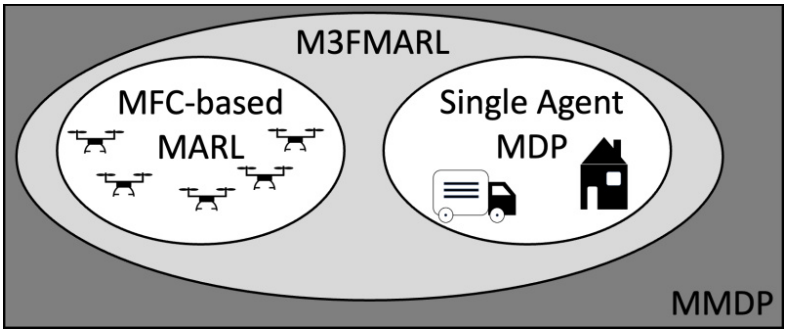
Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Major-Minor Mean Field Multi-Agent Reinforcement Learning



Project Manager
Kai Cui

Principal Investigator
Prof. Dr. Heinz Koeppl

Project Term
2023 - 2024

Clusters
Lichtenberg II Cluster Darmstadt

Software
PyTorch

Institute
Self-Organizing Systems

University
Technische Universität Darmstadt

Introduction

Multi-agent reinforcement learning (MARL) remains difficult to scale to many agents. Recent MARL using Mean Field Control (MFC) provides a tractable and rigorous approach to otherwise difficult cooperative MARL. However, the strict MFC assumption of many independent, weakly-interacting agents is too inflexible in practice. We generalize MFC to instead simultaneously model many similar and few complex agents – as Major-Minor Mean Field Control (M3FC). Theoretically, we give approximation results for finite agent control, and verify the sufficiency of stationary policies for optimality together with a dynamic programming principle. Algorithmically, we propose Major-Minor Mean Field MARL (M3FMARL) for finite agent systems instead of the limiting system. The algorithm is shown to approximate the policy gradient of the underlying M3FC MDP. Finally, we demonstrate its capabilities experimentally in various scenarios. We observe a strong performance in comparison to state-of-the-art policy gradient MARL methods. Our contribution can be summarized as follows: Existing MFC cannot model general agents and many aggregated agents simultaneously. In essence, we generalize the solution spaces of single-agent RL and MFC-based MARL – frameworks for cooperative MARL as depicted in Figure 1. This provides both tractability for many aggregated agents and generality for arbitrary general agents. Our contribution is briefly summarized into (i) formulating the first discrete-time MFC model with major agents, together with establishing its theoretical properties; (ii) providing a MFC-based MARL algorithm, which in contrast to prior work learns on the finite problem of interest; and (iii) we perform a significant empirical evaluation, also obtaining positive comparisons of MFCbased MARL against state of the art, whereas prior works on MFC were limited to verifying algorithms on one or two examples.

Methods

To begin, we extend standard MFC by modelling the presence of a major agent. The generalization to more than one major agent

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

is straightforward. This leads to our discretetime major-minor MFC (M3FC) model. Overall, we obtain a formulation that allows standard MARL handling of major agents, while tractably handling many minor agents via MFC-based techniques.

Results

As a first step, it is well known that stationary (time-independent) policies suffice for optimality in infinite-horizon discounted MDPs. This property is also verified for the M3FC MDP. For the following technical results, we assume standard Lipschitz conditions. To show the approximate optimality of M3FC solutions, we first obtain propagation of chaos – convergence of empirical MFs to the limiting MF. The result theoretically backs the reduction of multi-agent control to single-agent MDPs, as there is no loss of optimality in the finite problem by considering the M3FC problem. We assume standard Lipschitz conditions on policies. The above motivates M3FC by the following near optimality result of M3FC MDP solutions in the finite system, as it suffices to optimize over stationary M3FC policies. Therefore, one may solve difficult finite-agent MARL by detouring over the corresponding M3FC MDP as depicted in Figure 2, reducing to an MDP of a complexity independent of the number of agents N. Recalling the motivation of MFC, it is crucial to find tractable sample-based MARL techniques for both complex problems where other methods fail, and for problems where we have no access to the dynamics or reward model. In our work we analyze the proposed algorithm not on limiting M3FC MDPs, but on the more interesting finite M3FC system. In particular, if the M3FC MDP is known, one can instantiate finite systems of any size for training. We consider the following perspective: By our theory, the M3FC MDP is approximated well by the finite system. Therefore, we solve the limiting M3FC MDP by applying our proposed algorithm directly to finite M3FC systems.

Discussion

We have proposed a generalization of MDPs and MFC, enabling tractable state-of-theart MARL on general many-agent systems, with both theoretical and empirical support. Beyond the current model and its optimality guarantees, one could work on extended optimality conjectures, refined approximations, and local interactions. Algorithmically, M3FC MDP actions could move beyond binning the state space to gain performance, e.g. via kernels. Lastly, one may try to quantify convergence to the rate $O(1/\sqrt{N})$ for non-finite state spaces, as the current proof would need hard-to-verify assumptions.

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

Figures

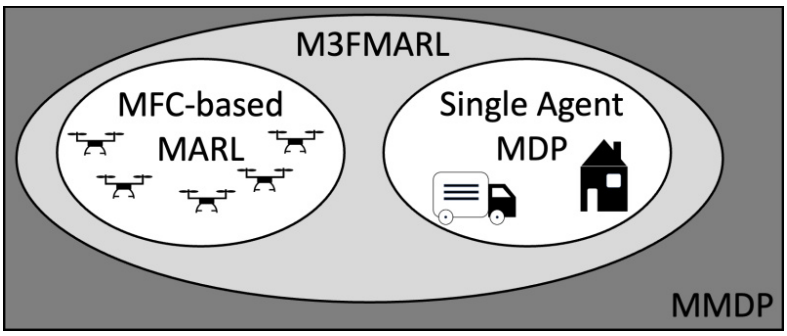


Figure 1: Our M3FC-based MARL generalizes MFC-based MARL and standard singleagent RL in the solution space of general MARL solutions, reducing the otherwise combinatorial nature of MARL to a tractable but still general setting.

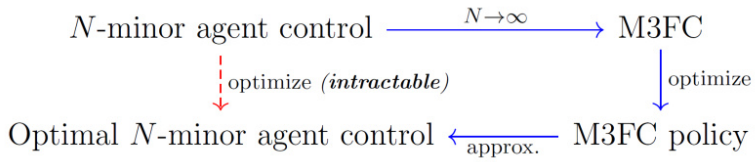


Figure 2: Approximation of intractable N-agent control by M3FC (blue path), the solution of which is near-optimal for large *N*.

Publications

Cui, K.; Fabian, C.; Tahir, A.; Koeppl, H.: “Major-Minor Mean Field Multi-Agent Reinforcement Learning,” in International Conference on Machine Learning (ICML) 2024. https://doi.org/10.48550/arXiv.2303.10665

Cui, K., International Conference on Machine Learning (ICML), Vienna, Austria, July 21 - July 27, 2024

Last Update: 2024-10-29 15:22