPC证明了在单次推出方案下，每个agent的策略可以保持单调改进的保证。进一步证明了在给定每个智能体的理论保证下，联合策略的单调界是单一推出算法中最紧密的，从而产生了有效的策略优化

2、

我们提出了A2PO，这是第一个逐个代理的顺序更新算法，它在每个代理的策略和联合策略上都保持单调的策略改进，并且在执行策略改进时不需要多次推出

3、

我们进一步研究了顺序策略更新方案、代理更新顺序和非平稳问题之间的联系，激发了两种新颖的方法:用于优化加速的半贪婪代理选择规则和用于缓解非平稳问题的自适应裁剪参数方法

下周任务计划完成情况：

详细阅读代码，争取彻底理解，跑出自己环境的实验。