Emergent Dominance Hierarchies in Reinforcement Learning Agents*

Extended Abstract

Ram Rachum
Independent Researcher
ram@rachum.com

Yonatan Nakar
Tel Aviv University

Bill Tomlinson
University of California, Irvine

Nitay Alon
Max Planck Institute, Bio-Cybernetics
Hebrew University of Jerusalem

Reuth Mirsky
Bar-Ilan University
Israel

ABSTRACT

Modern Reinforcement Learning (RL) algorithms are able to outperform humans in a wide variety of tasks. Multi-agent reinforcement learning (MARL) settings present additional challenges around cooperation in mixed-motive groups. Social conventions and norms, often inspired by human institutions, are used as tools for striking the balance between individual and group objectives.

We examine a fundamental social convention that underlies cooperation in animal and human societies: dominance hierarchies.

We adapt the ethological theory of dominance hierarchies to artificial agents, borrowing established terminology and definitions. We provide an environment we call Chicken Coop, and we demonstrate that populations of RL agents in that environment can invent, learn, enforce, and transmit a dominance hierarchy to new populations. The dominance hierarchies that emerge in it have a similar structure to those studied in chickens, mice, fish, and other species.

KEYWORDS

Multi-Agent Reinforcement Learning; Cultural Evolution; Dominance Hierarchy; Cooperative AI

ACM Reference Format:
Ram Rachum, Yonatan Nakar, Bill Tomlinson, Nitay Alon, and Reuth Mirsky. 2024. Emergent Dominance Hierarchies in Reinforcement Learning Agents: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

Research in the paradigm of Cooperative AI highlights the potential of AI agents that interact with each other in ways that are inspired by the interaction of biological life forms [7, 9, 36]. The social structures observed in animal and human societies allow the effective cooperation of groups comprised of wildly different personalities [42]. MARL presents an opportunity to implement simplified versions of those social structures (also labeled *institutions*), where the environment is closed and controlled, and humans are replaced by artificial agents [1, 17, 25].

*See additional definitions, plots, and discussion in Rachum et al. [26]

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

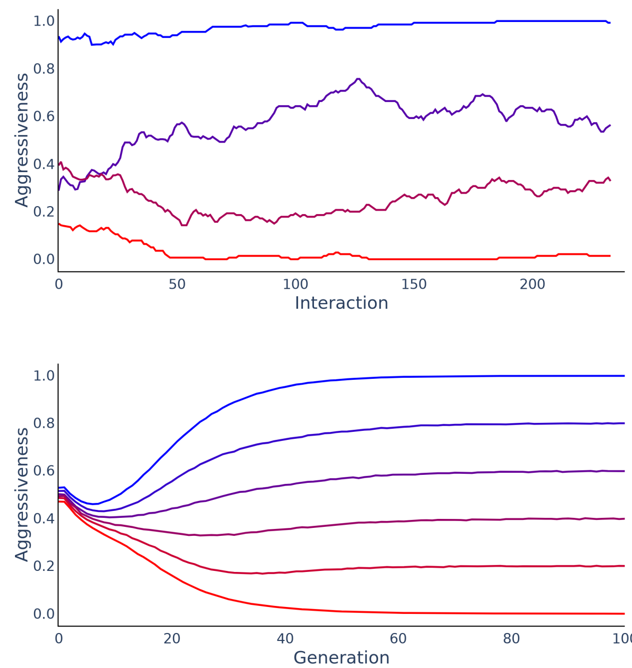

Figure 1: *Top:* Aggressiveness of 4 Leghorn chickens in 14 populations [4]. *Bottom:* Aggressiveness of 6 Chicken Coop RL agents across 300 populations, filtered to linear DHs. Results are averaged using each agent's rank as its identity.

In this work we study a primordial institution that underlies cooperation in animal and human societies: *dominance hierarchies*.

The field of dominance hierarchies began with Schjelderup-Ebbe [31], which described pecking orders in captive chicken societies. Over the past century, dominance hierarchies have been studied in canines [2, 13], birds [37], fish [10], primates [14, 30], and others, uncovering commonalities in conflict resolution and resource allocation across taxa [4]. In humans, dominance hierarchies appear in both explicit and implicit forms [15, 18, 23]. The "chain of command" that underpins large-scale human enterprise is an extension of human dominance behavior to groups that are far too large for any single individual to comprehend [6, 24].

We adapt the ethological theory of dominance hierarchies to artificial agents by modelling agonistic behavior in an N-player variant of the game of Chicken [27] that we call *Chicken Coop*. We borrow established terminology from animal study to allow the measurement of dominance behavior. We release the Chicken Coop environment under the MIT open-source license.[1] We train RL agents on the Chicken Coop environment to optimize their reward, observing three emergent phenomena:

(1) Agents collaboratively invent dominance hierarchies.
(2) Agents enforce dominance hierarchies on other agents.
(3) Agents transmit dominance hierarchies to new populations.

Vezhnevets et al. [35] and Wu et al. [39] study the efficacy of hierarchies of agents powered by Large Language Models (LLMs) [40] working together on a common task. With the advent of multimodal foundation models [21] and their seamless integration with LLMs, the potential usefulness of hierarchies of such agents is tremendous; we suggest that it may be maximized by a formal understanding and terminology of agent hierarchies.

2 DEFINITIONS

We assume a partially-observable stochastic game (POSG) [16, 41], such that for any two agents, there are two stable Nash equilibria NE_i and NE_j such that agent i's reward at NE_i is bigger than its reward at NE_j, and vice versa for agent j [19]. $\mathbb{T}$ is a set of timesteps.

We define dominance relationships between agents as a function of their aggressiveness levels towards each other. Similarly to Leibo et al. [20], we define an agent's aggressiveness by how frequently it chooses an action that reduces another agent's reward:

Definition 2.1 (Aggressiveness). Agent i's *aggressiveness* $g_i^{\mathbb{T}}$ is the portion of timesteps out of $\mathbb{T}$ in which it played $a^i \in NE_i$.

Definition 2.2 (Dominance relationship, dominant, subordinate). If the difference between agent i's aggressiveness and agent j's aggressiveness is above a certain threshold, we say that the two agents are in a *dominance relationship* (DR), with agent i being *dominant* and agent j being *subordinate*:

$$i \to j \quad \text{iff} \quad g_i^{\mathbb{T}} - g_j^{\mathbb{T}} > \eta, \quad \eta \in (0, 1]$$

Definition 2.3 (Dominance hierarchy). A *dominance hierarchy* $\mathcal{H}$ is a complete, directed graph where agents are represented as nodes and dominance relationships are represented as directed edges.

To investigate dominance hierarchies in groups of more than two individuals, we introduce the Chicken Coop environment:

Definition 2.4 (Chicken Coop). Chicken Coop is a POSG with N players. In each episode, agents are divided into random pairs. Each pair of agents plays one round of Chicken against each other, choosing either *hawk* or *dove* and receiving a reward in $\{W, T, L, C\}$. Each agent's sole observation is the identity of their opponent agent.

Additional definitions are made in Rachum et al. [26].

[1] Code and usage instructions are available at https://github.com/cool-RR/chicken-coop

3 EMPIRICAL RESULTS

We run experiments on the Chicken Coop environment using the PPO algorithm [32]. We use $N \in [6, 20], L \in [10, 300], \alpha \in [2^{-6}, 3^{-5}], \gamma = 0.99, \epsilon = 0.3$.

Each of the $L = 300$ runs of our experiment resulted in agents converging to a dominance hierarchy; between these 300 populations the hierarchies were divergent, as most populations developed a hierarchy that is unique to them. Similarly to animal groups, some of the formed hierarchies were linear (transitive), and some were non-linear [5, 11, 29]. In linear hierarchies, agents tend to maintain similarly-sized intervals in their mean aggressiveness (Figure 1.)

Inspired by the geometric study of intransitive policies in Czarnecki et al. [8], we note a resemblance in the occurrence of intransitive components (cycles) between Chicken Coop populations and those found in experiments with populations of CD-1 mice [33, 38].

Lastly, we run a two-stage experiment, where agents learn a dominance hierarchy in one Chicken Coop environment, and then a subset of them are transplanted into a new Chicken Coop environment with untrained agents, also known as *naive agents* [12]. We show that the *experienced agents* are able to teach the naive agents the same hierarchy that they've learned, showing cultural evolution similar to that presented in CGI Team [3], except in a much simpler environment with modest computational demands. This experiment may be repeated with the naive agents functioning as experienced agents in a new, third population, and so on to an indefinite string of populations, untethering the concept of dominance hierarchies from dependence on any specific host agents.

4 FUTURE WORK

Future work may explore the impact of various algorithms and parameters on the properties of dominance hierarchies. Preliminary experiments suggest that higher learning rates result in unstable dominance hierarchies, a phenomenon akin to *rank change* in animal societies [4, 28, 34]. We suggest that Opponent Shaping algorithms such as M-FOS [22], which consider the learning processes of other agents, could promote second-order dominance-seeking strategies, e.g., agent i may consider how to behave as to encourage agent j to place agent i at a high rank in the dominance hierarchy.

When humans work on problems as a group, we balance in-group intrigues against external pressures. This interplay between individual and group needs may play a crucial role in the success of our collective intelligence. Therefore, we propose augmenting a population of dominance-seeking agents with multimodal foundation models [21], and giving them real-world tasks to collaborate on. We hypothesize that the decisions made by such agents may be interpretable and corrigible, as human operators may recognize that the agents' decision process reflects their own.

ACKNOWLEDGMENTS

We thank the following colleagues for their advice and support in writing this paper: David Aha, Ivan Chase, Edgar Duéñez-Guzmán, Errol King, Joel Leibo, Olof Leimar, David Manheim, Markov, Georg Ostrovski, Jérémy Perret, Saul Pwanson, Venkatesh Rao and Eli Strauss.

This research was funded by The Association For Long Term Existence And Resilience (ALTER).

REFERENCES

[1] Stefano V. Albrecht, Filippos Christianos, and Lukas Schäfer. 2024. *Multi-Agent Reinforcement Learning: Foundations and Modern Approaches.* MIT Press.

[2] Roberto Bonanni, Simona Cafazzo, Arianna Abis, Emanuela Barillari, Paola Valsecchi, and Eugenia Natoli. 2017. Age-graded dominance hierarchies and social tolerance in packs of free-ranging dogs. *Behavioral Ecology* 28, 4 (2017), 1004–1020.

[3] CGI Team. 2022. Learning Robust Real-Time Cultural Transmission without Human Data. *arXiv:2203.00715* (2022).

[4] Ivan D Chase, Darius Coelho, Won Lee, Klaus Mueller, and James P Curley. 2022. Networks never rest: an investigation of network evolution in three species of animals. *Social Networks* 68 (2022), 356–373.

[5] Ivan D Chase, Craig Tovey, Debra Spangler-Martin, and Michael Manfredonia. 2002. Individual differences versus social dynamics in the formation of animal dominance hierarchies. *Proceedings of the National Academy of Sciences* 99, 8 (2002), 5744–5749.

[6] Tian Chen Zeng, Joey T Cheng, and Joseph Henrich. 2022. Dominance in humans. *Philosophical Transactions of the Royal Society B* 377, 1845 (2022), 20200451.

[7] Vincent Conitzer and Caspar Oesterheld. 2023. Foundations of cooperative AI. In *AAAI*, Vol. 37. 15359–15367.

[8] Wojciech M Czarnecki, Gauthier Gidel, Brendan Tracey, Karl Tuyls, Shayegan Omidshafiei, David Balduzzi, and Max Jaderberg. 2020. Real world games look like spinning tops. *Advances in Neural Information Processing Systems* 33 (2020), 17443–17454.

[9] Allan Dafoe, Edward Hughes, Yoram Bachrach, Tantum Collins, Kevin R. McKee, Joel Z. Leibo, Kate Larson, and Thore Graepel. 2020. Open Problems in Cooperative AI. arXiv:2012.08630 [cs.AI]

[10] Bruno O David and Rick J Stoffels. 2003. Spatial organisation and behavioural interaction of giant kokopu (Galaxias argenteus) in two stream pools differing in fish density. (2003).

[11] Han De Vries. 1995. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour* 50, 5 (1995), 1375–1389.

[12] Edgar A Duéñez-Guzmán, Kevin R McKee, Yiran Mao, Ben Coppin, Silvia Chiappa, Alexander Sasha Vezhnevets, Michiel A Bakker, Yoram Bachrach, Suzanne Sadedin, William Isaac, et al. 2021. Statistical discrimination in learning agents. *arXiv:2110.11404* (2021).

[13] Jennifer L Essler, Simona Cafazzo, Sarah Marshall-Pescini, Zsófia Virányi, Kurt Kotrschal, and Friederike Range. 2016. Play behavior in wolves: Using the '50:50' rule to test for egalitarian play styles. *PLoS One* 11, 5 (2016), e0154150.

[14] Jane Goodall. 1986. *The Chimpanzees of Gombe: Patterns of Behavior.* Belknap Press of Harvard University Press.

[15] Nir Halevy, Eileen Y Chou, Taya R Cohen, and Robert W Livingston. 2012. Status conferral in intergroup social dilemmas: behavioral antecedents and consequences of prestige and dominance. *Journal of personality and social psychology* 102, 2 (2012), 351.

[16] Eric A Hansen, Daniel S Bernstein, and Shlomo Zilberstein. 2004. Dynamic programming for partially observable stochastic games. In *AAAI*, Vol. 4. 709–715.

[17] Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Caglar Gulcehre, Pedro Ortega, DJ Strouse, Joel Z Leibo, and Nando De Freitas. 2019. Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *International conference on machine learning*. PMLR, 3040–3049.

[18] Keith Johnstone. 1981. *Impro: Improvisation and the Theatre.* Routledge. 33–74 pages.

[19] Elon Kohlberg and Jean-Francois Mertens. 1986. On the strategic stability of equilibria. *Econometrica: Journal of the Econometric Society* (1986), 1003–1037.

[20] Joel Z. Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel. 2017. Multi-agent Reinforcement Learning in Sequential Social Dilemmas. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems* (São Paulo, Brazil) *(AAMAS '17)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 464–473.

[21] Chunyuan Li, Zhe Gan, Zhengyuan Yang, Jianwei Yang, Linjie Li, Lijuan Wang, and Jianfeng Gao. 2023. Multimodal Foundation Models: From Specialists to General-Purpose Assistants. arXiv:2309.10020 [cs.CV]

[22] Christopher Lu, Timon Willi, Christian A Schroeder De Witt, and Jakob Foerster. 2022. Model-free opponent shaping. In *International Conference on Machine Learning*. PMLR, 14398–14411.

[23] Joe C Magee and Adam D Galinsky. 2008. Social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals* 2 (2008), 351–398. Issue 1.

[24] Henry Mintzberg. 1989. *The structuring of organizations.* Springer.

[25] Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *AAAI*, Vol. 32.

[26] Ram Rachum, Yonatan Nakar, Bill Tomlinson, Nitay Alon, and Reuth Mirsky. 2024. Emergent Dominance Hierarchies in Reinforcement Learning Agents. arXiv:2401.12258 [cs.MA]

[27] Anatol Rapoport and Albert M. Chammah. 1966. The Game of Chicken. *American Behavioral Scientist* 10, 3 (1966), 10–28. https://doi.org/10.1177/000276426601000303 arXiv:https://doi.org/10.1177/000276426601000303

[28] Amy Samuels, Joan B Silk, and Jeanne Altmann. 1987. Continuity and change in dominance relations among female baboons. *Animal Behaviour* 35, 3 (1987), 785–793.

[29] Nathan J Sanders and Deborah M Gordon. 2003. Resource-dependent interactions and the organization of desert ant communities. *Ecology* 84, 4 (2003), 1024–1031.

[30] Robert M Sapolsky. 2005. The influence of social hierarchy on primate health. *science* 308, 5722 (2005), 648–652.

[31] Thorleif Schjelderup-Ebbe. 1922. Beiträge zur Sozialpsychologie des Haushuhns. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane. Abt. 1. Zeitschrift für Psychologie* 88 (1922), 225.

[32] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv:1707.06347* (2017).

[33] Eli D Strauss, Alex R DeCasien, Gabriela Galindo, Elizabeth A Hobson, Daizaburo Shizuka, and James P Curley. 2022. DomArchive: a century of published dominance data. *Philosophical Transactions of the Royal Society B* 377, 1845 (2022), 20200436.

[34] Eli D Strauss and Daizaburo Shizuka. 2022. The dynamics of dominance: open questions, challenges and solutions. *Philosophical Transactions of the Royal Society B* 377, 1845 (2022), 20200445.

[35] Alexander Sasha Vezhnevets, John P. Agapiou, Avia Aharon, Ron Ziv, Jayd Matyas, Edgar A. Duéñez-Guzmán, William A. Cunningham, Simon Osindero, Danny Karmon, and Joel Z. Leibo. 2023. Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia. arXiv:2312.03664 [cs.AI]

[36] Eugene Vinitsky, Raphael Köster, John P Agapiou, Edgar A Duéñez-Guzmán, Alexander S Vezhnevets, and Joel Z Leibo. 2023. A learning agent that acquires social norms from public sanctions in decentralized multi-agent settings. *Collective Intelligence* 2, 2 (2023). https://doi.org/10.1177/26339137231162025

[37] John R Watson. 1970. Dominance-subordination in caged groups of house sparrows. *The Wilson Bulletin* (1970), 268–278.

[38] Cait M Williamson, Won Lee, Russell D Romeo, and James P Curley. 2017. Social context-dependent relationships between mouse dominance rank and plasma hormone levels. *Physiology & behavior* 171 (2017), 110–119.

[39] Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Beibin Li, Erkang Zhu, Li Jiang, Xiaoyun Zhang, Shaokun Zhang, Jiale Liu, Ahmed Hassan Awadallah, Ryen W White, Doug Burger, and Chi Wang. 2023. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation Framework. arXiv:2308.08155 [cs.AI]

[40] Zhiheng Xi, Wenxiang Chen, Xin Guo, Wei He, Yiwen Ding, Boyang Hong, Ming Zhang, Junzhe Wang, Senjie Jin, Enyu Zhou, Rui Zheng, Xiaoran Fan, Xiao Wang, Limao Xiong, Yuhao Zhou, Weiran Wang, Changhao Jiang, Yicheng Zou, Xiangyang Liu, Zhangyue Yin, Shihan Dou, Rongxiang Weng, Wensen Cheng, Qi Zhang, Wenjuan Qin, Yongyan Zheng, Xipeng Qiu, Xuanjing Huang, and Tao Gui. 2023. The Rise and Potential of Large Language Model Based Agents: A Survey. arXiv:2309.07864 [cs.AI]

[41] Yaodong Yang and Jun Wang. 2020. An overview of multi-agent reinforcement learning from game theoretical perspective. *arXiv:2011.00583* (2020).

[42] Siran Zhan, Namrita Bendapudi, and Ying-yi Hong. 2015. Re-examining diversity as a double-edged sword for innovation process. *Journal of Organizational Behavior* 36, 7 (2015), 1026–1049.