本周工作总结及计划：

1. 完成制动和保持距离2重场景下，中心式学习和分布式学习。用值分解方法（VDN）实现过程中，对于保持距离得场景无法得到收敛得结果。分析可知，VDN需要team reward 可分解成各个agent reward值得强假设。但保持距离得场景下，“保持”需要两车协同完成。因此在制动的场景下应用（各个agent目标相互独立）VND方法，实现了收敛的结果。中心式学习—分布式执行是MARL的学习范式，接下来用MADDPG方法实现保持距离的场景应用。并进一步将应用场景推广到协同保持车速，制动，超车等无人车控制常见场景上。
2. VND算法主要的不足，对值分解的线性假设。本质是agent reward对team reward的贡献问题，即 credit assignment。因此阅读了《2017-aaai-Counterfactual Multi-Agent Policy Gradients》。论文思想，将个体agent的行为，从联合行为中去除，选取其他动作代替，比较二者的差异（reward difference思想），从而界定个人行为对整体reward的贡献。有关划域的思想，就是credit assignment思想在大规模的agent场景下的进一步扩展。先将关系紧密的agent划为一个集合，在计算每个agent对整体的贡献。接下来继续阅读这部分文献。无人车场景有一个天然的划域规则就是车道线，但在超车，拥堵等场景下，各个车道线间会有协同或者竞争，思考如何根据这个特征建立模型。