

ACKNOWLEDGMENTS

This work is supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDA27030300) and the Program for National Nature Science Foundation of China (62073324).

REFERENCES

- [1] Rishabh Agarwal, Marlos C Machado, Pablo Samuel Castro, and Marc G Belle-mare. 2021. Contrastive behavioral similarity embeddings for generalization in reinforcement learning. *arXiv preprint arXiv:2101.05265* (2021).
- [2] Haitham Bou Ammar, Eric Eaton, Paul Ruvolo, and Matthew Taylor. 2014. Online multi-task learning for policy gradient methods. In *International conference on machine learning*. PMLR, 1206–1214.
- [3] Jacob Andreas, Dan Klein, and Sergey Levine. 2017. Modular multitask reinforcement learning with policy sketches. In *International Conference on Machine Learning*. PMLR, 166–175.
- [4] Alexei Baevski, Henry Zhou, Abdelrahman Mohamed, and Michael Auli. 2020. wav2vec 2.0: A framework for self-supervised learning of speech representations. *arXiv preprint arXiv:2006.11477* (2020).
- [5] Diana Borsa, Thore Graepel, and John Shawe-Taylor. 2016. Learning shared representations in multi-task reinforcement learning. *arXiv preprint arXiv:1603.02041* (2016).
- [6] Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. 2021. Decision transformer: Reinforcement learning via sequence modeling. *arXiv preprint arXiv:2106.01345* (2021).
- [7] Li Chenghao, Tonghan Wang, Chengjie Wu, Qianchuan Zhao, Jun Yang, and Chongjie Zhang. 2021. Celebrating diversity in shared multi-agent reinforcement learning. *Advances in Neural Information Processing Systems* 34 (2021), 3991–4002.
- [8] Michael Crawshaw. 2020. Multi-task learning with deep neural networks: A survey. *arXiv preprint arXiv:2009.09796* (2020).
- [9] Jingjing Cui, Yuanwei Liu, and Arumugam Nallanathan. 2019. Multi-agent reinforcement learning-based resource allocation for UAV networks. *IEEE Transactions on Wireless Communications* 19, 2 (2019), 729–743.
- [10] Carlo D’Eramo, Davide Tateo, Andrea Bonarini, Marcello Restelli, and Jan Peters. 2019. Sharing knowledge in multi-task deep reinforcement learning. In *International Conference on Learning Representations*.
- [11] Coline Devin, Abhishek Gupta, Trevor Darrell, Pieter Abbeel, and Sergey Levine. 2017. Learning modular neural network policies for multi-task and multi-robot transfer. In *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2169–2176.
- [12] Wei Du and Shifei Ding. 2020. A survey on multi-agent deep reinforcement learning: from the perspective of challenges and applications. *Artificial Intelligence Review* (2020).
- [13] Justin Fu, Aviral Kumar, Ofir Nachum, George Tucker, and Sergey Levine. 2020. D4rl: Datasets for deep data-driven reinforcement learning. *arXiv preprint arXiv:2004.07219* (2020).
- [14] Scott Fujimoto, David Meger, and Doina Precup. 2019. Off-policy deep reinforcement learning without exploration. In *International Conference on Machine Learning*. PMLR, 2052–2062.
- [15] Aditya Grover, Maruan Al-Shedivat, Jayesh Gupta, Yuri Burda, and Harrison Edwards. 2018. Learning policy representations in multiagent systems. In *International conference on machine learning*. PMLR, 1802–1811.
- [16] Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. 2019. Dream to control: Learning behaviors by latent imagination. *arXiv preprint arXiv:1912.01603* (2019).
- [17] Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, and Jimmy Ba. 2020. Mastering atari with discrete world models. *arXiv preprint arXiv:2010.02193* (2020).
- [18] Assaf Hallak, Dotan Di Castro, and Shie Mannor. 2015. Contextual markov decision processes. *arXiv preprint arXiv:1502.02259* (2015).
- [19] Siyi Hu, Fengda Zhu, Xiaojun Chang, and Xiaodan Liang. 2021. Updet: Universal multi-agent reinforcement learning via policy decoupling with transformers. *arXiv preprint arXiv:2101.08001* (2021).
- [20] Rahul Jain, Preeti Ranjan Panda, and Sreenivas Subramoney. 2017. Cooperative multi-agent reinforcement learning-based co-optimization of cores, caches, and on-chip network. *ACM Transactions on Architecture and Code Optimization (TACO)* 14, 4 (2017), 1–25.
- [21] Michael Janner, Qiyang Li, and Sergey Levine. 2021. Reinforcement Learning as One Big Sequence Modeling Problem. *arXiv preprint arXiv:2106.02039* (2021).
- [22] Aviral Kumar, Justin Fu, Matthew Soh, George Tucker, and Sergey Levine. 2019. Stabilizing Off-Policy Q-Learning via Bootstrapping Error Reduction. *Advances in Neural Information Processing Systems* 32 (2019), 11784–11794.
- [23] Aviral Kumar, Justin Fu, Matthew Soh, George Tucker, and Sergey Levine. 2019. Stabilizing off-policy q-learning via bootstrapping error reduction. *Advances in Neural Information Processing Systems* 32 (2019).
- [24] Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. 2020. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643* (2020).
- [25] Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. 2020. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643* (2020).
- [26] Linghui Meng, Muning Wen, Yaodong Yang, Chenyang Le, Xiyun Li, Weinan Zhang, Ying Wen, Haifeng Zhang, Jun Wang, and Bo Xu. 2021. Offline Pre-trained Multi-Agent Decision Transformer: One Big Sequence Model Conquers All StarCraftII Tasks. *arXiv preprint arXiv:2112.02845* (2021).
- [27] Ashvin Nair, Murtaza Dalal, Abhishek Gupta, and Sergey Levine. 2020. Accelerating online reinforcement learning with offline datasets. *arXiv preprint arXiv:2006.09359* (2020).
- [28] Junhyuk Oh, Satinder Singh, Honglak Lee, and Pushmeet Kohli. 2017. Zero-shot task generalization with multi-task deep reinforcement learning. In *International Conference on Machine Learning*. PMLR, 2661–2670.
- [29] Frans A Oliehoek and Christopher Amato. 2016. *A concise introduction to decentralized POMDPs*. Springer.
- [30] Shayegan Omidshafiei, Jason Papis, Christopher Amato, Jonathan P How, and John Vian. 2017. Deep decentralized multi-task multi-agent reinforcement learning under partial observability. In *International Conference on Machine Learning*. PMLR, 2681–2690.
- [31] Aaron Oord, Oriol Vinyals, and Koray Kavukcuoglu. 2019. Neural Discrete Representation Learning. (2019).
- [32] Sheril Ozair, Yazhe Li, Ali Razavi, Ioannis Antonoglou, Aaron van den Oord, and Oriol Vinyals. 2021. Vector Quantized Models for Planning. *arXiv: Learning* (2021).
- [33] Adolfo Perrusquia, Wen Yu, and Xiaoou Li. 2021. Multi-agent reinforcement learning for redundant robot control in task-space. *International Journal of Machine Learning and Cybernetics* 12, 1 (2021), 231–241.
- [34] Roxana Rădulescu, Patrick Mannion, Diederik M Roijers, and Ann Nowé. 2020. Multi-objective multi-agent decision making: a utility-based analysis and survey. *Autonomous Agents and Multi-Agent Systems* 34, 1 (2020), 1–52.
- [35] Ali Razavi, Aaron van den Oord, and Oriol Vinyals. 2019. Generating diverse high-fidelity images with vq-vae-2. In *Advances in neural information processing systems*. 14866–14876.
- [36] Mikayel Samvelyan, Tabish Rashid, Christian Schroeder De Witt, Gregory Farquhar, Nantas Nardelli, Tim GJ Rudner, Chia-Man Hung, Philip HS Torr, Jakob Foerster, and Shimon Whiteson. 2019. The starcraft multi-agent challenge. *arXiv preprint arXiv:1902.04043* (2019).
- [37] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- [38] Shagun Sodhani, Franziska Meier, Joelle Pineau, and Amy Zhang. 2021. Block Contextual MDPs for Continual Learning. *arXiv preprint arXiv:2110.06972* (2021).
- [39] Shagun Sodhani, Amy Zhang, and Joelle Pineau. 2021. Multi-Task Reinforcement Learning with Context-based Representations. *arXiv preprint arXiv:2102.06177* (2021).
- [40] Aravind Srinivas, Michael Laskin, and Pieter Abbeel. 2020. Curl: Contrastive unsupervised representations for reinforcement learning. *arXiv preprint arXiv:2004.04136* (2020).
- [41] Adam Stooke, Kimin Lee, Pieter Abbeel, and Michael Laskin. 2021. Decoupling representation learning from reinforcement learning. In *International Conference on Machine Learning*. PMLR, 9870–9879.
- [42] Fumihide Tanaka and Masayuki Yamamura. 2003. Multitask reinforcement learning on the distribution of MDPs. In *Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Vol. 3. IEEE, 1108–1113.
- [43] Wei-Cheng Tseng, Tsun-Hsuan Wang, Yen-Chen Lin, and Phillip Isola. 2022. Offline Multi-Agent Reinforcement Learning with Knowledge Distillation. In *Thirty-Sixth Conference on Neural Information Processing Systems*. <https://openreview.net/forum?id=yipUuqxveCy>
- [44] Wei-Cheng Tseng, Wei Wei, Da-Chen Juan, and Min Sun. 2021. Meta-cpr: Generalize to unseen large number of agents with communication pattern recognition module. *arXiv preprint arXiv:2112.07222* (2021).
- [45] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *arXiv preprint arXiv:1706.03762* (2017).
- [46] Huan Wang, Stephan Zheng, Caiming Xiong, and Richard Socher. 2019. On the generalization gap in reparameterizable reinforcement learning. In *International Conference on Machine Learning*. PMLR, 6648–6658.
- [47] Tonghan Wang, Heng Dong, Victor Lesser, and Chongjie Zhang. 2020. ROMA: Multi-Agent Reinforcement Learning with Emergent Roles. In *International Conference on Machine Learning*. PMLR, 9876–9886.
- [48] Tonghan Wang, Tarun Gupta, Anuj Mahajan, Bei Peng, Shimon Whiteson, and Chongjie Zhang. 2020. RODE: Learning Roles to Decompose Multi-Agent Tasks. In *International Conference on Learning Representations*.
- [49] Mengdi Xu, Yikang Shen, Shun Zhang, Yuchen Lu, Ding Zhao, Joshua Tenenbaum, and Chuang Gan. 2022. Prompting decision transformer for few-shot policy

- generalization. In *International Conference on Machine Learning*. PMLR, 24631–24645.
- [50] Ruihan Yang, Huazhe Xu, YI WU, and Xiaolong Wang. 2020. Multi-Task Reinforcement Learning with Soft Modularization. *Advances in Neural Information Processing Systems* 33 (2020), 4767–4777.
- [51] Yiqin Yang, Xiaoteng Ma, Li Chenghao, Zewu Zheng, Qiyuan Zhang, Gao Huang, Jun Yang, and Qianchuan Zhao. 2021. Believe what you see: Implicit constraint approach for offline multi-agent reinforcement learning. *Advances in Neural Information Processing Systems* 34 (2021).
- [52] Chao Yu, Akash Velu, Eugene Vinitzky, Yu Wang, Alexandre Bayen, and Yi Wu. 2021. The surprising effectiveness of mappo in cooperative, multi-agent games. *arXiv preprint arXiv:2103.01955* (2021).
- [53] Tianhe Yu, Aviral Kumar, Yevgen Chebotar, Karol Hausman, Sergey Levine, and Chelsea Finn. 2021. Conservative data sharing for multi-task offline reinforcement learning. *Advances in Neural Information Processing Systems* 34 (2021).
- [54] Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. 2020. Gradient Surgery for Multi-Task Learning. *Advances in Neural Information Processing Systems* 33 (2020).
- [55] Tianhe Yu, Deirdre Quillen, Zhanpeng He, Ryan Julian, Karol Hausman, Chelsea Finn, and Sergey Levine. 2020. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on Robot Learning*. PMLR, 1094–1100.
- [56] Amy Zhang, Rowan McAllister, Roberto Calandra, Yarin Gal, and Sergey Levine. 2020. Learning invariant representations for reinforcement learning without reconstruction. *arXiv preprint arXiv:2006.10742* (2020).
- [57] Tony Z Zhao, Jianlan Luo, Oleg Sushkov, Rugile Pevceviute, Nicolas Heess, Jon Scholz, Stefan Schaal, and Sergey Levine. 2021. Offline meta-reinforcement learning for industrial insertion. *arXiv preprint arXiv:2110.04276* (2021).
- [58] Yi Zhao, Haoyu Li, Cheng-I Lai, Jennifer Williams, Erica Cooper, and Junichi Yamagishi. 2020. Improved prosody from learned f0 codebook representations for vq-vae speech waveform reconstruction. *arXiv preprint arXiv:2005.07884* (2020).
- [59] Ming Zhou, Jun Luo, Julian Villella, Yaodong Yang, David Rusu, Jiayu Miao, Weinan Zhang, Montgomery Alban, Iman Fadakar, Zheng Chen, et al. 2020. Smarts: Scalable multi-agent reinforcement learning training school for autonomous driving. *arXiv preprint arXiv:2010.09776* (2020).