# Emergent Dominance Hierarchies in Reinforcement Learning Agents*

Ram Rachum
Bar-Ilan University
rachumr@biu.ac.il

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, and successful cooperation in mixed-motive groups of agents depends on a delicate balancing act between individual and group objectives. Social conventions and norms, often inspired by human institutions, are used as tools for striking this balance.

In this paper, we examine a fundamental, well-studied social convention that underlies cooperation in both animal and human societies: dominance hierarchies.

We adapt the ethological theory of dominance hierarchies to artificial agents, borrowing the established terminology and definitions with as few amendments as possible. We demonstrate that populations of RL agents, operating without explicit programming or intrinsic rewards, can invent, learn, enforce, and transmit a dominance hierarchy to new populations. The dominance hierarchies that emerge have a similar structure to those studied in chickens, mice, fish, and other species.

## KEYWORDS

Multi-Agent Reinforcement Learning, Reinforcement Learning, Cultural Evolution, Multi-Agent Systems, Social AI, Dominance Hierarchy, Cooperative AI

## 1 INTRODUCTION

Many animal species are able to collaborate, form groups and harness collective intelligence [15, 31]. Among those species, humans have achieved a scale and sophistication of collaboration that stands as one of the most profound and unparalleled phenomena on Earth. The grand feats of humanity were accomplished by groups of humans working together; what is grander is that these groups are composed of individuals with wildly different beliefs, motivations, and talents [85]. While differences between individuals provide an important source of diversity, they also lead to adversity and conflict, which can result in the violent demise of the social group [91]. Those human civilizations that tempered their internal conflicts while capitalizing on the diversity were the ones to survive and flourish [1, 7, 8]. The intricate social structures that make up these civilizations turn a potential runaway explosion into a controlled reaction, enabling groups of humans to achieve breakthroughs that exceed the combined capabilities of their individual members [55].

Multi-agent reinforcement learning (MARL) presents an opportunity to implement simplified versions of those intricate social structures, where the environment is closed and controlled, and humans are replaced by artificial agents. Research in the paradigm of Cooperative AI [14, 17, 18, 41, 59, 80] highlights the role of such social structures (also labeled *institutions*) in producing human-aligned AI systems. Existing research into reproducing such institutions for use in MARL settings includes language [54], voting systems [13, 88], auctions [56], reputation systems [60], and bargaining [44].

In this paper we take a step back from these cerebral and centralized institutions and turn our attention to a primordial institution that underlies cooperation in both animal and human societies: *dominance hierarchies*.

Our contribution in this work is twofold:

First, we adapt the ethological theory of dominance hierarchies to artificial agents by modeling conflicts between each pair of individuals as the classic game of Chicken [62].[1] To accommodate groups of more than two individuals, we generalize the game of Chicken into an $N$-player stochastic game that we call *Chicken Coop*. We borrow the established terminology and definitions from animal study to allow calculating dominance metrics from agents' action history. We release the Chicken Coop environment under the MIT open-source license.[2]

Second, we train RL agents on the Chicken Coop environment to optimize their score, observing the emergence of three behaviors:

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

These behaviors emerge from minimal environment rules and individual behavior determined by a stock RL algorithm with no intrinsic rewards.

We experiment with gradually ablating the agents' observations, showing a causal link between the agents' ability to identify the opposing agent and the emergence of dominance hierarchies. We compare the attributes of dominance hierarchies that emerge to

*A condensed version of this work has been presented as an extended abstract at the 2024 International Conference on Autonomous Agents and Multiagent Systems (AAMAS) [58]

*Proc. of the Adaptive and Learning Agents Workshop (ALA 2024), Avalos, Milec, Müller, Wang, Yates (eds.), May 6-7, 2024, Online, https://ala2024.github.io/.* 2024.

[1]We use the word "chicken" in this paper to refer to both the Game Theory concept and the beak-and-feathers animal. We capitalize "Chicken" to indicate the former.
[2]The code and installation instructions are available at https://github.com/cool-RR/chicken-coop

empirical results from groups of different animals, showing structural similarities in both aggressiveness levels and the distribution of intransitive triads.

This project seeks to lay the intellectual groundwork for dominance hierarchies to feature prominently in future MARL systems. Dominance hierarchies have been shown to provide substantial value to both animal and human communities [53, 69]. By providing MARL systems with the ability to coordinate their collective behavior via dominance hierarchies, we hope to enable those systems to reap similar benefits and integrate more seamlessly into existing human systems [63].

# 2 BACKGROUND

## 2.1 Dominance hierarchies in nature

When several animals live together in a social group they often find themselves in conflicts with other members of the group over resources such as food or mating partners. These conflicts can lead to physical injury and death. A *dominance relationship* [61] is a pattern in the conflicts between a pair of individuals, in which the *dominant* individual is highly likely to escalate violence and win the majority of resources, and the *subordinate* individual is highly likely to deescalate and yield any contested resources. This dynamic prevents the runaway escalation of violence that could otherwise lead to individual injury and a loss of group cohesion [69].

A *dominance hierarchy*, also known as a *pecking order*, is the aggregate of all the dominance relationships between each pair of agents in the group. It is defined by modelling dominance relationships as a total relation on the set of agents, where $d \rightarrow s$ means that individual $d$ is dominant over individual $s$. A dominance hierarchy is modelled as a complete, directed graph (tournament [22]) where agents are represented as nodes and dominance relationships are represented as directed edges [40]. This graph representation is more than just a map of agonistic behavior in the group; any skirmish between individuals $a$ and $b$ affects not only the edge connecting $a$ and $b$, but the edges connecting either $a$ or $b$ with each of the other individuals in the group, by the mechanisms of *winner effects* [9], *loser effects* [25] and *bystander effects* [26].

The field of dominance hierarchies traces its roots to Schjelderup-Ebbe [67], which described pecking orders in captive chicken societies. Over the past century, dominance hierarchies have been studied across a wide range of animal species including canines [6, 27], birds [82], elephants [84], fish [19] and primates [30, 66], uncovering commonalities in the structure and function of dominance hierarchies across taxa [10]. The study of dominance hierarchies continues to be a field of active research [71].

## 2.2 Dominance hierarchies in humans

In human societies, dominance hierarchies appear in many forms, ranging from explicit and rigid to implicit and fluid [33, 51]. The explicit forms, often appearing in corporations, military units, and sports teams, are laid bare for analysis, as they are announced with official titles, insignia patches, and organizational charts. The "chain of command" principle 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 [32, 53, 81].

Implicit forms of dominance hierarchies often go unspoken, making them more difficult to study; however, through subtlety, their effect on human behavior is profound [2, 74]. Johnstone [42] provides a manual for theater actors to faithfully reproduce real-world social scenarios, noting the importance of dominance hierarchies:

> "Try to get your status just a little above or below your partner's," I said, and I insisted that the gap should be minimal. The actors seemed to know exactly what I meant and the work was transformed. The scenes became 'authentic', and actors seemed marvellously observant. Suddenly we understood that every inflection and movement implies a status [...] It was hysterically funny, but at the same time very alarming. All our secret manoeuvrings were exposed. [...] Normally we are 'forbidden' to see status transactions except when there's a conflict. In reality status transactions continue all the time.

We note that, while dominance hierarchies form an important part of human social interactions, they are nevertheless just one component of the complex array of motivations that together produce social behavior [12].

# 3 RELATED WORK

While we advocate using dominance hierarchies to improve the performance of artificial agents, considerable work has been done in the opposite direction, applying MAS and MARL to improve our understanding of dominance hierarchies in animals. We use this distinction to roughly divide the work in the intersection of these fields to two categories, detailed below.

## 3.1 Applying MAS/MARL to dominance hierarchies

MAS models have been used for the modelling of dominance hierarchies in animals [29, 39] and in humans [57]. *DomWorld* [36] is a multi-agent environment developed to study the self-organization of social behavior and dominance hierarchies in primate groups. DomWorld demonstrates the decentralized nature of dominance hierarchies: each agent's actions are determined according to its own logic and observations, and the collective actions of all agents cause dominance hierarchies to emerge on the population level. These dominance hierarchies have similar properties to those found in real animal societies. Notably, DomWorld is not modelled as a stochastic game or any other game theoretic construct, as the agents have no concept of a reward signal that they're trying to maximize. Agents follow hard-coded logic to decide whether to attack an opponent or flee, helpfully described in a flowchart in Hemelrijk et al. [37].

Leimar [46] provides a proper MARL environment called *SocDom*, equipping the agents with an RL algorithm [72], albeit without neural network estimators. While the SocDom experiments show similar results to ours, the foci of the two environments differ: SocDom, like DomWorld, attempts to tailor the agents' behaviors to faithfully reproduce the dominance behavior seen in real animal societies. This is evident in SocDom's use of genetic algorithms in addition to RL. In contrast, our research is directed at boiling the mechanics of dominance hierarchies down to the simplest possible

set of premises, showing that they emerge even when using an off-the-shelf RL algorithm.

### 3.2 Applying dominance hierarchies to MAS/MARL

The use of dominance hierarchies for improving the performance of artificial agents was proposed in Hemelrijk [35], and in Tomlinson et al. [76], which suggested that dominance hierarchies could streamline negotiation, enable agents to form beneficial alliances, and make interfaces more intuitive.

Bakker et al. [5] employs a group of RL agents with a graph structure, which bears resemblance to a dominance hierarchy, besides being explicitly defined rather than emergent. The authors show how the dominance-like relationship between each pair of individuals results in increased cooperation from both sides.

Aroca-Ouellette et al. [3] demonstrates an explicit hierarchy where some agents have a different set of actions than others, depending on their position in the hierarchy. The authors show that the hierarchy improves the ability of humans to interact with the system, as humans find the agents more fluent and trustworthy.

## 4 DEFINITIONS

In this section we adapt the terms and definitions used in the study of dominance hierarchies in animal societies to the frameworks used in multi-agent reinforcement learning.

### 4.1 Dominance between two agents

Dominance hierarchies are a group phenomenon comprised of multiple dominance relationships, one between each pair of agents in the group. In this subsection we consider only those pairwise interactions; we define dominance relationships, the metrics of aggressiveness and rapport, the roles of dominant and subordinate, and the environment in which these occur.

Animals use dominance relationships to decide which individual will have access to an exclusive resource, such as food, a mating partner, or grooming by other individuals [38]. We model the environment that enables these interactions between agent $i$ and agent $j$ as a partially-observable stochastic game (POSG) [34, 89] which has exactly two stable [43] 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 agent $j$'s reward at $NE_j$ is bigger than its reward at $NE_i$.

We denote the actions that comprise $NE_i$ as $a_i^{NE_i}$ and $a_j^{NE_i}$.

Intuitively, a dominance relationship between two agents is the tendency of those agents to play the joint action of either $NE_i$ or $NE_j$, but not both. Similarly to Leibo et al. [45], we define an agent's aggressiveness by how frequently it chooses the action that reduces the other agent's reward:

*Definition 4.1 (Aggressiveness).* Given an agent $i$ and a set of timesteps $\mathbb{T}$, the agent's *aggressiveness* $g_i^{\mathbb{T}}$ is the portion of timesteps in which it played $a_i^{NE_i}$:

$$g_i^{\mathbb{T}} = \frac{|\{t \in \mathbb{T} \mid a_i^t = a_i^{NE_i}\}|}{|\mathbb{T}|}$$

*Definition 4.2 (Dominance relationship, dominant, subordinate).* Given two agents $i$ and $j$ and a set of timesteps $\mathbb{T}$ in which they played with each other, 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*:

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

In order to determine the existence of a dominance relationship between two agents without making assumptions about its polarity, we introduce a metric we call *rapport*:

*Definition 4.3 (Rapport).* For two agents $i$ and $j$ and timesteps $\mathbb{T}$, we define the *rapport* $\mathfrak{R}_{i,j}^{\mathbb{T}}$ as:

$$\mathfrak{R}_{i,j}^{\mathbb{T}} = |g_i^{\mathbb{T}} - g_j^{\mathbb{T}}|$$

A rapport $\mathfrak{R}_{i,j}^{\mathbb{T}} > \eta$ indicates either $i \rightarrow j$ or $i \leftarrow j$.

There are infinitely many environments that satisfy the conditions needed for dominance relationships to arise. Przepiorka et al. [57] defines the Hierarchy Contest Game (HCG), which is a generalization of the classic game of Chicken [62] to a game where players have different capabilities. They show how dominance hierarchies form in human play, whether the players start with knowledge of each other's capabilities or not. Because we wish to describe the simplest possible environment that can produce dominance relationships, we use the game of Chicken with no modifications. Its salient feature is the pair of symmetrically opposing, stable, pure-strategy Nash equilibria.

In the game of Chicken, each of the two agents plays either *hawk* or *dove*. An agent gets the most reward if it plays *hawk* while its partner plays *dove*; however, if both agents play *hawk*, they both get the lowest reward.

*Definition 4.4 (The game of Chicken).* The game of Chicken is defined as a normal-form game:

| Agent 1 \ Agent 2 | Dove | Hawk |
|---|---|---|
| Dove | $R, R$ | $S, T$ |
| Hawk | $T, S$ | $P, P$ |

The reward constants satisfy $T > R > S > P$. The $(S, T)$ and $(T, S)$ outcomes serve as the Nash equilibria $NE_i$ and $NE_j$ . Each agent's aggressiveness is simply its tendency to play *hawk*.

### 4.2 Dominance hierarchies

To model dominance relationships within an agent population, we generalize the game of Chicken to support more than two players, visualized with $N = 6$ in Figure 1:

*Definition 4.5 (Chicken Coop).* *Chicken Coop* is an $N$-player generalization of the game of Chicken. In each episode of Chicken Coop, the agents are divided into random pairs that are used for the entire episode. Each pair of agents plays one round of Chicken against each other, choosing either *hawk* or *dove* and receiving a reward in $\{T, R, S, P\}$. Each agent's sole observation is the identity of their opponent agent.

We extend the definitions in Subsection 4.1 from the Chicken game to the Chicken Coop game. For example, for each pair of agents $i$ and $j$ in the population, we determine the existence and

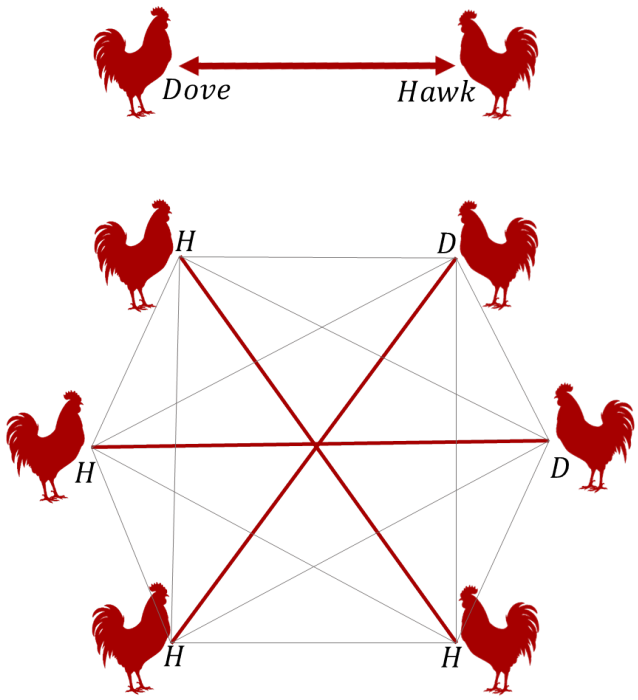


**Figure 1: The game of Chicken (top) compared with Chicken Coop (bottom). Each Chicken Coop episode can be described as $N/2$ chicken games between randomly-paired agents.**

polarity of their dominance relationship ($i \rightarrow j$ or $i \leftarrow j$) by evaluating the aggressiveness metrics $g_i^{\mathbb{T}}$ and $g_j^{\mathbb{T}}$ on the subset of $\mathbb{T}$ in which these two agents were randomly paired with each other. Finally, we aggregate these dominance relationships into a graph:

*Definition 4.6 (Dominance hierarchy).* A *dominance hierarchy* $\mathcal{H}$ is a complete, directed graph (also known as a tournament [22],) where agents are represented as nodes and dominance relationships are represented as directed edges.

## 5 METHODS

In all of the experiments below, unless stated otherwise, the RL agents are trained using the PPO algorithm [68] implementation provided by the RLlib framework [48]. We train $L = 300$ populations of $N = 6$ Chicken Coop agents, at a learning rate of $2 \times 10^{-6}$, a discount factor of $\gamma = 0.99$, and a clipping parameter of $\epsilon = 0.3$. We use a dominance threshold of $\eta = 0.55$ and these reward constants:

$$R = 0, \quad S = -2, \quad T = 5, \quad P = -10$$

Each agent has its own neural network with its own set of weights, which are separate from those of the other agents. Crucially, all of the agents in a given population are trained exclusively with the other agents in that population; we don't allow cross-population training, because we would like to allow each population to develop its own distinct dominance hierarchy.

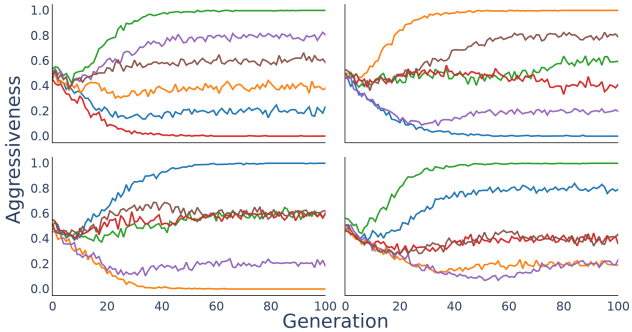


**Figure 2: Aggressiveness levels of each of the six agents (depicted as differently-colored lines) in four sample Chicken Coop populations.**

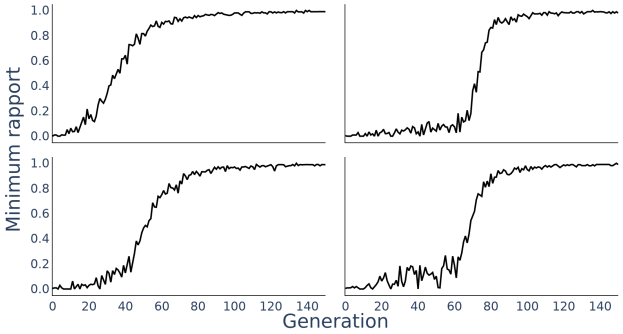


**Figure 3: Minimum rapport between each pair of the six agents in four sample Chicken Coop populations.**

Each generation is comprised of 512 episodes. After each generation, the policies update according to the actions and rewards in that generation.

## 6 RESULTS

### 6.1 Emergence of dominance hierarchies

We measure the aggressiveness of each agent in each population, aggregated from its interactions with each of the five other agents. In Figure 2 we plot the aggressiveness of all six agents in four sample populations. A few features can be visually discerned:

(1) The agents' aggressiveness values converge to approximately equally-spaced "ranks" in $[0, 1]$.
(2) Despite of this strong convergence, the mapping between agents and ranks appears completely arbitrary across the four sample populations.
(3) In some populations, the mapping between agents and ranks is a bijection, while in other populations, multiple agents converge to the same rank, leaving other ranks empty.

We offer the following explanation for these observations: when agent $i$'s aggressiveness approaches $k/N$, agent $i$ is playing a high proportion of *hawk* towards $k$ other agents, which in turn play a high proportion of *dove* towards agent $i$. Each of these $k$ agents

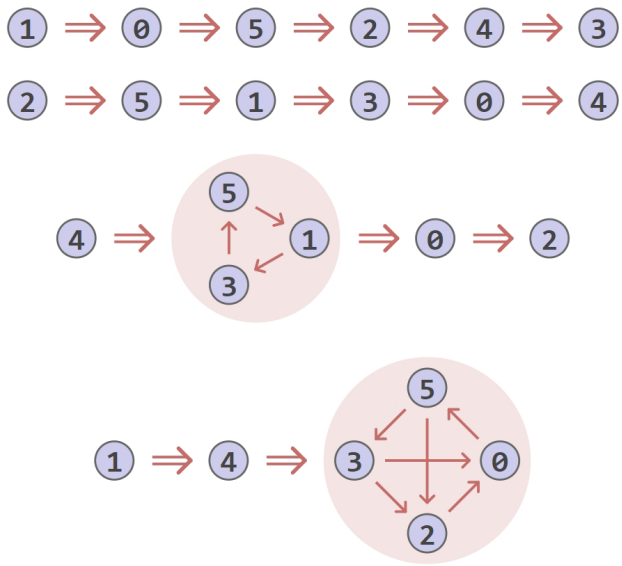


**Figure 4: Dominance hierarchies of four sample Chicken Coop populations, visualized in condensed form. A single arrow (→) represents a dominance relationship between two agents. A double arrow (⇒) represents a transitive relationship, i.e. each of the agents on the left side of the double arrow dominates each of the agents on the right side of that arrow.**

contributes $1/N$ to agent $i$'s aggressiveness. According to Definition 4.2, agent $i$ is dominating these $k$ agents. The remaining $N-k-1$ agents dominate agent $i$.

We verify this explanation by measuring the minimum rapport between each pair of agents for each of the four sample populations, observing that each agent develops a dominance relationship with each of the other agents in its population. (Figure 3.)

We aggregate these dominance relationships into dominance hierarchies, showing the results for our four sample populations in Figure 4. We observe that some of the dominance hierarchies form a perfect line, while others contain cycles. This distinction has been studied in animal societies. The former are called *linear dominance hierarchies* or *transitive dominance hierarchies*. The latter are called *nonlinear dominance hierarchies*, *near-linear dominance hierarchies* or *intransitive dominance hierarchies* [11, 20, 21, 65]. In linear dominance hierarchies, the agents may be assigned ranks $\{0, 1, 2, \ldots, N-1\}$ with rank 0 representing the agent that dominates all other agents, rank 1 representing the agent that dominates all other agents except that in rank 0, etc.

One profound property of the agents' learned behavior is its idiosyncraticity: (1) it converges to a stable dominance hierarchy, and (2) this hierarchy is not a single one, but rather varies between different trials. Our 300 populations converged to a total of 248 distinct dominance hierarchies (out of $2^{\frac{N(N-1)}{2}} = 2^{15}$ possible permutations). Most dominance hierarchies appeared in only one population, while the most common dominance hierarchies appeared in 3 populations. Despite the arbitrariness of that choice,

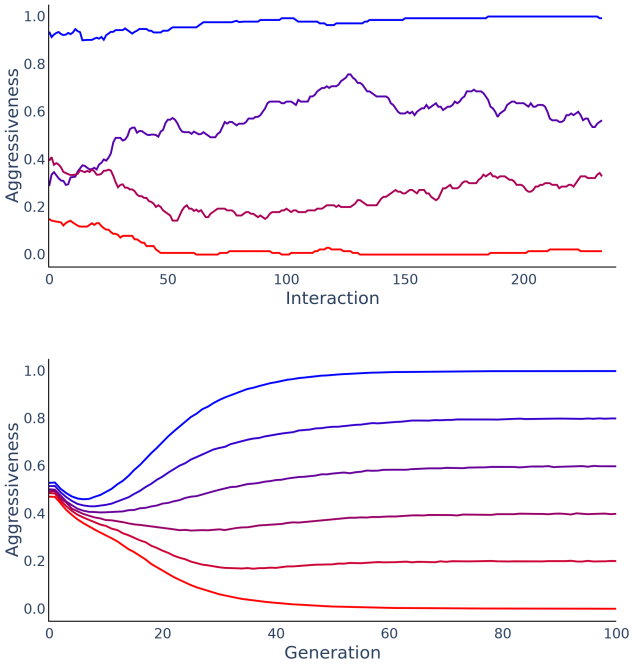


**Figure 5: Top: Mean aggressiveness levels of each of four Leghorn chickens in fourteen captive populations [10]. Bottom: Mean aggressiveness levels of each of six Chicken Coop agents across 300 populations, filtered to include only those that developed a linear dominance hierarchy. In both plots, the results are averaged using the agents' ranks as their identity.**

each population converges strongly to its chosen dominance hierarchy. Out of 248 distinct dominance hierarchies, only 17 were intransitive, i.e. containing dominance cycles.

## 6.2 Comparison to animal behavior

In Figure 5 we visually compare our aggressiveness results to empirical data from groups of actual chickens. In Chase et al. [10], 14 captive populations of 4 white Leghorn hens each were observed for aggressive behavior for two consecutive days. While there are considerable differences in the methodologies of that experiment and our study, the property of idiosyncraticity holds true in both of them. In both environments, the agents tend to maintain their position, regardless of how high or low it is in the hierarchy.

Inspired by the geometric study of intransitive policies in Czarnecki et al. [16], we visually compare the occurrence of intransitive components (cycles) between Chicken Coop populations and mice populations, shown in Figure 6. In Williamson et al. [83], 20 captive populations of 12 male CD-1 mice each were observed for aggressive behavior for 22 days.[3]. We compare these results to those from $L = 30$ populations of $N = 12$ Chicken Coop agents. In order to find the maximum similarity between the natural and artificial

[3]The data was conveniently aggregated by Strauss et al. [70]

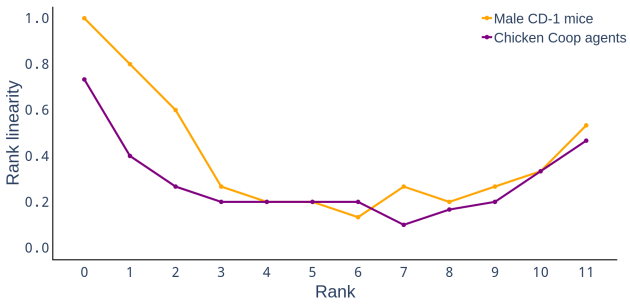


**Figure 6: The rank linearities of 12 different ranks, in both Chicken Coop agents (purple) and male CD-1 mice (orange)[83]. In both experiments, the top and bottom ranks were the least likely to be occupied by an intransitive component.**

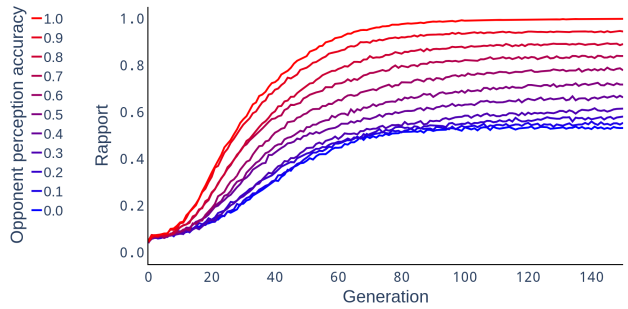


**Figure 7: The gradual rise of rapport in Chicken Coop populations, differentiated by how accurately the agents identify their opponents. The bluer lines represent bigger amounts of random noise added to the agents' observation signal.**

populations, we use a learning rate of $3 \times 10^{-5}$. We make this comparison by using a metric we define here as *rank linearity*[4], which expresses how likely each dominance rank is to be occupied by a single agent, rather than a $K$-way tie. For a single population, we can describe rank linearity as a boolean; for example, in the third population from the top in Figure 4, ranks $\{0, 4, 5\}$ are linear, while ranks $\{1, 2, 3\}$ are non-linear as they are part of a 3-way tie between agents 1, 3 and 5. For a set of populations, we define rank linearity as a number in $[0, 1]$ which is the probability of a rank being linear in a sample population.

## 6.3 Observation ablation

Tomlinson [77] proposes that an agent's perception of its opponent agent is a necessary condition for the formation of social relationships between agents. In this subsection we show a causal relationship between the ability of agents to identify their opponents and their tendency to form dominance hierarchies.

In the Chicken Coop environment, each agent observes the identity of its opponent as an index number between 0 and $N - 1$. We add gradually-rising amounts of random noise to the agents' observation, up to and including its complete obfuscation, and measure the resulting rapport in the population.

**Methodology:** We define an agent's *opponent perception accuracy* (OPA) as a number between 0 and 1. For each agent in each episode, the probability that its observation of its opponent is replaced with a random index number is $(1 - OPA)$. For each OPA value in $\{0, 0.1, 0.2, \ldots, 1\}$ we train $L = 10$ different populations of $N = 6$ Chicken Coop agents at a learning rate of $2 \times 10^{-6}$ and measure their rapport.

In Figure 7 we show a positive correlation between the agents' opponent perception accuracy and their rapport, and by extension, their tendency to form dominance hierarchies.

[4]The metric of rank linearity is similar to a metric called *dominance certainty*, introduced in McCowan et al. [52] and measured in rhesus macaque breeding groups.

## 6.4 Transmitting dominance hierarchies to new populations

In previous subsections we regarded Chicken Coop agents as the focal objects of study, while we considered the dominance hierarchies that they form as a pattern in their behavior. In this subsection we propose flipping this perspective on its head: We consider dominance hierarchies as the focal object while the agents are used as vessels that communicate the dominance hierarchy that they learned to a new population, promoting a cultural evolution paradigm in artificial intelligence [73, 75].

This transmission of dominance hierarchies to new populations is made possible by two properties of dominance hierarchies:

(1) The knowledge of each dominance hierarchy is replicated over all the agents that participate in it, with each agent having knowledge of a part of that dominance hierarchy in its neural network.
(2) Each agent learns to consistently play *hawk* with all its subordinates, and this tendency serves as a punishing mechanism [4], or a sanction. Any errant attempt by a subordinate to play *hawk* against a dominant would usually result in a *hawk*/*hawk* outcome, with both agents receiving the lowest possible reward of $P$. All non-bottom agents effectively act as independent enforcers of their dominance hierarchy, punishing any agent that disobeys it.

**Methodology:** We train $L = 10$ different populations of $N = 6$ Chicken Coop agents, similarly to Subsection 6.1, until they converge to dominance hierarchies. We call these *experienced populations*. For each of these $L$ populations, we perform the following procedure $M = 30$ times: for each $K \in \{0, 1, 2, \ldots, N-2\}$ we choose a random sample of $K$ of those agents to be our *experienced agents*. We transplant the $K$ experienced agents into a new population with $N - K$ *naive agents* [24] that have not been trained. Finally, we train the naive agents in the heterogeneous population at a learning rate of $2 \times 10^{-5}$ and observe their dominance behavior.

We label the indices of the experienced agents $I_e$, and the indices of the naive agents $I_n$.

To measure how close each new dominance hierarchy is to its original dominance hierarchy, we define a *dominance hierarchy*

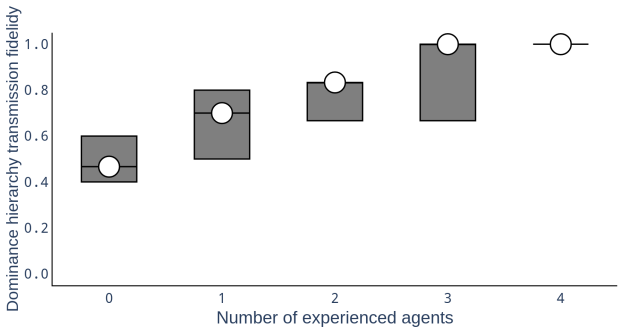


**Figure 8: The dominance hierarchy transmission fidelity (DHTF) of 300 different groups of experienced agents transplanted into 300 different naive populations. Higher values mean that the naive agents learned a dominance hierarchy similar to the one in the original experienced population. Circles indicate median DHTF while boxes indicate the two middle quartiles.**

*distance* (DHD) as the portion of dominance relationships that have the same polarity in both dominance hierarchies:

$$\mathrm{DHD}(\mathcal{H}, \mathcal{H}') = 1 - \frac{2 \cdot |E(\mathcal{H}) \cap E(\mathcal{H}')|}{N(N-1)} \tag{1}$$

Where $E(\mathcal{H})$ represents the set of directed edges in the dominance hierarchy as ordered pairs.

To restrict our measurement only to dominance relationships between pairs of naive agents, we define a *restricted dominance hierarchy distance* (RDHD), which is a distance measure similar to DHD, except that only dominance relationships between a specified set $I$ of agent indices are sampled for their polarity:

$$\mathrm{RDHD}(\mathcal{H}, \mathcal{H}', I) = 1 - \frac{2 \cdot |E(\mathcal{H}) \cap E(\mathcal{H}') \cap (I \times I)|}{|I|(|I|-1)} \tag{2}$$

Finally, we define the *dominance hierarchy transmission fidelity* as the complement to RDHD, where the distance is restricted to naive agents only:

*Definition 6.1 (Dominance hierarchy transmission fidelity).* The *dominance hierarchy transmission fidelity* (DHTF) between two dominance hierarchies $\mathcal{H}$ and $\mathcal{H}'$, with a common set of naive agents indexed by $I_n$, is a measure of how similar the dominance relationships are between the naive agents in the two different hierarchies:

$$\mathrm{DHTF}(\mathcal{H}, \mathcal{H}') = 1 - \mathrm{RDHD}(\mathcal{H}, \mathcal{H}', I_n) \tag{3}$$

$$= \frac{2 \cdot |E(\mathcal{H}) \cap E(\mathcal{H}') \cap (I_n \times I_n)|}{|I_n|(|I_n|-1)} \tag{4}$$

In Figure 8 we show the distribution of the dominance hierarchy transmission fidelities for different values of $K$. A mere $K = 2$ experienced agents are able to teach 4 naive agents their original dominance hierarchy with 80% median fidelity.

This transmission 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.

## 7 DISCUSSION AND FUTURE WORK

This paper thoroughly investigates dominance hierarchies in the context of multiagent RL. It first provides a formal definition and a game-theoretical representation of dominance hierarchies. It then introduces the phenomenon of dominance hierarchies in its most simplified form using the Chicken Coop environment, stripping away many of the complexities found in dominance hierarchies in biological life. Finally, it demonstrates how populations of RL agents can invent, learn, enforce, and transmit dominance hierarchies.

Concerning the empirical evaluation, this research was conducted using PPO agents with a single learning regime. This intentional decision was made, in order to mitigate any confounding effects from using a variety of learning mechanisms. However, future research could explore how different RL algorithms learn and adapt dominance hierarchies, as well as the effects of different learning parameters on the emergence and stability of these hierarchies.

Preliminary experiments with higher learning rates result in unstable dominance hierarchies, where agents seem to converge strongly to a certain rank for some number of generations, only to deviate from it and converge to a different rank. This is another unique behavior that exists in natural environments, where it is known as *rank change* [64, 71] or *dynamic stability* [10].

We would like to suggest several extensions to the Chicken Coop environment that may provide additional insights into MARL systems:

(1) **The Handicap Principle [90]:** Enhancing the model with a cost for being dominant. In this environment, an agent always gets more reward by being at a high rank rather than at a low rank. In animal and human societies, a high rank carries sacrifices and risks that may outweigh the benefits of the high rank.
(2) **Multiple Hierarchies:** In the Chicken Coop environment, all the agents participate in one dominance hierarchy that includes all of the agents. In animal and human societies, individuals participate in multiple dominance hierarchies for multiple groups of other individuals, holding different ranks in each one.
(3) **Societies with Mixed Incentives:** In this environment, the agents are fully dedicated to making decisions on how aggressive they'll be to other agents. In animal and human societies, individuals in a group cooperate towards shared goals in addition to their internal dominance struggles.
(4) **Opponent Shaping (OS):** Algorithms such as LOLA [28] and M-FOS [50] enable agents to take the learning process of the other agents into account in their own policy gradients. When applied to an environment that supports dominance hierarchies, they may promote second-order dominance-seeking strategies, e.g., they may prompt agent $i$ to consider how it can behave as to encourage agent $j$ to place agent $i$ at a high rank in the dominance hierarchy.

We propose that the study of dominance hierarchies between RL agents may be combined with research on large language models

(LLMs) and multimodal foundation models (MFMs), which have gained immense popularity due to their robustness in solving a wide variety of problems [47, 87]. Tufano et al. [78] and similar work [79, 86] present an innovative way to use LLMs: instead of sending queries to a single LLM, users converse with a group of LLM-based agents, organized in human-inspired hierarchies such as manager-workers and student-assistant-expert. The back-and-forth interaction between multiple LLM agents is effective at encouraging diversity [49], factuality, and reasoning capability [23], when compared to a single LLM-based agent. We suggest that the augmenting these LLM-based agents with RL or OS algorithms, specifically to maintain the social relationships between the agents, may result in hierarchies that are similar to those observed in biological life.

When humans work on problems as a group, we balance in-group intrigues with external pressures to provide good results. We suggest that this interplay between individual needs and group needs may play a crucial role in the success of our collective intelligence. We hypothesize that MFM-based agents operating under similar conditions may make real-world decisions in a way that is more 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] Alberto Alesina and Eliana La Ferrara. 2005. Ethnic diversity and economic performance. *Journal of economic literature* 43, 3 (2005), 762–800.
[2] Cameron Anderson and Gavin J Kilduff. 2009. The pursuit of status in social groups. *Current Directions in Psychological Science* 18, 5 (2009), 295–298.
[3] Stéphane Aroca-Ouellette, Miguel Aroca-Ouellette, Upasana Biswas, Katharina Kann, and Alessandro Roncone. 2023. Hierarchical Reinforcement Learning for Ad Hoc Teaming. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 2337–2339.
[4] Robert Axelrod. 1986. An evolutionary approach to norms. *American political science review* 80, 4 (1986), 1095–1111.
[5] Michiel A. Bakker, Richard Everett, Laura Weidinger, Iason Gabriel, William S. Isaac, Joel Z. Leibo, and Edward Hughes. 2021. Modelling Cooperation in Network Games with Spatio-Temporal Complexity. arXiv:2102.06911 [cs.MA]
[6] 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.
[7] Samuel Bowles. 2009. Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors? *Science* 324, 5932 (2009), 1293–1298.
[8] Robert Boyd and Peter J Richerson. 2009. Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 1533 (2009), 3281–3288.
[9] Ivan D Chase, Costanza Bartolomeo, and Lee A Dugatkin. 1994. Aggressive interactions and inter-contest interval: how long do winners keep winning?. *Animal Behaviour* 48, 2 (1994), 393–400.
[10] 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.
[11] 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.
[12] Tian Chen Zeng, Joey T Cheng, and Joseph Henrich. 2022. Dominance in humans. *Philosophical Transactions of the Royal Society B* 377, 1845 (2022), 20200451.
[13] Vincent Conitzer. 2006. *Computational aspects of preference aggregation*. Ph.D. Dissertation. Carnegie Mellon University.
[14] Vincent Conitzer and Caspar Oesterheld. 2023. Foundations of cooperative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 15359–15367.
[15] Iain D Couzin. 2009. Collective cognition in animal groups. *Trends in cognitive sciences* 13, 1 (2009), 36–43.
[16] 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.
[17] Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. 2021. Cooperative AI: machines must learn to find common ground. *Nature* 593, 7857 (2021), 33–36.
[18] 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]
[19] 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).
[20] Han De Vries. 1995. An improved test of linearity in dominance hierarchies containing unknown or tied relationships. *Animal Behaviour* 50, 5 (1995), 1375–1389.
[21] Han De Vries. 1998. Finding a dominance order most consistent with a linear hierarchy: a new procedure and review. *Animal behaviour* 55, 4 (1998), 827–843.
[22] Reinhard Diestel. 2017. *Graph Theory* (5 ed.). Springer.
[23] Yilun Du, Shuang Li, Antonio Torralba, Joshua B Tenenbaum, and Igor Mordatch. 2023. Improving Factuality and Reasoning in Language Models through Multiagent Debate. *arXiv preprint arXiv:2305.14325* (2023).
[24] 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 preprint arXiv:2110.11404* (2021).
[25] Lee Alan Dugatkin. 1997. Winner and loser effects and the structure of dominance hierarchies. *Behavioral Ecology* 8, 6 (1997), 583–587.
[26] Lee Alan Dugatkin. 2001. Bystander effects and the structure of dominance hierarchies. *Behavioral Ecology* 12, 3 (2001), 348–352.
[27] 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.
[28] Jakob Foerster, Richard Y. Chen, Maruan Al-Shedivat, Shimon Whiteson, Pieter Abbeel, and Igor Mordatch. 2018. Learning with Opponent-Learning Awareness. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems* (Stockholm, Sweden) *(AAMAS '18)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 122–130.
[29] Sergey Gavrilets, Edgar A Duéñez-Guzmán, and Michael D Vose. 2008. Dynamics of alliance formation and the egalitarian revolution. *PLoS One* 3, 10 (2008), e3293.
[30] Jane Goodall. 1986. *The Chimpanzees of Gombe: Patterns of Behavior*. Belknap Press of Harvard University Press. https://books.google.co.il/books?id=eloQAQAAMAAJ
[31] Deborah M Gordon. 1996. The organization of work in social insect colonies. *Nature* 380, 6570 (1996), 121–124.
[32] Sebastian Grueneisen and Michael Tomasello. 2017. Children coordinate in a recurrent social dilemma by taking turns and along dominance asymmetries. *Developmental psychology* 53, 2 (2017), 265.
[33] 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.
[34] Eric A Hansen, Daniel S Bernstein, and Shlomo Zilberstein. 2004. Dynamic programming for partially observable stochastic games. In *AAAI*, Vol. 4. 709–715.
[35] Charlotte K Hemelrijk. 1996. Dominance interactions, spatial dynamics and emergent reciprocity in a virtual world. In *Proceedings of the fourth international conference on simulation of adaptive behavior*, Vol. 4. Citeseer, 545–552.
[36] Charlotte K Hemelrijk. 2000. Social phenomena emerging by self-organisation in a competitive, virtual world ("DomWorld"). In *Learning to behave Workshop II: Internalising knowledge. Ieper, Belgium*. 11–19.
[37] Charlotte K Hemelrijk, Peter M Kappeler, and Ivan Puga-Gonzalez. 2017. The self-organization of social complexity in group-living animals: Lessons from the DomWorld model. In *Advances in the Study of Behavior*. Academic Press, 361–405.
[38] S Peter Henazi and Louise Barrett. 1999. The value of grooming to female primates. *Primates* 40 (1999), 47–59.
[39] Yuko Ishiwaka, Xiao S Zeng, Shun Ogawa, Donovan Michael Westwater, Tadayuki Tone, and Masaki Nakada. 2022. DeepFoids: Adaptive Bio-Inspired Fish Simulation with Deep Reinforcement Learning. *Advances in Neural Information*

*Processing Systems* 35 (2022), 18377–18389.
[40] Geoffrey J Iverson and Donald Stone Sade. 1990. Statistical issues in the analysis of dominance hierarchies in animal societies. *Journal of Quantitative Anthropology* 2, 1 (1990), 61–83.
[41] 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.
[42] Keith Johnstone. 1981. *Impro: Improvisation and the Theatre*. Routledge. 33–74 pages.
[43] Elon Kohlberg and Jean-Francois Mertens. 1986. On the strategic stability of equilibria. *Econometrica: Journal of the Econometric Society* (1986), 1003–1037.
[44] Sarit Kraus. 1997. Negotiation and cooperation in multi-agent environments. *Artificial Intelligence* 94, 1 (1997), 79–97. https://doi.org/10.1016/S0004-3702(97)00025-8 Economic Principles of Multi-Agent Systems.
[45] 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.
[46] Olof Leimar. 2021. The evolution of social dominance through reinforcement learning. *The American Naturalist* 197, 5 (2021), 560–575.
[47] 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]
[48] Eric Liang, Richard Liaw, Philipp Moritz, Robert Nishihara, Roy Fox, Ken Goldberg, Joseph E. Gonzalez, Michael I. Jordan, and Ion Stoica. 2018. RLlib: Abstractions for Distributed Reinforcement Learning. arXiv:1712.09381 [cs.AI]
[49] Tian Liang, Zhiwei He, Wenxiang Jiao, Xing Wang, Yan Wang, Rui Wang, Yujiu Yang, Zhaopeng Tu, and Shuming Shi. 2023. Encouraging Divergent Thinking in Large Language Models through Multi-Agent Debate. *arXiv preprint arXiv:2305.19118* (2023).
[50] 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.
[51] 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.
[52] Brenda McCowan, Jessica Vandeleest, Krishna Balasubramaniam, Fushing Hsieh, Amy Nathman, and Brianne Beisner. 2022. Measuring dominance certainty and assessing its impact on individual and societal health in a nonhuman primate model: a network approach. *Philosophical Transactions of the Royal Society B* 377, 1845 (2022), 20200438.
[53] Henry Mintzberg. 1989. *The structuring of organizations*. Springer.
[54] Igor Mordatch and Pieter Abbeel. 2018. Emergence of grounded compositional language in multi-agent populations. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 32.
[55] Michael Muthukrishna and Joseph Henrich. 2016. Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371, 1690 (2016).
[56] Simon Parsons, Juan A Rodriguez-Aguilar, and Mark Klein. 2011. Auctions and bidding: A guide for computer scientists. *ACM Computing Surveys (CSUR)* 43, 2 (2011), 1–59.
[57] Wojtek Przepiorka, Charlotte Rutten, Vincent Buskens, and Aron Szekely. 2020. How dominance hierarchies emerge from conflict: A game theoretic model and experimental evidence. *Social science research* 86 (2020), 102393.
[58] Ram Rachum, Yonatan Nakar, Bill Tomlinson, Nitay Alon, and Reuth Mirsky. 2024. Emergent Dominance Hierarchies in Reinforcement Learning Agents. In *Proceedings of the 2024 International Conference on Autonomous Agents and Multiagent Systems* (Auckland, New Zealand) *(AAMAS '24)*. International Foundation for Autonomous Agents and Multiagent Systems.
[59] David Radke, Kate Larson, Tim Brecht, and Kyle Tilbury. 2023. Towards a Better Understanding of Learning with Multiagent Teams. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. ijcai.org, 271–279. https://doi.org/10.24963/ijcai.2023/31
[60] Sarvapali D Ramchurn, Claudio Mezzetti, Andrea Giovannucci, Juan A Rodriguez-Aguilar, Rajdeep K Dash, and Nicholas R Jennings. 2009. Trust-based mechanisms for robust and efficient task allocation in the presence of execution uncertainty. *Journal of Artificial Intelligence Research* 35 (2009), 119–159.
[61] Anatol Rapoport. 1949. Outline of a probabilistic approach to animal sociology: I. *The bulletin of mathematical biophysics* 11 (1949), 183–196.
[62] 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
[63] Byron Reeves and Clifford Nass. 1996. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK* 10, 10 (1996).
[64] 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.
[65] Nathan J Sanders and Deborah M Gordon. 2003. Resource-dependent interactions and the organization of desert ant communities. *Ecology* 84, 4 (2003), 1024–1031.
[66] Robert M Sapolsky. 2005. The influence of social hierarchy on primate health. *science* 308, 5722 (2005), 648–652.
[67] Thorleif Schjelderup-Ebbe. 1922. Beiträge zur Sozialpsychologie des Haushuhns.[Observation on the social psychology of domestic fowls.]. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane. Abt. 1. Zeitschrift für Psychologie* 88 (1922), 225.
[68] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
[69] J Maynard Smith and George R Price. 1973. The logic of animal conflict. *Nature* 246, 5427 (1973), 15–18.
[70] 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.
[71] 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.
[72] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
[73] Cultural General Intelligence Team, Avishkar Bhoopchand, Bethanie Brownfield, Adrian Collister, Agustin Dal Lago, Ashley Edwards, Richard Everett, Alexandre Frechette, Yanko Gitahy Oliveira, Edward Hughes, et al. 2022. Learning Robust Real-Time Cultural Transmission without Human Data. *arXiv preprint arXiv:2203.00715* (2022).
[74] Larissa Z Tiedens and Alison R Fragale. 2003. Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of personality and social psychology* 84, 3 (2003), 558.
[75] Bill Tomlinson. 2009. A Proximate Mechanism for Communities of Agents to Commemorate Long Dead Ancestors. *Journal of Artificial Societies and Social Simulation* 12, 1 (2009), 7. https://www.jasss.org/12/1/7.html
[76] Bill Tomlinson, Bruce Blumberg, and Bradley Rhodes. 2000. How is an agent like a wolf?: Dominance and submission in multi-agent systems. In *International ICSC Symposium on Multi-Agents and Mobile Agents in Virtual Organizations and E-Commerce (MAMA'2000)*.
[77] William Michael Tomlinson. 2002. *Synthetic social relationships for computational entities*. Ph.D. Dissertation. Massachusetts Institute of Technology.
[78] Michele Tufano, Anisha Agarwal, Jinu Jang, Roshanak Zilouchian Moghaddam, and Neel Sundaresan. 2024. AutoDev: Automated AI-Driven Development. arXiv:2403.08299 [cs.SE]
[79] 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]
[80] 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
[81] Christopher R von Rueden, Daniel Redhead, Rick O'Gorman, Hillard Kaplan, and Michael Gurven. 2019. The dynamics of men's cooperation and social status in a small-scale society. *Proceedings of the Royal Society B* 286, 1908 (2019), 20191367.
[82] John R Watson. 1970. Dominance-subordination in caged groups of house sparrows. *The Wilson Bulletin* (1970), 268–278.
[83] 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.
[84] G Wittemyer and Wayne M Getz. 2007. Hierarchical dominance structure and social organization in African elephants, Loxodonta africana. *Animal Behaviour* 73, 4 (2007), 671–681.
[85] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *Science* 330, 6004 (2010), 686–688.
[86] 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]
[87] 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]

[88] Yue Xu, Zengde Deng, Mengdi Wang, Wenjun Xu, Anthony Man-Cho So, and Shuguang Cui. 2020. Voting-Based Multi-Agent Reinforcement Learning for Intelligent IoT. arXiv:1907.01385 [cs.LG]

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

[90] Amotz Zahavi and Avishag Zahavi. 1999. *The handicap principle: A missing piece of Darwin's puzzle.* Oxford University Press.

[91] 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.