FRIteza

Predator–prey survival pressure

Rafaelle Lacraz, Titouan Steyer

Collective Behaviour Course, Research Seminar Report

December 2025

Supervisor: Iztok Lebar Bajec

This project explores how collective behaviors can emerge in predator–prey systems driven by minimal survival incentives. Following the framework of Li et al. [1], we implement a 2D continuous environment with multiple prey and a few predators, each governed by physical dynamics and subject to active and passive forces. The reward function is deliberately simple: predators gain a reward when catching prey, and prey receive a penalty when caught. Using public code as a starting point, we reproduce baseline behavior and visualize system evolution under random motion. Quantitative metrics such as the Degree of Sparsity (DoS) and Degree of Alignment (DoA) allow us to characterize emergent group dynamics. This report serves as a progress checkpoint toward building an adaptive reinforcement learning model that captures realistic group behavior.

1 Introduction

Collective behaviours such as flocking or coordinated pursuit are commonly observed in biological systems and have long been studied through simple interaction-based models. These approaches successfully reproduce group-level patterns, but rely on predefined social rules and do not explain how collective behaviour may emerge autonomously.

Multi-agent reinforcement learning (MARL) provides an alternative framework, in which coordinated behaviours can arise from individual objectives without explicitly encoding social interactions. Recent studies have shown that even minimal survival-based rewards may lead to swarm-like dynamics, raising questions about the respective roles of learning, incentives and environmental constraints in multi-agent reinforcement learning settings [2].

Li et al. [1] proposed a minimalist predator–prey system driven solely by survival pressure, where predators are rewarded for captures and prey are penalised when caught. Despite the simplicity of the reward structure, collective organisation emerges in a continuous two-dimensional environment with periodic boundary conditions.

In this work, we reproduce this framework and analyse how environmental structure influences emergent collective behaviour. Using quantitative metrics, we compare periodic and bounded environments to assess how physical constraints and wall interactions shape prey organisation and predator performance.

2 Methods

2.1 Environment and Implementation

To reproduce the conditions described in Li et al. [1], we implemented a continuous two-dimensional square environment in which predators and prey move and interact. Two types of boundary conditions are supported:

Periodic boundaries (torus). When an agent crosses one side of the domain, it reappears on the opposite side with the same velocity. This configuration removes edge effects and matches the setup of the original paper.

Solid boundaries (walls). Agents remain confined inside the square and bounce when reaching the boundary. Collisions are modeled using a repulsive spring force ($k = 50$ N/m), preventing agents from crossing the walls.

At the beginning of each simulation, predators ($n_0 = 3$) and prey ($n_1 = 10$) are placed at random positions with random headings. All agents are represented as disks with mass $m = 1$ kg. Predators and prey differ slightly in their locomotion capabilities, with maximum speeds of 0.35 m/s and 0.30 m/s respectively. Each agent receives two bounded control inputs: a forward acceleration and a rotational command.

The physics of the environment follow Newtonian dynamics and include:

- **active forces** generated by the agent's own propulsion,
- **viscous drag**, proportional to velocity,
- **collision forces** with other agents or with boundaries, based on Hooke's law.

For prey, we optionally implement Couzin-type social rules (repulsion, alignment, attraction) to test deterministic collective behaviour before applying reinforcement learning.

Overall, the implementation remains faithful to the physical model described in Li et al., while allowing different environmental configurations to be studied.

2.2 Collective Behavior Metrics

In addition to visual inspection, we use quantitative metrics to assess emergent behaviours within agent populations. Following Li et al. [1], we compute the *Degree of Sparsity* (DoS) and the *Degree of Alignment* (DoA), which characterise spatial cohesion and local heading synchronisation.

Degree of Sparsity (DoS). $\text{DoS} \in [0, 1]$ quantifies spatial dispersion by averaging the normalised distance to the nearest neighbour among conspecifics:

$$\text{DoS} = \frac{1}{TND} \sum_{t=1}^{T} \sum_{j=1}^{N} \left\| \mathbf{x}_j(t) - \mathbf{x}_{k(j)}(t) \right\|$$

where $\mathbf{x}_j(t) \in \mathbb{R}^2$ is the position of agent j at time t, $k(j)$ its nearest neighbour, T the episode length, N the number of agents, and D the environment size. Low DoS values indicate cohesive groups, while high values correspond to dispersed configurations.

Degree of Alignment (DoA). $\text{DoA} \in [0, 1]$ measures local alignment between neighbouring agents and is defined as:

$$\text{DoA} = \frac{1}{2TN} \sum_{t=1}^{T} \sum_{j=1}^{N} \left\| \mathbf{h}_j(t) - \mathbf{h}_{k(j)}(t) \right\|$$

where $\mathbf{h}_j(t)$ is the unit heading vector of agent j. Higher DoA values correspond to stronger local alignment.

2.3 Reward Function

We adopt a minimal reward structure consistent with Li et al. [1]:

- predators receive $+1$ when catching a prey,
- prey receive -1 when caught.

No additional incentives related to cohesion or alignment are introduced. We used a public GitHub repository [4] as a starting point for part of the implementation.

Author contributions. Rafaelle Lacraz implemented and trained the predator MADDPG model. Titouan Steyer developed the bounded environments and simulation logic. Both authors contributed to analysis and manuscript writing.

3 Results

3.1 Baseline observations without learning

We first analyse predator–prey interactions in a confined environment with solid boundaries, in the absence of learning. All agents follow random motion policies, providing a baseline reference without structured interactions.

In this setting, prey remain largely dispersed and do not form persistent collective structures. Although temporary local aggregations may arise due to spatial proximity or wall reflections, these configurations are unstable and quickly dissolve. Quantitatively, the Degree of Sparsity (DoS) remains rather high throughout the simulation, while the Degree of Alignment (DoA) stays close to zero, indicating the absence of sustained collective organisation.

Similar behaviour is observed in a periodic (torus) environment, confirming that random motion alone is insufficient to generate collective dynamics, regardless of boundary conditions.

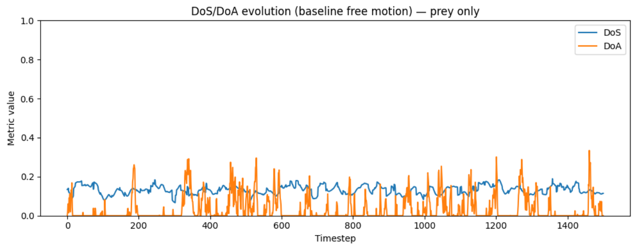

Figure 1: Evolution of DoS and DoA for prey under random-motion policies in a periodic environment.

3.2 Deterministic rules and confined environments

We next investigate the effect of deterministic interaction rules in a confined environment. Prey dynamics are governed by Couzin-type social rules, while predators actively pursue prey.

Under these conditions, prey rapidly form cohesive and persistent clusters, characterised by low spatial dispersion and strong local alignment. The presence of walls reinforces these structures by constraining agent motion and promoting repeated interactions. Predators tend to focus their pursuit on dense prey groups, leading to frequent capture events.

In contrast, in a periodic environment, similar deterministic rules lead to collective motion without confinement effects, highlighting how environmental boundaries modulate, but do not suppress rule based collective behaviour.

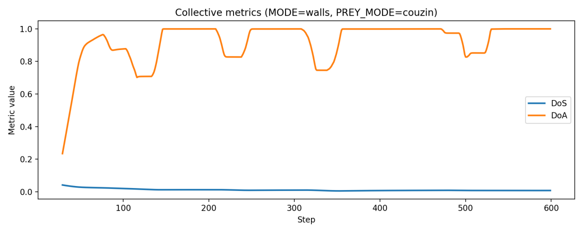

Figure 2: Evolution of the DoS and DoA for prey under deterministic rules and confined environment.

3.3 Reinforcement learning experiments

Finally, we examine collective dynamics emerging from reinforcement learning in the confined environment. Predator agents are trained using a MADDPG framework with minimal survival-based rewards, without explicit communication or social incentives, following the philosophy introduced by Li et al. [1].

We successfully establish a stable training pipeline and observe clear learning signals compared to the random baseline. Predators exhibit improved pursuit behaviour, and learning consistently drives the system away from disordered motion toward compact collective states. In particular, the Degree of Sparsity decreases markedly, reaching values comparable to, or even lower than, those obtained under deterministic interaction rules, indicating that spatial cohesion emerges robustly despite the absence of explicit social rewards.

However, the collective regimes obtained differ from those reported by Li et al. [1]. While alignment increases during early stages of learning, revealing transient coherent motion, it is not stabilised over longer timescales. We attribute this primarily to stronger physical constraints in the confined environment: action regularisation and wall-related penalties significantly limit exploration and favour conservative motion strategies over aggressive pursuit.

As a result, convergence toward highly aligned and persistent collective hunting behaviours is hindered. Qualitative inspection of simulations suggests that confined environments promote rotating or locally coordinated patterns that are not fully captured by global alignment metrics such as the Degree of Alignment. In contrast, alignment is more stable in periodic environments, confirming the central role of boundary conditions and penalty design in shaping learned collective regimes.

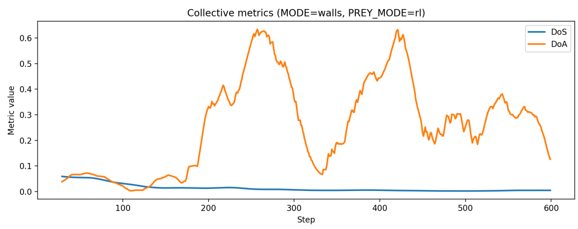

Figure 3: Evolution of the DoS and DoA for prey under confined environment with reinforcement learning.

4 Discussion

Our results show that reproducing the collective regimes reported by Li et al. [1] becomes challenging when extending the model to bounded environments with stronger physical constraints. Baseline simulations without learning behave as expected, with high DoS and negligible alignment, while both deterministic and learning-based settings clearly depart from this null collective regime.

In the deterministic Couzin-based setting, collective organisation is strong and stable. As shown in Figure 2, the Degree of Sparsity rapidly drops toward zero and the Degree of Alignment remains close to one, corresponding to a compact and strongly aligned prey group. This behaviour is consistent with the imposed interaction rules and confirms that the metrics correctly capture collective cohesion and alignment.

In the reinforcement learning setting, results are more contrasted. The Degree of Sparsity reaches very low values (around 0.1, Figure 3), comparable or even lower than those reported by Li et al. [1]. This increased compactness can be explained by the stronger penalty terms used in our model, which constrain motion more tightly and force prey to remain closer together. From a spatial cohesion perspective, learning therefore succeeds in reproducing, and even amplifying, one key aspect of the collective regime.

In contrast, alignment is not sustained. During the early phase of training (roughly the first 200 steps), the Degree of Alignment increases, indicating the emergence of coherent motion. However, this alignment subsequently degrades as training progresses, and long-term learning does not stabilise a strongly aligned collective state despite maintaining low sparsity. Qualitative inspection of simulations during this early phase reveals transient but coherent collective motion, consistent with the temporary rise in DoA, although these patterns are not sustained.

We explored several variations of the experimental setup, including changes in agent speed, world size, and physical scaling parameters, as well as an alternative boundary model based on purely elastic wall reflections.

None of these modifications led to a consistent improvement in long-term alignment. Similar difficulties in stabilising cooperative or aligned behaviours under sparse rewards have also been reported in other multi-agent predator–prey learning settings [3].

This suggests that the collective regimes observed by Li et al. [1] are highly sensitive to the balance between reward sparsity, physical constraints, and boundary conditions.

References

[1] Zhengyang Li, Emanuele Cuccoli, Marco Villani, and Andrea Cavagna. Emergence of collective behaviour in coevolving predator–prey systems. *New Journal of Physics*, 25(9):093028, 2023.

[2] Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*, 2017.

[3] Frederik Ritz, Daniel Ratke, Tho Phan, Laurent Belzner, and Claudia Linnhoff-Popien. A sustainable ecosystem through emergent cooperation in multi-agent reinforcement learning. In *Proceedings of the Artificial Life Conference*, 2021.

[4] xxnnnnn. Predatorprey_rl_reproduction. `https://github.com/xxnnnnn/PredatorPrey_RL_Reproduction`, 2024.