

# Social Behavior as a Key to Learning-Based Multi-Agent Pathfinding Dilemmas (Abstract Reprint)

Chengyang He<sup>1</sup>, Tanishq Duhan<sup>1</sup>, Parth Tulsyan<sup>1</sup>, Patrick Kim<sup>1</sup>, Guillaume Sartoretti<sup>1</sup>

<sup>1</sup>Department of Mechanical Engineering, College of Design and Engineering, National University of Singapore, Singapore

**Abstract Reprint.** This is an abstract reprint of the journal article by He, Duhan, Tulsyan, Kim, and Sartoretti (2025).

## Abstract

The Multi-agent Path Finding (MAPF) problem involves finding collision-free paths for a team of agents in a known, static environment, with important applications in warehouse automation, logistics, or last-mile delivery. To meet the needs of these large-scale applications, current learning-based methods often deploy the same fully trained, decentralized network to all agents to improve scalability. However, such parameter sharing typically results in homogeneous behaviors among agents, which may prevent agents from breaking ties around symmetric conflict (e.g., bottlenecks) and might lead to live-/deadlocks. In this paper, we propose SYLPH, a novel learning-based MAPF framework aimed to mitigate the adverse effects of homogeneity by allowing agents to learn and dynamically select different social behaviors (akin to individual, dynamic roles), without affecting the scalability offered by parameter sharing. Specifically, SYLPH offers a novel hierarchical mechanism by introducing Social Value Orientation (SVO) as a temporally extended latent variable that plays a central role in both policy generation and reward assignment. To support this hierarchical decision-making process, we introduce Social-aware Multi-Policy PPO (SMP3O), a reinforcement learning method that ensures stable and effective training through a mechanism for the cross-utilization of advantages. Moreover, we design an SVO-based learning tie-breaking algorithm, allowing agents to proactively avoid collisions, rather than relying solely on post-processing techniques. As a result of this hierarchical decision-making and exchange of social preferences, SYLPH endows agents with the ability to reason about the MAPF task through more latent spaces and nuanced contexts, leading to varied responses that can help break ties around symmetric conflicts. Our comparative experiments show that SYLPH achieves state-of-the-art performance, surpassing other learning-based MAPF planners in random, room-like, and maze-like maps, while our ablation studies demonstrate the advantages of each component in SYLPH. We finally experimentally validate our trained

policies on hardware in three types of maps, showing how SYLPH allows agents to find high-quality paths under real-life conditions. Our code and videos are available at: [marmotlab.github.io/mapf\\_sylph](https://marmotlab.github.io/mapf_sylph).

## References

He, C.; Duhan, T.; Tulsyan, P.; Kim, P.; and Sartoretti, G. 2025. Social behavior as a key to learning-based multi-agent pathfinding dilemmas. *Artificial Intelligence*, 348: 104397.