ABOUT ME

Jiayu Chen

2023-09-26

Self-introduction

- I'm Jiayu Chen, a Ph.D. candidate at the school of Industrial Engineering (Operation Research), Purdue University (Aug 2020 -- May 2024).
- I obtained my Bachelor Degree from College of Engineering, Peking University (Sep 2016 – Jul 2020).
- My research focuses on the algorithm design and applications of reinforcement learning (RL).
- My current research focus is unsupervised reinforcement learning and hierarchical imitation learning (IL).

Unsupervised RL

- In typical RL, agents are trained with task-specific rewards. There could be several disadvantages:
  - It is usually challenging to design effective, task-related reward functions.
  - For each new task, an agent needs to be retrained from scratch, which is resource-consuming.

- Unsupervised RL allows the agent to gain prior knowledge of the environment, without access to the reward function or any task-specific information. The learning outcome (e.g., temporal abstractions or thorough exploration) can benefit multiple downstream tasks in the same environment and has the potential to be transferred into new scenarios.

Unsupervised RL

- Unsupervised skill discovery allows the agent to discover task-agnostic skills while interacting with the environment. These skills can be viewed as sub-policies (i.e., policy segments), and adopted to policy learning in multiple downstream tasks. The agent can now select (make decisions) among pre-learned skills rather than primitive actions, which can greatly improve the learning efficiency.

- Based on the objective for skill discovery, the related algorithm can be generally divided into two categories: Laplacian-based skill discovery and variational skill discovery. We have works done in each category.

Hierarchical IL

- Imitation Learning enables the agent to recover a policy from a dataset of expert demonstrations, through matching the state-action distribution from the trained policy and the expert dataset.

- Hierarchical IL involves skills in the learning process. Given that long-horizon, complex tasks usually contain completion of a series of subtasks, the agent should form skills for each subtask and, in the meantime, learn a high-level policy to coordinate the switch of these skills. Thus, the learning target is now a hierarchical policy, of which the low-level parts correspond to the skills.

- Compared with standard IL, Hierarchical IL requires the agent to identify the potential subtask decomposition and learn skills for each subtask, which is a lot more challenging in the algorithm design. We have a work done in this direction, and another work as an extension to multi-task learning scenarios.

Publications

- **Unsupervised RL:**

- **Jiayu Chen,** Marina Wagdy Wadea Haliem, Tian Lan, and Vaneet Aggarwal, "Multi-agent Deep Covering Option Discovery", accepted in Proc. ICML Reinforcement Learning for Real Life Workshop, Jul 2021.

- **Jiayu Chen,** Jingdi Chen, Tian Lan, and Vaneet Aggarwal, "Learning Multi-agent Options for Tabular Reinforcement Learning using Factor Graphs", accepted to IEEE Transactions on Artificial Intelligence (TAI), Jul 2022.

- **Jiayu Chen,** Jingdi Chen, Tian Lan, and Vaneet Aggarwal, "Scalable Multi-agent Option Discovery based on Kronecker Graphs", accepted in Proc. NeurIPS, Dec 2022.

- **Jiayu Chen,** Vaneet Aggarwal, and Tian Lan, "A Unified Algorithm Framework for Unsupervised Discovery of Skills based on Determinantal Point Process", accepted in Proc. NeurIPS 2023.

Publications

- **Hierarchical IL:**

- **Jiayu Chen,** Tian Lan, and Vaneet Aggarwal, “Option-Aware Adversarial Inverse Reinforcement Learning for Robotic Control”, accepted in Proc. IEEE ICRA, Jun 2023.

- **Jiayu Chen,** Vaneet Aggarwal, and Tian Lan, ``Hierarchical Adversarial Inverse Reinforcement Learning", accepted in IEEE Transactions on Neural Networks and Learning Systems (TNNLS), Aug 2023. (A completed version of the first one.)

- **Jiayu Chen,** Dipesh Tamboli, Tian Lan, and Vaneet Aggarwal, “Multi-task Hierarchical Adversarial Inverse Reinforcement Learning”, accepted in Proc. ICML, Jul 2023.

Publications

- **Applications of RL and IL:**

- Pin Wang, Dapeng Liu, **Jiayu Chen,** Hanhan Li, and Ching-Yao Chan, "Decision Making for Autonomous Driving via Augmented Adversarial Inverse Reinforcement Learning", accepted in Proc. IEEE ICRA, Jun 2021. (Autonomous Driving)

- **Jiayu Chen,** Abhishek K. Umrawal, Tian Lan, and Vaneet Aggarwal, "DeepFreight: A Model-free Deep-reinforcement-learning-based Algorithm for Multi-transfer Freight Delivery", accepted in Proc. ICAPS, Aug 2021. (Intelligent Transportation System)

- Dipesh Tamboli, **Jiayu Chen,** Kiran Pranesh Jotheeswaran, Denny Yu, Vaneet Aggarwal, "Reinforced Sequential Decision-Making for Sepsis Treatment: The PosNegDM Framework with Mortality Classifier and Transformer", submitted to IEEE Journal of Biomedical and Health Informatics. (Healthcare)

- Chang-Lin Chen, Hanhan Zhou, **Jiayu Chen, et. al,** "Learning-based Two-tiered Online Optimization of Region-wide Datacenter Resource Allocation", submitted to IEEE Transactions on Network and Service Management. (Datacenter Resource Allocation)

Publications

- **Other works:**

- **Jiayu Chen,** Tian Lan, and Vaneet Aggarwal, “Hierarchical Deep Counterfactual Regret Minimization”, submitted to JMLR.

- **Jiayu Chen*,** Hanhan Zhou*, Yonsheng Mei, Jingdi Chen, Vaneet Aggarwal, Tian Lan, “Generative Models for Offline Policy Learning: A Survey”, in progress.

- …

Detailed Introductions

- **(Scalable) Multi-agent Skill Discovery:**

- **Objective:** In Laplacian-based skill discovery, skills are built to improve the connectedness of the state space and thus shorten the exploration time of the agent. In particular, the Fiedler vector is estimated for the state transition graph, which provides an embedding of the distance among states. Then, the most distant states indicated by the Fiedler vector are connected with skills. Our work firstly extends Laplacian-based skill discovery to multi-agent cooperative scenarios, to benefit the co-exploration and coordination among the agents.

- **Challenge:** Estimation of the Fiedler vector for the joint-state transition graph, of which the size grows exponentially with the number of agents.

- **Novelty:** We propose to model the joint-state transition graph as a Kronecker graph, which allows us to estimate the joint Fiedler vector based on the Laplacian spectrum of each individual state transition graph. We further propose an NN-based manner to estimate the Laplacian spectrum, so that our algorithm does not rely on eigen decompositions and can be applied to continuous robotic control tasks with large-scale state spaces.

Detailed Introductions

- **A Unified Framework for Unsupervised Skill Discovery:**

- **Objective:** Variational skill discovery algorithms learn skills through a mutual-information loss to improve the diversity among the skills. However, the coverage of each skill is usually overlooked. While, Laplacian-based skill discovery encourages the learned skill to traverse multiple state clusters to improve the connectedness. However, only one skill is discovered at a time, lacking diversity. We aim at integrating Laplacian-based and variational skill discovery to acquire skills with both superior diversity and coverage.

- **Challenge:** Explicit measure and optimization of the skill diversity and coverage in a unified framework.

- **Novelty:** We propose to adopt Determinantal Point Process (DPP) to model the (single-skill) coverage as the number of modes in its trajectory and (multi-skill) diversity as the number of modes in the set of all trajectories. Further, we integrate the DPP-based objectives with a mutual-information loss to establish the skill-policy mapping, which enables us to learn multiple skills at the sample time. In this way, we provide a unified framework of unsupervised skill discovery for both coverage and diversity.

Detailed Introductions

- **Hierarchical Adversarial Inverse Reinforcement learning:**

- **Objective/Challenge:** Adversarial Inverse Reinforcement learning (AIRL) is a state-of-the-art (SOTA) IL algorithm. We aim to propose a Hierarchical IL algorithm built on AIRL, while solving the disadvantages of pre-proposed Hierarchical IL algorithms, such as failing to enhance the casual relationship between the trajectory and the skill sequence and failing to optimize the high-level and low-level policy simultaneously.

- **Novelty:** We adopt the one-step option framework and redefine the AIRL objectives in an extended state-action space with the skill variables. To avoid degeneration and enhance the casual relationship between the skill variables and induced trajectories, we further introduce objective terms based on directed information as regularizers. To fit our algorithm to trajectories without skill annotations, we also provide an EM-adaption of our algorithm.

Detailed Introductions

- **Multi-task Hierarchical AIRL:**

- **Objective/Challenge:** Given a mixture of expert demonstrations from multiple tasks but without any (task or skill) annotation, we aim to recover a multi-task hierarchical policy that can be directly applied on a class of tasks, which is a pioneering research work in this area and a step towards general-purpose robots.

- **Novelty:** Based on Hierarchical AIRL, we further introduce a task variable and establish the task-policy mapping by maximizing the mutual information between the task variable and its corresponding trajectories. Given multiple components in the overall framework, we formally propose a hierarchical learning framework to update them coherently. The learned hierarchical policy shows superior performance on challenging, long-horizon tasks. Also, the low-level components of the hierarchical policy constitutes transferrable skills for even tasks out of the class.

Detailed Introductions

- **Deepfreight:**

- **Objective/Challenge:** A scalable, efficient, and reliable freight delivery scheduling system.

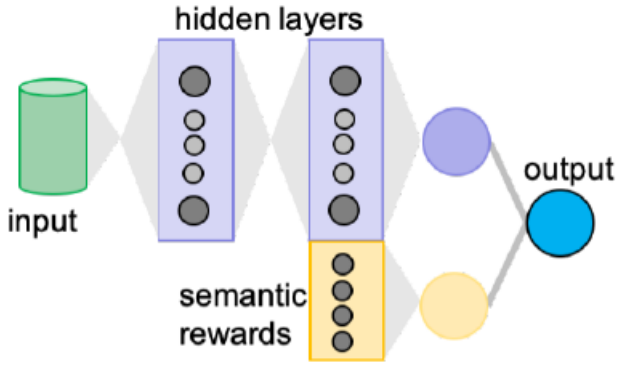
- **Novelty:** We propose to solve the freight delivery problem as two sub-problems: truck dispatch and package matching. The truck dispatch decisions are given by agents trained with multi-agent reinforcement learning (MARL). We adopt the centralized training and decentralized execution (CTDE) scheme, so the coordination among agents are taken into consideration when training, while joint dispatch decisions are decomposed to each individual truck when utilizing the learned policy, so the decision making is highly scalable. For package matching, we adopt a fast greedy heuristics based on the graph search, which allows multiple transfer of each package for efficiency. Further, we propose to model this problem as a mixed integer programming (MIP) problem and adopt an MIP solver in conjunction with MARL to filter out inefficient dispatch decisions and realize 100% delivery success, making the overall system reliable.

Detailed Introductions

- **Decision Making for Autonomous Driving via augmented AIRL:**

- **Objective/Challenge:** Learning human-like lane-changing behaviors with an emphasis on the efficiency, safety, and comfort.

- **Novelty:** Besides expert demonstrations, we introduce extra supervision on the objectives mentioned above, through introducing semantic rewards. While the terms are defined by users, the weight for each term is learned as part of the neural network in an augmented AIRL framework.



Detailed Introductions

- **Hierarchical Deep Counterfactual Regret Minimization:**

- **Objective:** Professionals in a field typically possess robust domain-specific skills, which they can employ to compose comprehensive strategies for tackling diverse and intricate task scenarios. In this work, we explore how to enable AI agents to do decision-making in imperfect information games (e.g., play pokers) while forming skills or with predefined skills.

- **Challenge:** Since there are no previous works on this direction, theoretical foundations need to be laid and then deep learning extensions should be developed for scalable applications.

- **Novelty:** Built on one of the most successful family of algorithms in imperfect information games – CFR, we propose HDCFR which is developed with gradual extensions. In particular, we extend the definitions of extensive-form games with the option framework, propose hierarchical CFR updating rules with convergence guarantee, employ a low-variance Monte Carlo sampling for model-free learning, and finally incorporate deep learning techniques for scalability.

Future Research Directions

Please view of my research statement available at: **https://lucascjysdl.github.io/files/Research_Statement.pdf**

Thank you!

I appreciate the provided opportunity to show my research work.

Jiayu Chen

chen3686@purdue.edu

lucascjysdl.github.io