Diffusion-Guided Multi-Arm Decentralized Motion Planning

Viraj Parimi[1], Annabel Gomez[1], Howard Chen[2], Andreas Hoffman[3], Brian C. Williams[1]
[1]MIT [2]Imperial College [3]Aescape

Generate motion plans for multiple arms in a decentralized fashion using diffusion models guided by reinforcement learning

Motivation

- Propose a unified multi-arm diffusion policy capable of handling dynamic obstacles (e.g moving robots), eliminating the need for traditional roadmaps and extensive collision annotations
- Leverage diffusion models' generative capabilities to directly produce feasible actions for agents without relying on precomputed roadmaps
- Capture multi-modal distributions prevalent in multi-arm tasks during learning
- Use diffusion model to predict anticipated actions of neighbors for improved planning

Next Steps

- Receding horizon control for reactive execution
- Combine learned diffusion model with multi-agent path finding planners to improve scalability for a larger number of agents

References

- Jingkai Chen, Jiaoyang Li, Yijiang Huang, Caelan Garrett, Dawei Sun, Chuchu Fan, Andreas Hofmann, Caitlin Mueller, Sven Koenig, and Brian C. Williams. Cooperative task and motion planning for multi-arm assembly systems, 2022.
- Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion, 2024.
- Huy Ha, Jingxi Xu, and Shuran Song. Learning a decentralized multi-arm motion planner, 2020.
- Zhendong Wang, Jonathan J Hunt, and Mingyuan Zhou. Diffusion policies as an expressive policy class for offline reinforcement learning, 2023.
- Zhengbang Zhu, Minghuan Liu, Liyuan Mao, Bingyi Kang, Minkai Xu, Yong Yu, Stefano Ermon, and Weinan Zhang. Madiff: Offline multi-agent learning with diffusion models, 2024.

Approach

- Extend diffusion models performance in out-of-distribution settings using reinforcement learning
  - Integrate actor-critic framework where the critic evaluates the quality of actions improving model's robustness without limiting model expressiveness
- Utilize UNet architecture as noise-predictor network
  - MLP-based architecture hindered actor performance
- Combine multi-arm observations into a single state-encoder using LSTM cells