

- [3] Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemyslaw Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, et al. 2019. Dota 2 with large scale deep reinforcement learning. *arXiv preprint arXiv:1912.06680* (2019).
- [4] Noam Brown and Tuomas Sandholm. 2019. Superhuman AI for multiplayer poker. *Science* 365, 6456 (2019), 885–890.
- [5] Kalesha Bullard, Douwe Kiela, Franziska Meier, Joelle Pineau, and Jakob Foerster. 2021. Quasi-Equivalence Discovery for Zero-Shot Emergent Communication. (March 2021). [arXiv:2103.08067 \[cs.MA\]](https://arxiv.org/abs/2103.08067)
- [6] Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. 2002. Deep blue. *Artificial intelligence* 134, 1-2 (2002), 57–83.
- [7] Rodrigo Canaan, Xianbo Gao, Julian Togelius, Andy Nealen, and Stefan Menzel. 2020. Generating and adapting to diverse ad-hoc cooperation agents in hanabi. *arXiv preprint arXiv:2004.13710* (2020).
- [8] Rodrigo Canaan, Julian Togelius, Andy Nealen, and Stefan Menzel. 2019. Diverse agents for ad-hoc cooperation in hanabi. In *2019 IEEE Conference on Games (CoG)*. IEEE, 1–8.
- [9] Micah Carroll, Rohin Shah, Mark K Ho, Tom Griffiths, Sanjit Seshia, Pieter Abbeel, and Anca Dragan. 2019. On the utility of learning about humans for human-ai coordination. *Advances in Neural Information Processing Systems* 32 (2019), 5174–5185.
- [10] Benjamin Eysenbach, Abhishek Gupta, Julian Ibarz, and Sergey Levine. 2019. Diversity is All You Need: Learning Skills without a Reward Function. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=Sjx63jRqFm>
- [11] Jakob Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. 2018. Counterfactual multi-agent policy gradients. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.
- [12] Jakob N Foerster, Francis Song, Edward Hughes, Neil Burch, Iain Dunning, Shimon Whiteson, Matthew Botvinick, and Michael Bowling. 2018. Bayesian Action Decoder for Deep Multi-Agent Reinforcement Learning. (Nov. 2018). [arXiv:1811.01458 \[cs.MA\]](https://arxiv.org/abs/1811.01458)
- [13] Jonathan Gray, Adam Lerer, Anton Bakhtin, and Noam Brown. 2020. Human-Level Performance in No-Press Diplomacy via Equilibrium Search. *arXiv preprint arXiv:2010.02923* (2020).
- [14] Jayesh K Gupta, Maxim Egorov, and Mykel Kochenderfer. 2017. Cooperative multi-agent control using deep reinforcement learning. In *International Conference on Autonomous Agents and Multiagent Systems*. Springer, 66–83.
- [15] Elad Hazan, Sham Kakade, Karan Singh, and Abby Van Soest. 2019. Provably efficient maximum entropy exploration. In *International Conference on Machine Learning*. PMLR, 2681–2691.
- [16] Hengyuan Hu and Jakob N Foerster. 2020. Simplified Action Decoder for Deep Multi-Agent Reinforcement Learning. In *International Conference on Learning Representations*. <https://openreview.net/forum?id=B1xm3RVtwB>
- [17] Hengyuan Hu, Adam Lerer, Brandon Cui, David Wu, Luis Pineda, Noam Brown, and Jakob Foerster. 2021. Off-Belief Learning. [arXiv:2103.04000 \[cs.AI\]](https://arxiv.org/abs/2103.04000)
- [18] Hengyuan Hu, Alexander Peysakhovich, Adam Lerer, and Jakob Foerster. 2020. “Other-Play” for Zero-Shot Coordination. In *Proceedings of Machine Learning and Systems 2020*. 9396–9407.
- [19] Mykel J Kochenderfer. 2015. *Decision making under uncertainty: theory and application*. MIT press.
- [20] Adam Lerer, Hengyuan Hu, Jakob Foerster, and Noam Brown. 2020. Improving policies via search in cooperative partially observable games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 7187–7194.
- [21] Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. *arXiv preprint arXiv:1706.02275* (2017).
- [22] Andrei Lupu, Brandon Cui, Hengyuan Hu, and Jakob Foerster. 2021. Trajectory diversity for zero-shot coordination. In *International Conference on Machine Learning*. PMLR, 7204–7213.
- [23] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *nature* 518, 7540 (2015), 529–533.
- [24] Shakir Mohamed and Danilo Jimenez Rezende. 2015. Variational information maximisation for intrinsically motivated reinforcement learning. *arXiv preprint arXiv:1509.08731* (2015).
- [25] Hirotaka Osawa. 2015. Solving hanabi: Estimating hands by opponent’s actions in cooperative game with incomplete information. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.
- [26] David Premack and Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1, 4 (1978), 515–526.
- [27] Neil Rabinowitz, Frank Perbet, Francis Song, Chiyuan Zhang, SM Ali Eslami, and Matthew Botvinick. 2018. Machine theory of mind. In *International conference on machine learning*. PMLR, 4218–4227.
- [28] Cinjon Resnick, Wes Eldridge, David Ha, Denny Britz, Jakob Foerster, Julian Togelius, Kyunghyun Cho, and Joan Bruna. 2018. Pommernan: A multi-agent playground. *arXiv preprint arXiv:1809.07124* (2018).
- [29] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, et al. 2020. Mastering atari, go, chess and shogi by planning with a learned model. *Nature* 588, 7839 (2020), 604–609.
- [30] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. Mastering the game of Go with deep neural networks and tree search. *nature* 529, 7587 (2016), 484–489.
- [31] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharrshan Kumaran, Thore Graepel, et al. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144.
- [32] Ho Chit Siu, Jaime D Pena, Kimberlee C Chang, Edenna Chen, Yutai Zhou, Victor J Lopez, Kyle Palko, and Ross E Allen. 2021. Evaluation of Human-AI Teams for Learned and Rule-Based Agents in Hanabi. *arXiv preprint arXiv:2107.07630* (2021).
- [33] DJ Strouse, Kevin McKee, Matt Botvinick, Edward Hughes, and Richard Everett. 2021. Collaborating with Humans without Human Data. *Advances in Neural Information Processing Systems* 34 (2021).
- [34] Justin K Terry, Nathaniel Grammel, Ananth Hari, and Luis Santos. 2020. Parameter sharing is surprisingly useful for multi-agent deep reinforcement learning. *arXiv e-prints* (2020), arXiv:2005.
- [35] Mark JH van den Bergh, Anne Hommelberg, Walter A Kusters, and Flora M Spiekma. 2016. Aspects of the cooperative card game hanabi. In *Benelux Conference on Artificial Intelligence*. Springer, 93–105.
- [36] Oriol Vinyals, Igor Babuschkin, Wojciech M Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H Choi, Richard Powell, Timo Ewalds, Petko Georgiev, et al. 2019. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature* 575, 7782 (2019), 350–354.
- [37] Joseph Walton-Rivers, Piers R Williams, and Richard Bartle. 2019. The 2018 hanabi competition. In *2019 IEEE Conference on Games (CoG)*. IEEE, 1–8.
- [38] Joseph Walton-Rivers, Piers R Williams, Richard Bartle, Diego Perez-Liebana, and Simon M Lucas. 2017. Evaluating and modelling hanabi-playing agents. In *2017 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, 1382–1389.
- [39] Rose E Wang, Sarah A Wu, James A Evans, Joshua B Tenenbaum, David C Parkes, and Max Kleiman-Weiner. 2020. Too many cooks: Coordinating multi-agent collaboration through inverse planning. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 2032–2034.
- [40] David H Wolpert and Kagan Tumer. 2002. Optimal payoff functions for members of collectives. In *Modeling complexity in economic and social systems*. World Scientific, 355–369.
- [41] Chih-Kuan Yeh, Cheng-Yu Hsieh, and Hsuan-Tien Lin. 2018. Automatic bridge bidding using deep reinforcement learning. *IEEE Transactions on Games* 10, 4 (2018), 365–377.