Part A

Josyula, Gopala Krishna - 934375321

September 2022

1 Reinforcement Learning

1. CF Hayes, M Reymond, DM Roijers, E Howley, Distributional monte carlo tree search for risk-aware and multi-objective reinforcement learning, AAMAS20, Citations:12
2. TD Simão, N Jansen, MTJ Spaan, AlwaysSafe: Reinforcement learning without safety constraint violations during training, AAMAS21, Citations: 15
3. N Anastassacos, J García, S Hailes, Cooperation and reputation dynamics with reinforcement learning, AAMAS21, Citations: 5
4. H Xu, R Wang, L Raizman, Z Rabinovich, Transferable environment poisoning: Training-time attack on reinforcement learning, AAMAS21, Citations:10
5. K Chen, S Guo, T Zhang, S Li, Temporal watermarks for deep reinforcement learning models, AAMAS21, Citations:11

2 Neural Networks & Deep Learning

1. FT Johora, H Cheng, JP Müller, M Sester, An agent-based model for trajectory modelling in shared spaces: a combination of expert-based and deep learning approaches, AAMAS20, Citation: 11
2. Fredrik Präntare, Herman Appelgren, Mattias Tiger, David Bergström and Fredrik Heintz, Learning Heuristics for Combinatorial Assignment by Optimally Solving Subproblems, Citations: 0
3. Naman Shah and Siddharth Srivastava, Using Deep Learning to Bootstrap Abstractions for Hierarchical Robot Planning, AAMAS22, Citations: 4
4. Luca Viano, Yu-Ting Huang, Parameswaran Kamalaruban, Craig Innes, Subramanian Ramamoorthy and Adrian Weller, Robust Learning from Observation with Model Misspecification, AAMAS22, Citations: 1
5. Mathieu Reymond, Eugenio Bargiacchi, Ann Nowé,Pareto Conditioned Networks, AAMAS22, Citations: 1

3 Reward Shaping

1. H Zou, T Ren, D Yan, H Su, J Zhu, Learning task-distribution reward shaping with meta-learning, AAMAS21, Citations: 6

2. B Xiao, Q Lu, B Ramasubramanian, A Clark, Fresh: Interactive reward shaping in high-dimensional state spaces using human feedback, AAMAS20, Citations: 14
3. RT Icarte, TQ Klassen, R Valenzano, Reward machines: Exploiting reward function structure in reinforcement learning, JAIR22, Citations: 38
4. Neary, C., Xu, Z., Wu, B., Topcu, U., Reward machines for cooperative multiagent reinforcement learning, AAMAS21, Citations: 17
5. Baicen Xiao, Bhaskar Ramasubramanian and Radha Poovendran, Agent-Temporal Attention for Reward Redistribution in Episodic Multi-Agent Reinforcement Learning, AAMAS22, Citations: 0

4 Potential Functions

1. P Atreya, J Biswas , State Supervised Steering Function for Sampling-based Kinodynamic Planning, AAMAS22, Citations:0
2. N Zerbel, K Tumer, The power of suggestion, AAMAS22, Citations: 1
3. G Rockefeller, S Khadka, K Tumer, Multi-level fitness critics for cooperative coevolution, AAMAS21, Citations: 9
4. J Burden, SK Siahroudi, Latent Property State Abstraction For Reinforcement learning, AAMAS21, Citations: 21
5. WF Sun, CK Lee, CY Lee, A Distributional Perspective on Value Function Factorization Methods for Multi-Agent Reinforcement Learning, AAAMAS21, Citations: 2

5 Co-evolutionary Learning

I could not find sufficient papers on this topic published within the past two years, I include the most cited papers I have found

1. Nicholas Zerbel, Kagan Tumer, The Power of Suggestion, AAMAS20, Citations: 1
2. JJ Chung, D Miklić, L Sabattini, The impact of agent definitions and interactions on multiagent learning for coordination, AAMAS19, Citations: 7
3. Bara, J., Turrini, P. and Andrighetto, G., . Enabling imitation based cooperation in dynamic social networks. AAMAS22, Citations:0
4. Chung, J.J., Miklić, D., Sabattini, L., Tumer, K. and Siegwart, R., The impact of agent definitions and interactions on multiagent learning for coordination in traffic management domains. AAMAS20 Citations: 6.
5. Jorge Carvalho Gomes, Pedro Mariano, Anders Lyhne Christensen, Cooperative Coevolution of Partially Heterogeneous Multiagent Systems, AAMAS15, Citations: 22

6 Teaming or Team Formation

1. Aaquib Tabrez, Matthew B. Luebbers and Bradley Hayes, Descriptive and Prescriptive Visual Guidance to Improve Shared Situational Awareness in Human-Robot Teaming, AAMAS22, Citations:2
2. Esmaeil Seraj, Zheyuan Wang, Rohan Paleja, Daniel Martin, Matthew Sklar, Anirudh Patel and Matthew Gombolay, Learning Efficient Diverse Communication for Cooperative Heterogeneous Teaming, AAMAS22, Citations:3
3. Athina Georgara, Juan Antonio Rodriguez Aguilar and Carles Sierra,Building contrastive explanations for multi-agent team formation, AAMAS22, Citations: 0
4. Yutong Wang and Guillaume Sartoretti, FCMNet: Full Communication Memory Net for Team-Level Cooperation in Multi-Agent Systems, AAMAS22, Citations: 0
5. Esmaeil Seraj, Zheyuan Wang, Rohan Paleja, Daniel Martin, Matthew Sklar, Anirudh Patel, and Matthew Gombolay., Learning Efficient Diverse Communication for Cooperative Heterogeneous Teaming, AAMAS20, Citations: 3

7 Game Theory

1. Julien Pérolat, Sarah Perrin, Romuald Elie, Mathieu Laurière, Georgios Piliouras, Matthieu Geist, Karl Tuyls and Olivier Pietquin, Scaling Mean Field Games with Online Mirror Descent, AAMAS22, Citations:19
2. Gianlorenzo D'Angelo, Esmaeil Delfaraz and Hugo Gilbert, Computation and Bribery of Voting Power in Delegative Simple Games, AAMAS22, Citations: 1
3. Denizalp Goktas, Jiayi Zhao and Amy Greenwald, Robust No-Regret Learning in Min-Max Stackelberg Games, AAMAS22, Citations: 2
4. Kaisheng Wu, Liangda Fang, Liping Xiong, Zhao-Rong Lai, Yong Qiao, Kaidong Chen, Fei Rong, Automatic Synthesis of Generalized Winning Strategy of Impartial Combinatorial Games, AAMAS20, Citations: 2
5. Ian Gemp, Rahul Savani, Marc Lanctot, Yoram Bachrach, Thomas Anthony, Richard Everett, Andrea Tacchetti, Tom Eccles and Janos Kramar, Sample-based Approximation of Nash in Large Many-Player Games via Gradient Descent, AAMAS22, Citations: 3

8 Computational Social Choice

1. Liangde Tao, Lin Chen, Lei Xu, Weidong Shi, Ahmed Sunny and Md Mahabub Uz Zaman, How Hard is Bribery in Elections with Randomly Selected Voters, AAMAS22, Citations:0
2. S Olafsson, BC Wallace, TW Bickmore, Towards a Computational Framework for Automating Substance Use Counseling with Virtual Agents. AAMAS20, Citations 14
3. Oliviero Nardi, Arthur Boixel and Ulle Endriss, A Graph-Based Algorithm for the Automated Justification of Collective Decisions, AAMAS22, Citations: 5
4. Roberto Lucchetti, Stefano Moretti and Tommaso Rea, Coalition Formation Games and Social Ranking Solutions, AAMAS22, Citations: 0
5. Felix Brandt, Martin Bullinger, Patrick Lederer, On the Indecisiveness of Kelly-Strategyproof Social Choice Functions, AAMAS21, Citations: 5

9 Mechanism Design, Auctions

1. Aravind Srinivasan and Pan Xu, The Generalized Magician Problem under Unknown Distributions and Related Applications, Citations:0
2. Dengji Zhao, Mechanism Design Powered by Social Interactions, AAMAS21, Citations: 8
3. Rahul Chandan, Dario Paccagnan, Jason R. Marden, Tractable mechanisms for computing near-optimal utility functions, AAMAS21, Citations: 4
4. Alina Filimonov, Reshef Meir, Strategyproof Facility Location Mechanisms on Discrete Trees, AAMAS21, Citations: 2
5. Jiarui Gan , Edith Elkind , Sarit Kraus , Michael Wooldridge, Mechanism Design for Defense Coordination in Security Games, Citations: 6

10 Multi-Robot Coordination

1. Mirko Salaris, Alessandro Riva , Francesco Amigoni, Politecnico di Milano, Multirobot Coverage of Modular Environments, AAMAS20, Citations: 2
2. Haris Aziz, Hau Chan, Ágnes Cseh, Bo Li, Fahimeh Ramezani, Chenhao Wang, Multi-Robot Task Allocation-Complexity and Approximation, Citations: 8
3. Jingqing Ruan et.al, GCS: Graph-Based Coordination Strategy for Multi-Agent Reinforcement Learning,AAMAS22 Citations: 4
4. Charlie Street, Bruno Lacerda, Manuel Mühlig, Nick Hawes, Multi-Robot Planning Under Uncertainty with Congestion-Aware Models, AAMAS20, Citations: 15
5. Guohui Ding et.al, Distributed Reinforcement Learning for Cooperative Multi-Robot Object Manipulation, AAMAS20 Citations: 16

Paper Reviews Multiagent Systems

Josyula, Gopala Krishna - 934375321

September 2022

1 Topic Neural Networks or Deep Learning

1.1 Graph Neural Networks for Decentralized Multi-Robot Path Planning, Qingbiao Li1, Fernando Gama, Alejandro Ribeiro, Amanda Prorok, AAMAS 2020, Citations AAMAS:3 IROS:77

The paper in discussion is [1]

- Big Picture: This paper deals with decentralized multirobot path planning with local communication, they outline the information that is crucial to communicate between the robots from local observations to navigate in a constrained workspace successfully. They solve this problem by introducing feature extraction from CNN[1] and a GNN[2] for communication modeling and an MLP[3] which is a decision maker of motion planning commands.
- Why Care?: Navigating a team of robots has been a problem of historical importance in robot control and navigation due to its prevalence in many target industries like construction, mapping, exploration, and delivery etc interest and holds a major share of today's market.
- What is hard or challenging about solving this problem?: Solving for conflict-free and optimal paths in a multi-robot system is considered NP-hard [1], although significant progress has been made towards alleviating the computational load these approaches still scale poorly in environments with a high number of potential path conflicts.
- Current Approaches: Optimization based methods can generally be solved without the use of communication and taking purely local objectives into account, where global objectives cannot be explicitly optimized [2]. Learning-based approaches perform better than classical optimization-based approaches like ORCA[3] and handle high dimensional problems however it is not a decentralized approach and cannot exploit the local communications well.
- Gaps : Optimization based approaches were considering only local objectives which do not address global optimality, cannot scale to high dimensional problems and learning based approaches were not decentralized and both methods do not exploit local communication model.
- Addressing the GAP: Authors use a GNN to address the communication model. The GNN communicates only with nearest neighbors for decision making, the features for GNN are extracted from a CNN, which do not have a global reference frame and only consider local, relative information about their positions and goals.

[1] Convolutional Neural Networks
[2] Graph Neural Networks
[3] Multilayer Perceptron

- Cool/Notable things: Their results show performance close to that of expert algorithm like discrete ORCA and also demonstrate their results on unseen data with larger groups of robots with a success rate of 95.8% for 14 agents and 99% for 4 agents in random configurations.
- Outcome of this work: Authors demonstrate collision avoidance on multirobot systems with time delayed communication, i.e. local messages do not arrive instantly and also demonstrate results of position swapping, braking dead locks between robots with 100% success rate for fixed configurations.

2 Topic: Multi-Robot Coordination

2.1 GCS: Graph-Based Coordination Strategy for Multi-Agent Reinforcement Learning, Jingqing Ruan et.al AAMAS22 Citations: 4

This section discuss the paper [4]

- Big Picture: Many real-world scenarios involve a team of agents coordinating their policies to achieve a shared goal. This paper proposes an approach to achieve behavior/action coordination among multiagent systems through graphs.
- Why Care? : This work has many real-life applications such as traffic control, multiplayer games, and social dilemma. Poorly coordinated agents can cause lot of issues, Taking traffic flow as an example, when multiple vehicles are trying to cross an intersection without traffic lights, most likely, the traffic will become congested if all vehicles take actions simultaneously without a rational sequence.
- What is hard or challenging about solving this problem? : Knowing the intent of other agents and changing the behavior of agents according to the globally shared objectives.
- Current Approaches: Most current approaches treat the entire MAS[4] as a single agent and optimize a centralized objective for all the agents. BiAC[5] approach mainly focuses on coordination of the asynchronous decisions of two agents. The multi-agent rollout algorithm[6] provides a theoretical view of executing a local rollout with some coordinating information but is limited to an agent-by-agent decision dependency structure.
- Gaps : Previous works were mostly centralized. Although works like [5] and [6] investigate the execution order of two agent and multiagent systems, they don't capture the underlying dynamic decision dependency structure of a MAS.
- Addressing the GAP: Authors propose a graph-based coordination strategy(GCS) that learns coordinated behaviors through factorizing joint team policy into a graph generator and a graph-based coordinated policy. A graph generator aims to learn an action coordination graph (ACG) that properly represents the decision dependency. GCS coordinates the dependent behaviors among agents exploiting the underlying decision dependency. This allows for a decentralized and action-coordinated strategy
- Cool/Notable things: Apart from not being centralized and extracting coordinated action strategies, the authors demonstrate their work on the Google Research Football RL environment [7]
- Outcome of this work: A graph-based multi-robot coordination strategy where the joint team policy is factored into a graph generator and which outputs DAGs[5]to capture the underlying decision structure and a graph coordinated policy controller, with experiments and results with demonstrated performance increase in Google Research Football and Cooperative Navigation environments compared with baselines.

[4]Multiagent System
[5]Directed Acyclic Graphs

3 Topic: Reinforcement Learning

3.1 Distributed Reinforcement Learning for Cooperative Multi-Robot Object Manipulation, Guohui Ding et.al. AAMAS20 Citations: 16

This section discuss the paper [8]

- Big Picture: The paper studies two approaches to distributed reinforcement learning, a game theoretic approach and a distributed approximate RL for training robot arms to manipulate an object cooperatively
- Why Care?: Reinforcement learning is a successful paradigm applied to many control problems however, distributed and cooperative approaches in RL are extremely challenging to solve and require scalable solutions.
- What is hard or challenging about solving this problem?: The robots may have heterogeneous physical constraints and possibly partial or asymmetric observations of different robots. Parallel implementations of single agent RL suffer learning instabilities.
- Current Approaches: Distributed approximate RL applies single agent q-learning to individual agents but suffers from no convergence. Authors, therefore, use Game theoretic RL which ensures to achieve Nash-equilibia of the agents[9]
- Gaps : In Distributed RL the reward is given a sum of individual agent rewards which cannot be accountable for single agent optimal policy.
- Addressing the GAP: Authors propose game theoretic RL which models the Q-value function updates in such a way that a nash equilibrium is achieved, this approach is called Nash-Q-Learning.
- Cool/Notable things: Application of game theoretic RL models to object manipulation with asymmetric observations and heterogeneous capabilities, demonstration of the results in a classic bi-robot object manipulation task.
- Outcome of this work: A comparison of distributed approximate RL approaches and the Nash Q-learning authors demonstrate that Game Theoritic-RL is sensitive to reward structure while DA-RL is robust to reward structure with this reward structure authors also conclude that GT-RL performs better than DA-RL when an appropriate reward structure is chosen they show that there is more than 80% success rate for object manipulation task with GT-RL with experimental evidence.

References

[1] Q. Li, F. Gama, A. Ribeiro, and A. Prorok, "Graph neural networks for decentralized path planning," in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, ser. AAMAS '20. Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems, 2020, p. 1901–1903.

[2] M. Rotkowitz and S. Lall, "A characterization of convex problems in decentralized control," *IEEE Transactions on Automatic Control*, vol. 50, no. 12, pp. 1984–1996, 2005.

[3] A. Rajeswaran, K. Lowrey, E. Todorov, and S. M. Kakade, "Towards generalization and simplicity in continuous control," *CoRR*, vol. abs/1703.02660, 2017. [Online]. Available: http://arxiv.org/abs/1703.02660

[4] J. Ruan, Y. Du, X. Xiong, D. Xing, X. Li, L. Meng, H. Zhang, J. Wang, and B. Xu, "Gcs: Graph-based coordination strategy for multi-agent reinforcement learning," *arXiv preprint arXiv:2201.06257*, 2022.

[5] H. Zhang, W. Chen, Z. Huang, M. Li, Y. Yang, W. Zhang, and J. Wang, "Bi-level actor-critic for multi-agent coordination," 2019. [Online]. Available: https://arxiv.org/abs/1909.03510

[6] D. P. Bertsekas, "Multiagent rollout algorithms and reinforcement learning," *CoRR*, vol. abs/1910.00120, 2019. [Online]. Available: http://arxiv.org/abs/1910.00120

[7] Google, "Google research football," 2019. [Online]. Available: https://github.com/google-research/football

[8] G. Ding, J. J. Koh, K. Merckaert, B. Vanderborght, M. M. Nicotra, C. Heckman, A. Roncone, and L. Chen, "Distributed reinforcement learning for cooperative multi-robot object manipulation," *CoRR*, vol. abs/2003.09540, 2020. [Online]. Available: https://arxiv.org/abs/2003.09540

[9] J. Hu and M. P. Wellman, "Nash q-learning for general-sum stochastic games," *J. Mach. Learn. Res.*, vol. 4, no. null, p. 1039–1069, dec 2003.