**多智能体强化学习（MARL）计划安排**

* **计划安排表**

|  |  |  |
| --- | --- | --- |
| 持续时间 | 人员 | 计划安排 |
| 2-3个月 | 全部 | 共同阅读相关文献资料，规定在特定时间内阅读相同材料，定期安排讨论与分享（如一周两到三次），每月做一次月末总结，同时学习过程中要做必要的笔记。相关资料阅读内容，可大体按照如下次序进行：   1. Review 2. Framework 3. Joint action learning 4. Cooperation and competition 5. Coordination 6. Security 7. Self-Play 8. Learning to communicate 9. Transfer Learning 10. Imitation and Inverse RL 11. Meta Learning 12. Application   每部分至少精读优秀文献两篇，略读若干，以夯实基础，触类旁通，此乃“厚积”。 |
| 2个月 | 1-2人 | 前期选取合适仿真平台，学习认知架构，进行仿真研究。具体平台的选取暂定，可供挑选的仿真平台如下几种：   1. NetLogo 2. Repast 3. Arena 4. Neural MMO 5. 其他 |
| 1-2人 | 学习编程算法，在后期平台选取后可以尽快实现第一阶段文献中经典算法的复现。若凡灵光乍现，可推陈出新，便是“薄发” |
| 1个月 | 全部 | 机器人方面入门，为将多智能体算法从仿真转载到实物上做准备。 |
|  | 全部 | 购买设备，搭载多智能体的机器人平台，实现仿真到实物的过渡。期间对MARL理论的进一步深入研究可同时进行，并行不悖。希望理论与实践双管齐下，研究可更上一层楼。科研永无止境，倘若非要加上一个期限，我希望是一万年。 |

* **文献材料获取**

1. 在教材上，国外出版的有关多智能体强化学习的书籍有
2. [Deep Multi-Agent Reinforcement Learning](https://ora.ox.ac.uk/objects/uuid:a55621b3-53c0-4e1b-ad1c-92438b57ffa4) by Jakob N Foerster, 2018. PhD Thesis.
3. [Multi-Agent Machine Learning: A Reinforcement Approach](https://onlinelibrary.wiley.com/doi/book/10.1002/9781118884614) by H. M. Schwartz, 2014.
4. [Multiagent Reinforcement Learning](http://www.ecmlpkdd2013.org/wp-content/uploads/2013/09/Multiagent-Reinforcement-Learning.pdf) by Daan Bloembergen, Daniel Hennes, Michael Kaisers, Peter Vrancx. ECML, 2013.
5. [Multiagent systems: Algorithmic, game-theoretic, and logical foundations](http://www.masfoundations.org/download.html) by Shoham Y, Leyton-Brown K. Cambridge University Press, 2008.
6. 在文献上，一些顶会、顶刊的论文自不待言，除此之外，GitHub上有一个MARL的论文收集项目在使用时可能比较方便，第一阶段的阅读顺序也是出自于此：<https://github.com/LantaoYu/MARL-Papers>，。
7. 在视频学习上，bilibili和YouTube上有诸如斯坦福大学或DeepMind等的RL课程视频，可统一安排学习。

* **仿真平台初步介绍**

1. **NetLogo**：如官网所说，NetLogo是一个“multi-agent programmable modeling environment”，可以模拟自然现象和社会现象的建模环境，特别适合对复杂系统进行建模，并且编程语法非常简单，功能却非常强大，此外有着丰富的文档、教程以及模型库。
2. **Repast**：Repast Suite平台已经持续开发了15年以上，是一个有着一系列高级，免费和开放源代码的基于代理的建模和仿真平台。
3. **Arena**：Arena是第一个用于多主体智能研究的通用评估平台，通过借助构建多智能体环境的构建工具包，允许轻松构建新的多智能体问题。
4. **Neural MMO**： Neural MMO是OpenAI发布的—个强化学习的大型多智能体游戏环境。这一平台可以在持久性和开放式任务中支持大量可变数量的智能体。这一多智能体的环境可以探索更兼容和高效的整体环境，力求在复杂度和智能体人数上获取难得的平衡。

仿真平台进一步的选取需要等第一阶段工作完成后再进行，彼时若发现其他合适的平台环境，也可加入考虑范围内。

* **软硬件环境购买**

1. 前期主要的任务是文献的阅读，大部分教材和资料都可以在网络上寻得，因此并没有特别的花费。
2. 中期的仿真平台的搭建与模拟，大部分仿真平台开源免费，只有对硬件的要求提出了一定的要求（但可能也不是太高，比如Neural MMO，只用单个桌面CPU就可以训练有效的策略）。
3. 后期实物机器人平台的搭载可能是主要的花费项，这体现在对机器人功能上提出了要求，但是具体要实现哪些功能以及选取什么种类的机器人还需进一步研究后确定。

综上，在软硬件购买上，主要体现在中期运算硬件和后期机器人平台的搭建上，但是目前看来，由于处于前期，这些东西还不必急于购买，等第一阶段的文献阅读完成后再做挑选即可。

* **其他**

由于目前处于前期工作，只对文献阅读方面做了相对较详细的安排，在后面进一步的学习之中，将务必对之后的工作做更加详细的安排与规划。

此外，计划倘若不能严格执行，便是一纸空文，因此必须严于律己。而当现实情况与计划安排出现较大出入时，可以在共同讨论下对其做适当的修正。

* **寄语**

王安石说，世之奇伟、瑰怪，非常之观，常在于险远，而人之所罕至焉，故非有志者不能至也。今我等当做有志之人，莫做碌碌之辈，宵衣旰食，披荆斩棘，虽尽吾志而不能至者，可以无悔矣，其孰能讥之？MARL之路任重而道远，吾将上下而求索。