本周工作：

这周跑了20-30组实验，想把COMA加entropy的版本调出来，加entropy是为了提高智能体的探索，防止决策陷入局部最优，但是我加了之后反而陷入了局部最优。因为loss = coma\_loss + \alpha \* entropy，我分析可能是这个系数 \alpha比较大，所以我将coma\_loss 进行了均值为0，方差为1的归一化，发现学习速度可以提升了，但是还是会陷入局部最优，今天跑实验的时候发现同样的参数，跑出来的结果差的也还是挺多的，所以决定以后做实验跑多几组。

同时还基于COMA引入了自适应entropy，发现效果不是特别好，尝试采用对coma改进一下再用自适应entropy。

下周工作：

参照之前那篇nips的工作，完成对coma的改进，同时再做实验测试一下加entropy后的效果。