Swarming behaviour in predator-prey model

Ariana Kržan, Tina Brdnik, and Vito Levstik

Collective behaviour course research seminar report

Collective animal behaviour, especially swarming in predator-prey dynamics, offers insights into survival strategies that emerge under evolutionary pressures. This report outlines the initial objectives and foundational concepts for simulating predator-prey. Inspired by previous work, we examine how survival pressures can drive emergent group behaviours in prey through reinforcement learning. We begin with an overview of related work, from classic rule-based models to more recent reinforcement learning approaches, highlighting advances that allow agents to adapt to changing environments. Our primary objective is to recreate a reinforcement learning-based model where predator-prey interactions lead to swarming and evasion behaviours. The model will then extend to include environmental obstacles and an additional species, enabling us to investigate the interplay between interspecies interactions and survival strategies.

Simulation | swarming behaviour | predator | prey

The sudden emergence of swarming behaviours in animals is one of the most striking examples of collective animal behaviour. These behaviours have been extensively studied for their implications for the evolution of cooperation, social cognition and predator–prey dynamics[1]. Swarming, which appears in many different species like starlings, herrings, and locusts, has been linked to several benefits including enhanced foraging efficiency, improved mating success, and distributed problem-solving abilities. Furthermore, they are hypothesized to help with improving group vigilance, reducing the chance of being encountered by predators, diluting an individual's risk of being attacked, enabling an active defence against predators and reducing predator attack efficiency by confusing the predator. [2].

In this project we will be taking inspiration from the work of Li et al. (2023) and Olson et al. (2013) to explore how survival pressures can drive the emergence of swarming behaviour. The first goal will be to create a realistic simulation where both prey and predators learn to adapt through reinforcement learning based on their drive to survive. Modelling these interactions, we will observe how simple survival pressures can lead to evolution of more complex behaviours like flocking, swirling and edge predation.

Then, we will extend our research by evolving out existing model by introducing new environmental obstacles and new species to observe how interspecies interactions lead to new survival strategies.

Related Work

The modeling of swarming behavior has evolved from static rule-based frameworks to adaptive reinforcement learning (RL) models, with intermediate advances in topological and vision-based approaches enhancing realism.

Rule-Based Models. Early models like Aoki's Zone Model (1982) defined interaction zones—repulsion, alignment, and attraction—based on proximity [3]. Vicsek's Model (1995) and Reynolds' Boids Model (1987) added alignment, cohesion, and separation rules to simulate group dynamics [4][5]. However, these fixed-rule systems lack adaptability to dynamic environments.

Topological and Vision-Based Models. Topological models incorporated sensory constraints for more realism. Hemelrijk & Hildenbrandt (2008) introduced a perception model where agents respond only to neighbors within a variable radius [6]. Kunz & Hemelrijk (2012) added visual occlusion, simulating sensory limits [7]. While more realistic, these models remain static compared to adaptive RL approaches.

Learning-Based Models. Learning-based methods allow agents to adapt dynamically, producing emergent behaviors like flocking and evasion.

- **Olson et al. (2013):** Used genetic algorithms (GA) to model predator confusion, where prey evolved clustering to reduce predation risk [1].

- **Lowe et al. (2017):** Introduced MADDPG, an RL algorithm enabling agents to learn strategies in mixed cooperative and competitive environments [8].

This research explores the emergence of collective behaviors in predator-prey dynamics using reinforcement learning to simulate how survival pressures drive adaptive behaviors like swarming and evasion. By modeling multi-species interactions in a complex environment, the study extends current understanding of evolutionary survival strategies and the role of learning in shaping group dynamics.

Simulation | swarming behaviour | predator | prey

- **Li et al. (2023):** Demonstrated prey swarming in an RL model, where agents maximize survival rewards and evade predators adaptively [2].

Methods

Our proposed methodology aims to simulate swarming behaviours in a predator-prey environment using reinforcement learning (RL). We will define and test a RL-based model where agents, such as prey and predators interact within a two-dimensional space. The goal is to observe how different pressures and interactions influence collective behaviours like swarming, evasion, and strategic movement.

Environment Setup. The simulation will take place in a 2D environment with open and confined spaces. The confined space will have stiff boundaries, meaning that agents will bounce off them when they collide, whereas the open space will have periodic boundaries, meaning that agents will reappear on the opposite side when they cross the boundary. Such setup with periodic boundaries serves as an approximation of an infinite space, allowing agents to move freely without encountering physical borders. Later on, we wish to place random obstacles, which will be distributed across the space to create a complex and realistic setting that challenges the agents to adapt their movement and coordination.

We will apply the perception and action models from Li et al. (2023) [2] to guide agent interactions in the simulation.

- **Perception Model**: Each agent detects others only within a specified range and limited to a maximum number of nearby agents, simulating real-world sensory limitations.

- **Action Model**: Agents adjust their movement through forward propulsion and directional changes, governed by RL policies that optimize goals like survival and prey capture.

Agent Types and Behaviour.

- **Prey**: These agents aim to survive by avoiding predators and moving as a group.

- **Predators**: Predators are designed to pursue and catch prey.

- **New Species**: We may introduce a third type of agent, such as a neutral species, scavenger or competitor, which will have its own survival or resource-based objectives.

Reinforcement Learning Framework.

- **Algorithm**: We plan to use the Multi-Agent Deep Deterministic Policy Gradient (MADDPG) algorithm.

- **Reward Structure**:

  - **Prey**: Rewarded for survival over time, with penalties for being caught.

  - **Predators**: Rewarded for capturing prey, with penalties for colliding with obstacles.

  - **New Species**: Rewarded based on interactions like resource competition or cooperation with other agents.

- **Training Setup**: Agents will be trained through episodic simulations, allowing them to learn and adapt from each episode's interactions. We will vary conditions to observe how changes influence learned behaviours.

Proposed Methodology for Verification. To verify the behavior of our model, we will adopt the methodology described in Li et al. (2023) [2], utilizing two key metrics: the Degree of Alignment (DoA) and the Degree of Separation (DoS).

- **Degree of Sparsity (DoS):** This metric measures the spatial aggregation of agents, capturing how densely the agents cluster together. It is defined as:

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

where: $\mathbf{x}_j(t)$ is the position of the j-th agent at time step t, $\mathbf{x}_k(t)$ is the position of the nearest neighbor $k = \arg\min_k \|\mathbf{x}_j(t) - \mathbf{x}_k(t)\|$, T is the episode length, N is the total number of agents, and D is the maximum possible distance between two agents in the environment.

A smaller DoS value indicates denser clustering, while a value of 0 represents all agents aggregating at a single point.

- **Degree of Alignment (DoA):** This metric quantifies the alignment of the agents' headings, assessing how consistently agents move in the same direction. It is defined as:

$$\text{DoA} = \frac{1}{2TN} \sum_{t=1}^{T} \sum_{j=1}^{N} \|\mathbf{h}_j(t) + \mathbf{h}_k(t)\|$$

where: $\mathbf{h}_j(t)$ is the heading of the j-th agent at time step t, $\mathbf{h}_k(t)$ is the heading of the nearest neighbor of agent j (the same nearest neighbor as in the DoS calculation), T is the episode length, and N is the total number of agents.

Higher DoA values indicate stronger alignment in agent movement. It is important to note that the DoA measures local alignment between neighboring agents rather than the mean heading of the entire group, making it more suitable for detecting relative alignment within swarms.

By analyzing these metrics during and after training, we aim to verify whether our model reproduces swarming behaviors.

Results

In the initial phase of our project, we implemented a basic model where we created our environment with periodic borders and successfully populated it with agents which followed a reward system following the article. At this stage we only implemented active forces with fixed values. Our results looked promising.

Now we have improved the model by first adding passive forces. We also fine-tuned some parameters such as the agents' speed, size and passive forces. The passive forces include dragging forces, elastic forces between contacting agents, and elastic forces between agents and boundaries. Next we initiated active forces as random instead of fixed valued. That was crucial for our next step, since we were gonna build our reinforcment model based on them.

Lastly we added a RL component. The RL component was built to fine-tune the active forces a_f and a_f to enable the agents to move efficiently and maximize their results. The a_f component controls the agent's forward movement, aligned with its current heading, while the a_r component adjusts the agent's heading through rotational actions. However so far, after successfully training our model numerous times over 1000 episodes, the results were not great. Instead of the agents moving to maximize their reward, they end up moving in circles.
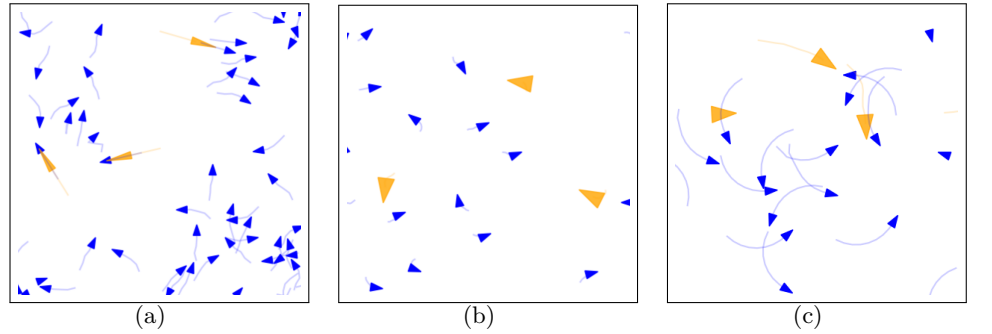
Figure 1. (a) Our model with no RL component, following the reward system and fixed active forces. (b) Our model with no RL component, with added fixed passive forces and random initialization of active forces (c) Our model with RL.

Recognizing these limitations, we decided to adopt the RL model from the Li et al. article [2]. Using their code, we were able to successfully run simulations of their model in our environment. This provided a functioning baseline for comparison and further experimentation. Preliminary results from this approach, however, did not fully replicate the swarming behaviors reported in their study. Our results show significant variability in both the Degree of Separation (DoS) and Degree of Alignment (DoA) across episodes, with no clear trends indicating improvement. The fluctuating

DoS suggests that agents fail to form stable clusters, while the inconsistent DoA values indicate a lack of coordinated alignment among agents. These results suggest that the model has not converged effectively, likely due to issues with the reward structure or suboptimal parameter tuning. Further refinement is needed to achieve consistent swarming behaviors.
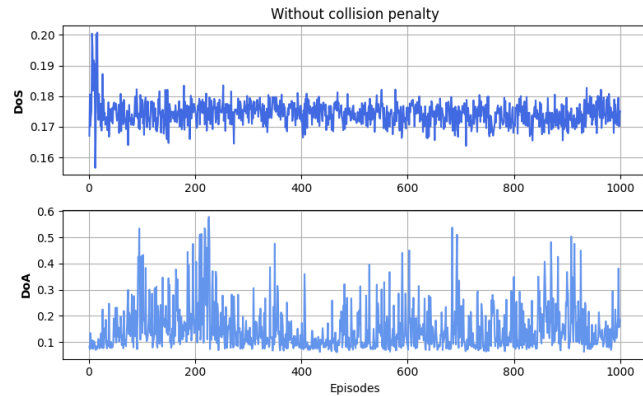
Figure 2. Degree of Alignment (DoA) and Degree of Separation (DoS) across episodes without collision penalty. The DoA remains highly variable, while the DoS shows no clear clustering trend.

Discussion

In this stage of the project, we successfully ran the code from the main article but encountered challenges with the agents' behavior. The DoA and DoS metrics remain highly variable, indicating that swarming behaviors have not yet emerged. These results suggest the need for further tuning of the reward structure and additional training.

Despite the setbacks, we are motivated to continue refining the model and addressing these issues. Our next steps are to improve the code to produce results similar to the article, introduce obstacles to the environment, and add a new species, which we have yet to finalize.

1. Olson RS, Hintze A, Dyer FC, Knoester DB, Adami C (2013) Predator confusion is sufficient to evolve swarming behaviour. *Journal of The Royal Society Interface* 10(85):20130305.
2. Li J, Li L, Zhao S (2023) Predator–prey survival pressure is sufficient to evolve swarming behaviors. *New Journal of Physics* 25(9):092001.
3. AOKI I (1982) A simulation study on the schooling mechanism in fish. *NIPPON SUISAN GAKKAISHI* 48(8):1081−1088.
4. Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O (1995) Novel type of phase transition in a system of self-driven particles. *Physical Review Letters* 75(6):1226−1229.
5. Reynolds CW (1987) Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.* 21(4).
6. Hemelrijk CK, Hildenbrandt H (2008) Self-organized shape and frontal density of fish schools. *Ethology* 114(3):245−254.
7. Kunz H, Hemelrijk CK (2012) Simulations of the social organization of large schools of fish whose perception is obstructed. *Applied Animal Behaviour Science* 138(3):142−151. Special Issue: Living In Large Groups.
8. Lowe R et al. (2017) Multi-agent actor-critic for mixed cooperative-competitive environments. *Proceedings of the 31st International Conference on Neural Information Processing Systems*.