阶段研究报告

**目前研究进展**

完成了NIPS 2020上发表的the state of the art多智能体强化学习算法LICA的复现工作，并对其提出的创新点自适应entropy进行了测试，其在星际争霸的好几个地图上都有很好的效果。

**下一步研究计划**

结合之前遇到的问题，LICA属于on-policy算法，对样本的利用效率较低，QMIX属于off-policy算法，对样本的效率利用率高，但是性能不是特别好，目前qmix的更新操作中有一个取max的操作，然而q值本来就是具有很大噪声的，对噪声取max，其结果必然会产生高估，因此目前打算先在简单场景上验证qmix的价值函数是高估的，加入double qmix或进一步研究，缓解高估问题。然后在复杂的星际环境中进行进一步的验证，并在部分地图上效果超过LICA。争取可以赶上AAAI 2022.