本周工作：

本周主要做了多智能体强化学习（MARL）的overview，MARL目前算法主要分为基于模型和不基于模型的，基于模型的主流算法是Counterfactual Regret Minimization (CFR)，在德州扑克游戏上已经打败了人类，不基于模型的主流算法是基于值的VDN、QMIX和基于策略的COMA，今年的ICML上还发现了一系列的基于这些算法的论文。这些算法的实验平台主要是星际争霸。

下周工作：

跑ROMA的代码，借鉴其是怎么修改模型以及代码的。