The Influence of Social Display in Competitive Multiagent Learning

October 2014

Pedro Sequeira, Francisco S. Melo and Ana Paiva

In this work, we analyze the impact of simple **social signaling** mechanisms in the performance of learning agents in **competitive multiagent settings**. The social signals influence the total amount of reward received throughout time.

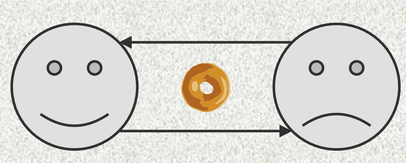
In a series of experiments, we measure the **combined performance** of agent populations according to distinct **selection paradigms**.

The results of our study show that agents learn to **collectively coordinate their feeding behavior** by trading-off immediate benefit for long-term social welfare.

The individuals with the **highest fitness** appear in the **most socially-aware** populations.

Social inequalities arise where the more fit agents benefit from their higher status being appreciated by other members of the population.

Social Display



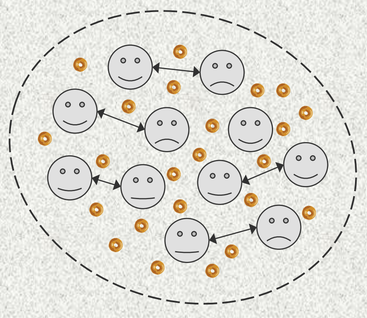
Signals allow for **coordination** by informing action opportunities to others, allowing groups of agents to **perform better as a whole** (*e.g.*, cooperative hunting, gathering, breeding, etc.).

Signals carry information about the **sender's observed state** such that it may change the **receiver's actions**.

Repeated interactions between signaling individuals make each **adjust its own strategy** towards maximizing its payoff given some fitness function.

Competitive Multiagent Learning

Populations of K agents, food exists for $K \times P_{eat\_max}$.



Partially-Observable Markov Game (POMG)

Each agent $k \in K$ is paired with another agent, ℓ according to $P_{meet}(k,\ell)$

Partial state view (observation):

- Own satiation status, $U_k = \{\text{FULL}_k, \text{ALMOST\_HUNGRY}_k, \text{HUNGRY}_k\}$
- Other's status, $O_{k,\ell} = \{\text{NOTHING}, \text{OTHER\_HUNGRY}_\ell, \text{OTHER\_FULL}_\ell\}$

Select action according to current policy from

$$\mathcal{A}_k = \{Eat, Nothing\}$$

Intrinsically-motivated Reinforcement Learning (IMRL)

Agents receive a reward $r = \phi_{\text{soc}}\,\theta_{\text{soc}} + \phi_{\text{fit}}\,\theta_{\text{fit}}$ combining:

- Food-related, **extrinsic** reward:

$$\phi_{\text{fit},k}(t) = \begin{cases} 1/K_{eat}(t), & \text{if } U_k(t+1) = \text{FULL}_k \\ -1/K, & \text{if } U_k(t+1) = \text{HUNGRY}_k \\ 0, & otherwise \end{cases}$$

- Social, **intrinsic** reward

$$\phi_{\text{soc},k}(t) = \begin{cases} 1, & \text{if } U_\ell(t+1) \in \{\text{ALMOST\_HUNGRY}_\ell, \text{FULL}_\ell\} \\ 0, & \text{if } O_{k,\ell}(t) = \text{NOTHING} \\ -1, & \text{if } U_\ell(t+1) = \text{HUNGRY}_\ell \end{cases}$$

Population reward function according to $[\theta_{\text{soc},k}, \theta_{\text{fit},k}]$, represents preferences.

Each agent is a **self-interested, independent** learner, each following and learning its own policy so to maximize the reward r.

Agent performance according to:

$$f_k = \sum_{\tau} \phi_{\text{fit}_k}(\tau)$$

Population Fitness

Measure the population performance, selection paradigms:

- **Average** fitness: $f_{avg} = \frac{1}{K}\sum_{k=1}^{K} f_k$
- **Maximal** fitness: $f_{max} = \max_k f_k$
- **Std. dev.** Fitness: $f_{std\_dev} = -\sqrt{\frac{1}{K}\sum_{k=1}^{K}(f_k - f_{avg})^2}$

Experimental Procedure

Objectives

- Study complex dynamics arising from the agents actions.
- See conditions for the emergence of cooperation.
- How population fitness varies according to selection.

Sensitive analysis

- $\theta_{\text{soc}} = \{0, 0.2, 0.4, 0.6, 0.8\}$
- $\theta_{\text{fit}} = 1 - \theta_{\text{soc}}$
- $P_{meet} = \{0, 0.1, 0.2, \ldots, 1\}$
- $P_{eat\_max} = \{0.1, 0.2, \ldots, 1\}$

Experimental setup

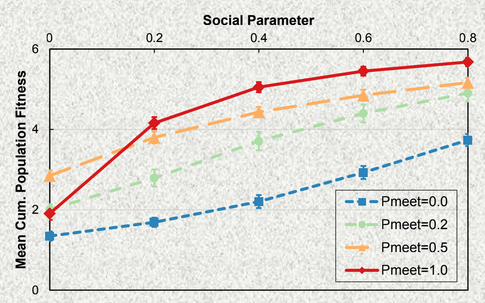
- 3300 populations of K = 100 agents
- 60 independent Monte-Carlo trials
- 1.000 time steps
- Policy-gradient algorithm

Results

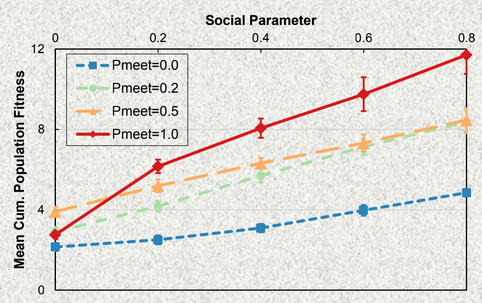
Social Influence on Population Fitness

Positive impact on the population fitness.

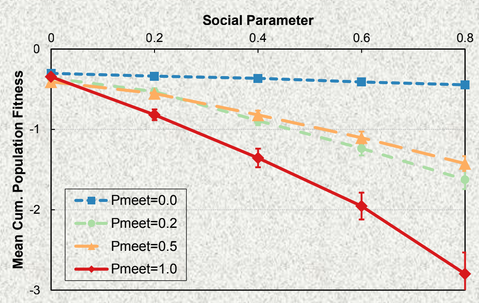
Agents achieve coordination, *i.e.* eat when necessary.



Avg. Pop. Fitness



Max. Pop. Fitness

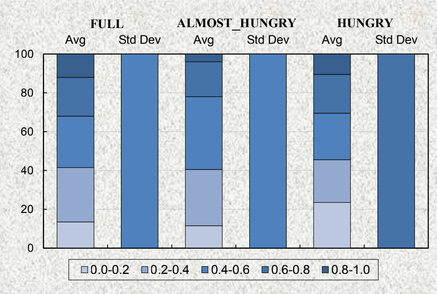


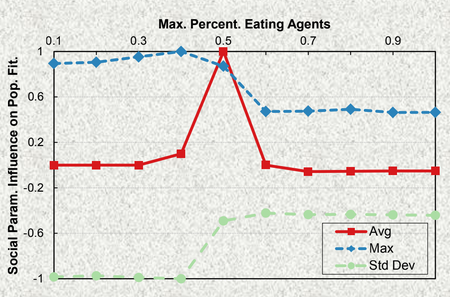
Std. Dev. Pop. Fitness

Social Inequalities

Fittest individuals, uneven populations, mixed strategies.

More uniform populations based on selfish agents.





Impact of Available Resources

Avg. fitness: social signals important when half of populations is allowed to eat.

Max. fitness: most important when few resources are available

Appearance of "popular" or "high-status" individuals.

http://gaips.inesc-id.pt/~psequeira

This work was supported by national funds through FCT under project PEst-OE/EEI/LA0021/2013.