Sharing is Not Needed: Modeling Animal Coordinated Hunting with Reinforcement Learning

Minglu Zhao[1], Ning Tang[1], Annya L. Dahmani[2], Yixin Zhu[1], Federico Rossano[3], Tao Gao[1,4]

[1]Department of Statistics, UCLA, [2]Department of Psychology, UCLA,
[3]Department of Cognitive Sciences, UCSD, [4]Department of Communication, UCLA

Motivation

Modeling

Human social intelligence:
Cooperation requires:
- Joint commitment
- Fairness
- Social punishment
- ...

?

Multi-agent reinforcement learning (MARL):
- Coordination achieved through sharing rewards

Is coordination modeled through sharing rewards robust?

Animal cognition

Chimpanzees have roles in coordinated hunting, indicating **cooperation**

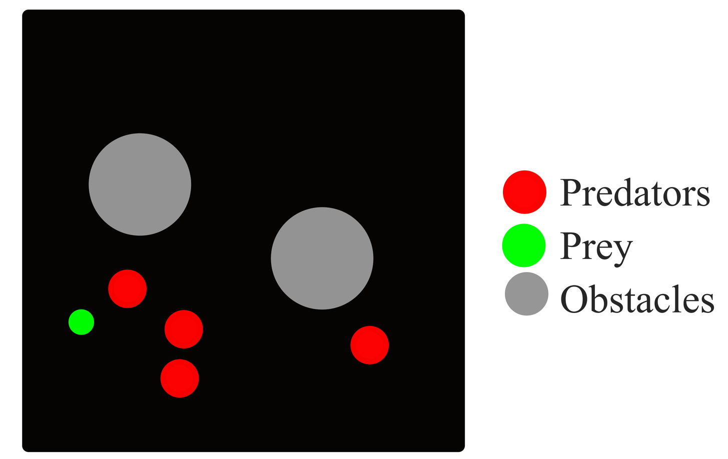
?

Chimpanzees coordinate to maximize **individual** benefits

Is sharing rewards necessary for coordination?

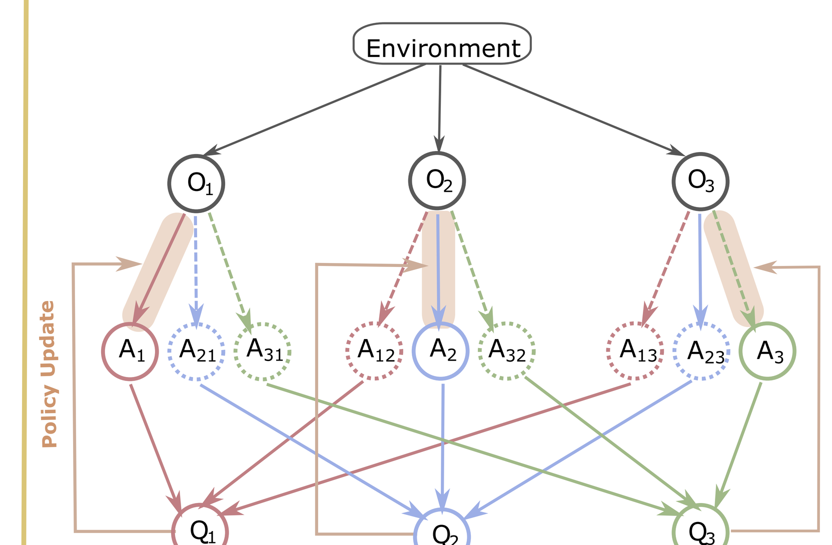
Methods

Task: Coordinated hunting



- Multiple predators chase one prey
- Predators have **individual** action costs
- Predators rewarded only after **kills**

Reinforcement learning model



Multi-agent Deep Deterministic Policy Gradient (MADDPG)

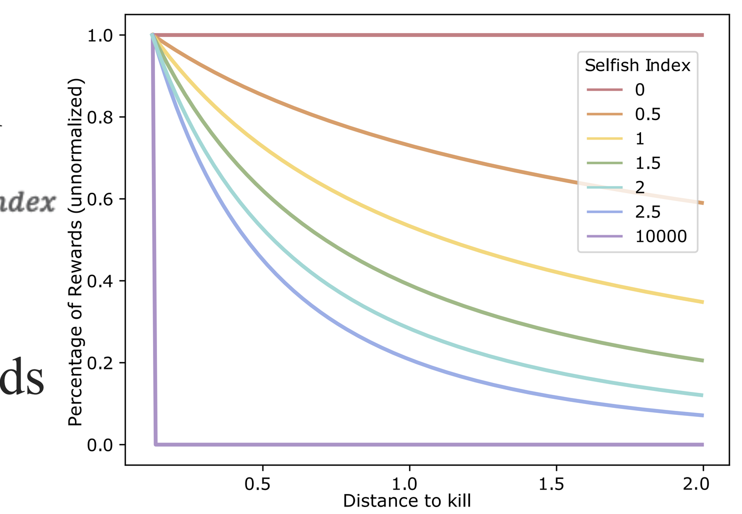
Experiment design

Manipulate evolutionarily important factors:

1. **Reward distribution:** through selfish index

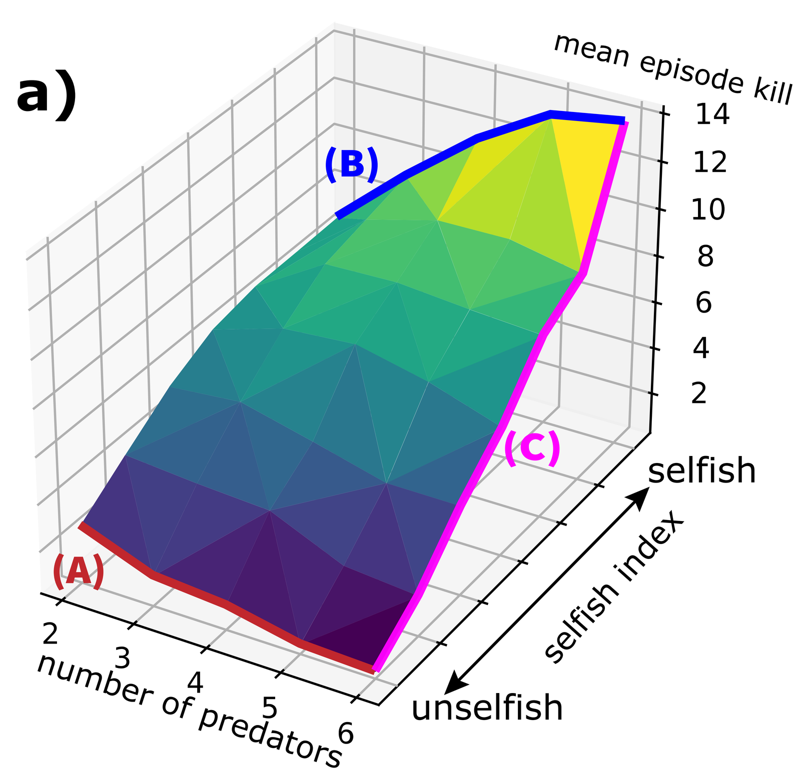
$$percentageOfRewards \propto (distanceToKill + 1 - k)^{-selfishIndex}$$

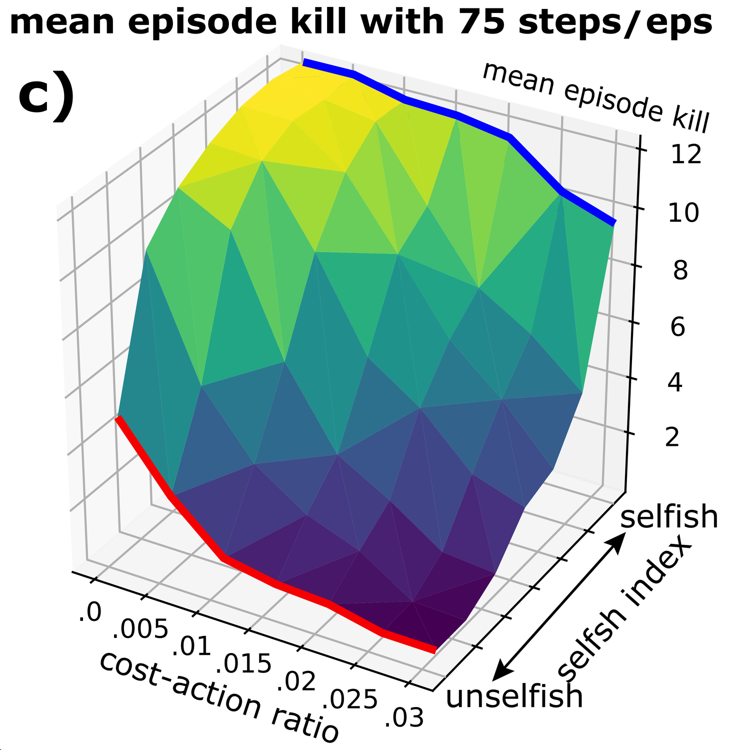
Selfish index = 0: unselfish, share rewards
Selfish index = inf: selfish, individual rewards



2. **Free-rider effects:** through action cost, proportional to action force exerted
3. **Hunting party size**: through number of predators
4. **Hunting difficulty**: through prey speed

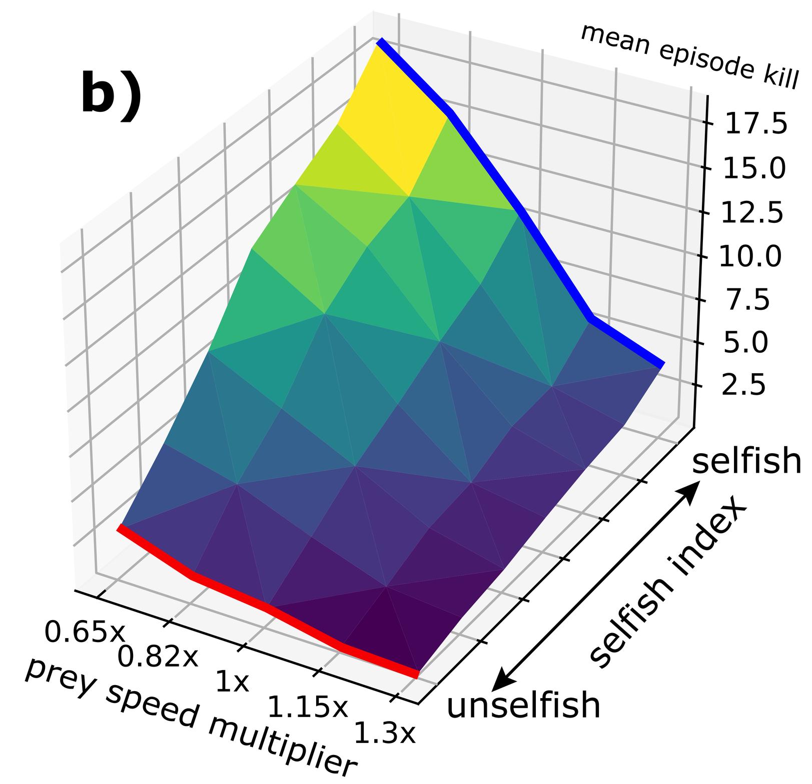
Main experiment results

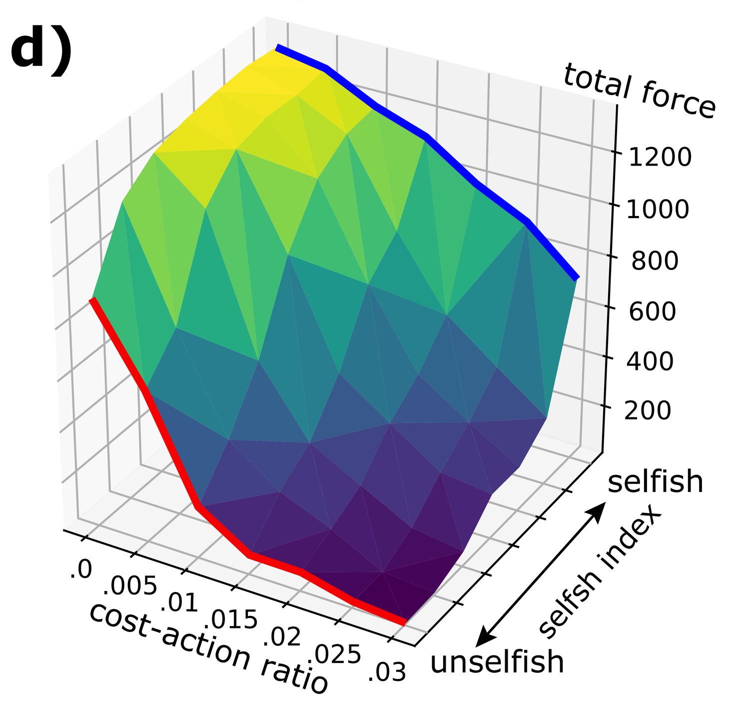




Findings

1. **Figure a:**
   - Performance of selfish agents increases linealy with the group size (line B).
   - Unselfish agents' performance remains the same or drops when having a larger group (line A);
   - More selfish, better performance (line C).
2. **Figure b:**
   - Prey speed increases → predators' performance decreases
3. **Figure c:**
   - Action cost increases → performance decreases
   - Increasing action costs hurts unselfish agents more than selfish agents
4. **Figure d:**
   - More selfish agents have their action force less sensitive to action costs. The most unselfish agents decide almost not to move at all when there is a small action cost
   - **Such a result strongly indicates the presence of the free-rider problem under the reward-sharing mechanism**

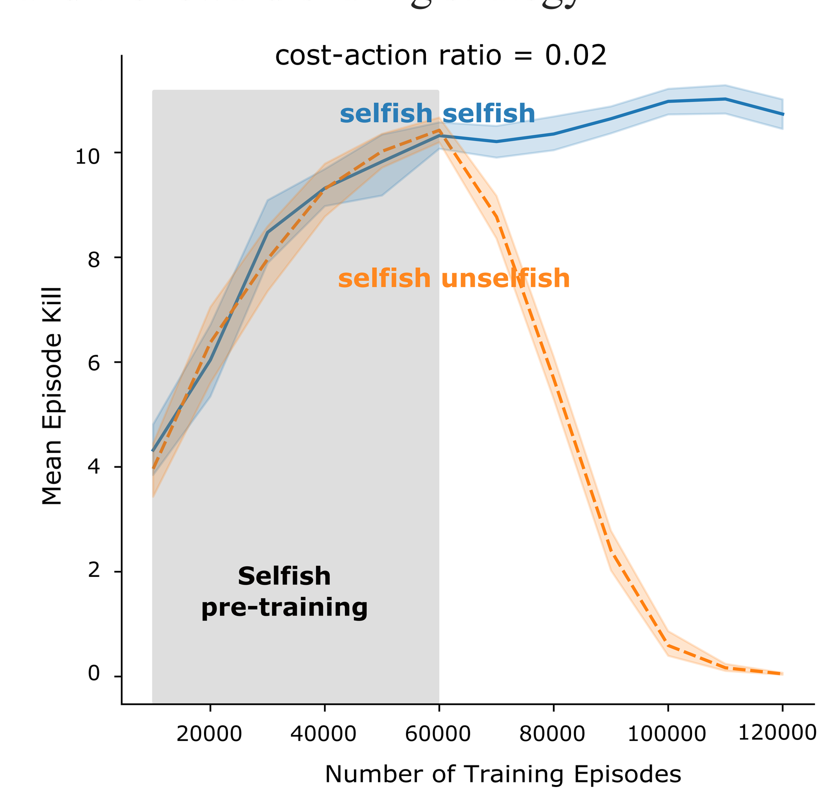




Equilibrium testing experiment

Motivation: From main results, sharing rewards cannot **generate** robust coordination. Then can it **maintain** successful coordination?

Methods: Use models of **already-coordinated** selfish agents and train them for **another around** with both the individual-reward strategy and the reward-sharing strategy



Findings

- Performance of agents pretrained with individual rewards for 60k episodes significantly decreased after training with shared rewards for another 60k episodes
- Successful coordination through sharing rewards is **not an equilibrium**, since all agents' policies deviate from it in further training.

Conclusions

Sharing rewards is **neither necessary nor sufficient** for modeling animal coordinated hunting with reinforcement learning:

Not necessary
- Models without any sharedness (selfish agents) achieve good training results

Not sufficient
- Free-rider problem in sharing rewards
- Unselfish agents' hunting performance plateaus at small group size
- Coordination through sharing rewards is not a Nash equilibrium

Animal cognition perspective:
Support the argument that chimpanzees can coordinate in hunting even only with selfish motivations

Computational perspective:
Structures involving developmental insights required to model human-like cooperation